A Many-Sorted Epistemic Logic for Chromatic Hypergraphs

Éric Goubault ✉ 🏠 iD
LIX, CNRS, École Polytechnique, IP-Paris, Palaiseau Cedex, France

Roman Kniazev ✉ 🏠 iD
LIX, CNRS, École Polytechnique, IP-Paris, Palaiseau Cedex, France
Université Paris-Saclay, ENS Paris-Saclay, CNRS, LMF, 91190, Gif-Sur-Yvette, France
Université Paris Cité, CNRS, IRIF, F-75013, Paris, France

Jérémy Ledent ✉ 🏠 iD
Université Paris Cité, CNRS, IRIF, F-75013, Paris, France

──────── **Abstract** ────────

We propose a many-sorted modal logic for reasoning about knowledge in multi-agent systems. Our logic introduces a clear distinction between participating agents and the environment. This allows to express local properties of agents and global properties of worlds in a uniform way, as well as to talk about the presence or absence of agents in a world. The logic subsumes the standard epistemic logic and is a conservative extension of it. The semantics is given in chromatic hypergraphs, a generalization of chromatic simplicial complexes, which were recently used to model knowledge in distributed systems. We show that the logic is sound and complete with respect to the intended semantics. We also show a further connection of chromatic hypergraphs with neighborhood frames.

1 Introduction

Epistemic logic is a modal logic that allows formal reasoning about knowledge and belief in multi-agent systems. At its core lies the *knowledge operator* denoted $K_a\varphi$, which means that "agent a knows that the formula φ holds", or simply "a knows φ" for short. Since the work of [16], epistemic logic has expanded in many directions, studying various operators such as common knowledge [11], distributed knowledge [6], as well as studying how knowledge evolves over time, as in public announcement logics [7], temporal epistemic logics [12], or dynamic epistemic logics [5]. A common feature of those approaches is that they rely on the classic "possible worlds" semantics of normal modal logics, based of Kripke structures. Indeed, a model for multi-agent epistemic logic $\mathsf{S5}_n$ usually consists of a set of *possible worlds* W, and for each agent a, an equivalence relation $\sim_a \subseteq W \times W$ called the *indistinguishability relation* of agent a. The intended semantics of such a model is that an agent a *knows* that a formula φ is true, then φ is true in every possible world that is indistinguishable from the real world for that agent. Formally, this means that the satisfaction relation is defined as follows, given a model M and a world $w \in W$ (thought of as the "real world"):

$$M, w \models K_a\varphi \quad \text{iff} \quad M, w' \models \varphi \text{ for every world } w' \in W \text{ such that } w \sim_a w' \tag{1}$$

A recent line of work [9, 4, 3, 10, 18, 8] has been developing a new notion of model for epistemic logic called *simplicial models*. This approach was closely inspired by connections with distributed computing, where simplicial complexes have been very successful in modeling various models of computation [13]. Compared to Hintikka's possible worlds semantics, this new approach represents a shift in perspective. Rather than focusing on the *worlds* (a.k.a. global states, in distributed computing terms) as the primary object of study, we instead focus on the agents' *points of view* about the world (a.k.a. local states). A possible world can then be defined as a set of *compatible* points of view, one for each agent.

Worlds and views.   To illustrate the distinction between possible worlds, and local views about the possible worlds, we use a classic example originally from distributed computing [14]. Since this paper is not concerned with distributed computing, we instead tell a story about a card game. Assume there are three agents, $\mathcal{A} = \{a, b, c\}$, and a deck of four cards, $\{1, 2, 3, 4\}$. We deal one card to each agent, and keep the remaining card hidden, so that each agent only knows its own card. Our goal is to model this (static) epistemic situation.

In the standard Kripke model semantics, the main step is to identify the possible worlds: each possible distribution of the cards constitutes a possible world. There are 24 such distributions, which we can denote by $W = \{123, 124, 132, 134, 142, 143, 213, 214, \ldots\}$, where for example the world "123" denotes the situation where agent a (resp. b, c) has received card number 1 (resp. 2, 3). To get a Kripke model, one must also define the indistinguishability relations for each agent. For instance, we have $123 \sim_a 132$. Indeed, from the point of view of agent a, who holds the same card number 1 in both of those worlds, these two worlds are indistinguishable. One can check that $\sim_a$ defined in this way is indeed an equivalence relation, with four equivalence classes, and similarly for $\sim_b$ and $\sim_c$.

In simplicial models, however, the central notion is that of *local view*. From the point of view of agent a, who sees only his own card, there are four possible situations: he can be given cards 1, 2, 3 or 4. We call these the *views* of a, denoted by $V_a = \{1_a, 2_a, 3_a, 4_a\}$. Similarly for the other two agents, we have $V_b = \{1_b, 2_b, 3_b, 4_b\}$ and $V_c = \{1_c, 2_c, 3_c, 4_c\}$. In order to get a model, one must moreover define which of those views are *compatible*. Indeed, a possible world can now be seen as a set of views, one for each agent. But not every combination is allowed: in our example, $\{1_a, 1_b, 1_c\}$ is not a compatible set of views, because at most one agent can be given the card number 1. As before, there are 24 sets of compatible views, corresponding to the 24 possible worlds: $\{1_a, 2_b, 3_c\}$, $\{1_a, 2_b, 4_c\}$, etc.

Surprisingly, shifting our focus from worlds to local views reveals an underlying geometric structure in this model. Indeed, the structure described above, with a set of views and a n-ary compatibility relations between those views, is known in mathematics as an *abstract simplicial complex*[1]. Simplicial complexes provide a combinatorial description of topological spaces. In the picture below, each local view is represented as a vertex, and each set of compatible views is represented as a triangle between the corresponding three vertices. The resulting shape is that of a triangulated torus, represented on the right in 3D view, and on the left as a flattened view with some repeated vertices. Notice that we use colors to indicate agent names, as we will do in the rest of the paper.

---

[1] Technically, we did not explicitly require the downward-closure property of a simplicial complex. We will come back to this later when we move on to hypergraph models.
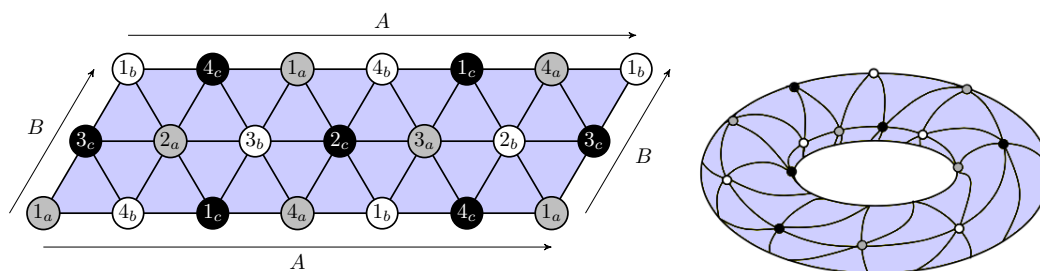
This idea led to the definition of a *pure simplicial model* in [9]. Readers unfamiliar with simplicial complexes need not worry about technical details, as we will be using hypergraphs instead in this paper. In the following, we assume the number of agents is $|\mathcal{A}| = n + 1$.

▶ **Definition 1** ([9]). *A pure simplicial model $M = (V, S, \chi, \ell)$ is given by:*
- *(V, S) is a pure simplicial complex of dimension n.*
- *$\chi : V \to \mathcal{A}$ assigns agents to vertices, s.t. every simplex has vertices of different colors.*
- *$\ell : V \to 2^{\mathsf{Ap}}$ assigns sets of atomic propositions to vertices.*

For this notion of model, the satisfaction relation $M, w \models \varphi$ is defined below, by induction on the structure of the formula φ. Note that now, the worlds w, w' are facets of the simplicial model, i.e., sets of compatible vertices. With that in mind, the previous condition (1) defining the knowledge operator in Kripke models, is replaced by condition (2). There, the condition $a \in \chi(w \cap w')$ can be read as "w and w' share an a-colored vertex", that is, agent a has the same local point of view in both worlds. Another discrepancy with respect to Kripke models is the case of atomic propositions: since the labelling ℓ decorates vertices, not worlds, with truth values of atomic propositions, we say that a proposition is true in a world when it is true in one of the vertices.

$$
\begin{aligned}
M, w &\models p & &\text{iff} & p &\in \bigcup_{v \in w} \ell(v) \\
M, w &\models \varphi \wedge \psi & &\text{iff} & M&, w \models \varphi \text{ and } M, w \models \psi \\
M, w &\models \neg \varphi & &\text{iff} & M&, w \not\models \varphi \\
M, w &\models K_a \varphi & &\text{iff} & M&, w' \models \varphi \text{ for every world } w' \text{ such that } a \in \chi(w \cap w') \quad (2)
\end{aligned}
$$

In Definition 1, the requirement to be "pure of dimension n" ensures that every agent is present in every world of the model. While this is a standard assumption in the epistemic logic literature, it is often not the case in distributed computing. Indeed, when we study computational models where processes may crash, one usually ends up with an impure simplicial model (see e.g. [15]).

Impure simplicial models. The idea of having a different set of agents (processes) in different possible worlds (executions) is ubiquitous in distributed computing. This situation might occur when a process crashes during the execution of a protocol; or simply when the set of participating agents is not known in advance (say, a server concurrently answering requests from various clients). In reference to the idea of crashed processes, and to be consistent with previous work on the topic [3, 10, 18, 8], we will say that an agent can be "alive" or "dead" in a given world. But note that for the time being, we only model static situations, so we could also say that agents can be "present" or "absent".

In the epistemic logic literature, the topic of non-participating agents has not been thoroughly studied. It was briefly considered, e.g. in [6], where it is called a "nonrigid set of agents". This formalism is not very handy to work with, as it simply consists of extra data

on top of the model indicating which agents are alive; and it is easy to circumvent the issue entirely by considering a special local state for crashed processes. However, in simplicial models, it is quite natural and straightforward to model worlds with non-participating agents: we simply remove from Definition 1 the requirement that the model must be pure. This simple idea led to a line of research on *impure simplicial models* [3, 10, 18, 8].

While it is clear which class of models we want to consider, we quickly run into issues when we try to define the semantics of epistemic logic formulas on these models. The crux of the matter is that we have to decide how to define the satisfaction relation $w \models K_a \varphi$, in a world w where agent a is dead. Two ways of dealing with this have been proposed.

The first approach, called the three-valued semantics [3, 18], claims that such a formula should be undefined in world w. Formulas can then be either true, false or undefined, hence the name "three-valued". Defining when formulas are well-defined is not trivial, as knowledge operators can be nested and evaluating the satisfaction relation will explore the various possible worlds of the model. So one first needs to inductively define a judgment $w \bowtie \varphi$, meaning that the formula φ is well-defined in world w; and then we can define the satisfaction relation $w \models \varphi$ on top of it. The resulting logic called $\mathsf{S5}_n^\bowtie$ is fully axiomatized in [18]. It is a non-normal modal logic, where axiom **K** : $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$, and even modus ponens, do not always hold. However, it retains the axiom of truth, $\mathsf{T} : K_a\varphi \to \varphi$.

The second approach, called the two-valued semantics [10, 8], is obtained by closely following the correspondence with Kripke models. As such, it yields a normal modal logic, where axiom K holds. However, the axiom of truth is lost, and only a weaker version remains, saying that alive agents are truthful: $\mathsf{alive}(a) \Rightarrow (K_a\varphi \Rightarrow \varphi)$. The reason for this is that dead agents know every formula: when agent a is dead in world w, the judgment $w \models K_a\varphi$ is vacuously true, because there is no world w' satisfying the condition in (2). The resulting logic is called $\mathsf{KB4}_n$; however, small design choices in how we define the models can result in additional axioms, as has been thoroughly investigated in [8].

This situation is quite unsatisfactory, as both approaches seem to have pros and cons, and there is no obvious way to tell which one might turn out to be more useful in practice. In this paper, we introduce a many-sorted logic that avoids entirely the problem of undefined formulas. Rather than being a third proposal, it subsumes and unifies the previous two approaches. But before we present the syntax of our logic, let us argue in favor of moving from simplicial complexes to hypergraphs.

From simplicial complexes to hypergraphs. In pure simplicial models (Definition 1), the vertices represent points of view of individual agents, and the facets (a.k.a. maximal simplexes) represent the possible worlds. There are also simplexes of lower dimension, but they do not seem to have a meaning. They are only here to preserve the geometric structure of the model: in order to have a triangle, we must also have the three edges of the triangle.

When we move on to impure simplicial models, we are now allowing worlds of different dimensions. So there is no good reason why only the facets of the model should represent worlds. In fact, as was shown in [10], having only facets as worlds results in some dubious axioms. The idea that all simplexes, not only the facets, could represent worlds was briefly discussed in [4], and later studied in [3, 18]. The idea was further generalized in [8], where models are equipped with additional data (called a covering) allowing to explicitly say which simplexes are worlds or not. Thus, models can be *minimal* (only the facets are worlds), *maximal* (all simplexes are worlds), or anything in-between. For instance in Figure 1 (left), the model has only one facet (the blue triangle) but three worlds: w_1 is the triangle itself, where three agents are alive; w_2 is an edge, where only two agents are alive; and w_3 is a vertex, where only one agent is alive.
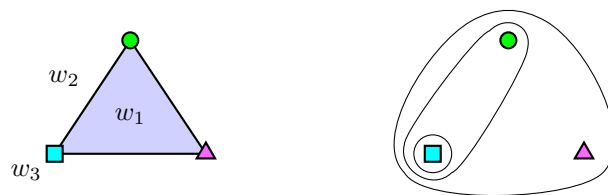
Figure 1 A simplicial model with three worlds, and the equivalent hypergraph representation.

Still, the simplicial model depicted on the left has some simplexes that play no epistemic role (the two edges and two vertices with no label). This introduces a strange discrepancy between simplexes that represent worlds, and simplexes that do not. Our proposal, which is quite mild from a technical standpoint, is to simply get rid of those extra simplexes. The resulting structure is called a hypergraph.

▶ **Definition 2.** *A (simple) hypergraph H is given by a pair (V, E), where V is a set of vertices, and $E \subseteq 2^V$ is a set of hyperedges (or just "edges", when clear from context).*

Note that the only difference between a hypergraph and a simplicial complex, is that the set E of hyperedges does not have to be downward-closed. A hypergraph is depicted in Figure 1 (right). It consists of three vertices $V = \{x, y, z\}$ (named from left to right), and three hyperedges $E = \{\{x\}, \{x, y\}, \{x, y, z\}\}$. Hyperedges are represented by a closed curve around the corresponding vertices, not unlike a Venn diagram. By moving from simplicial complexes to hypergraphs, we seem to lose the geometric intuition behind simplicial complexes. But it can be easily recovered by computing the downward closure of E: if (V, E) is a hypergraph, then $(V, {\downarrow}E)$ is a simplicial complex.

▶ Remark 3. For readers familiar with the hierarchy of models introduced in [8], hypergraph models fit nicely within that picture, as a strict subclass of epistemic covering models. They are not as general as simplicial sets, as there cannot be complex connectivity between hyperedges. They can be either minimal, or maximal, or in-between. Simple hypergraphs as in Definition 2 give rise to proper models, but as we will see later, we can also model non-proper behavior by allowing several hyperedges with the same sets of vertices.

Many-sorted epistemic logic. We now describe the main contribution of the paper. We propose a new syntax for epistemic logic formulas where agents can be dead or alive. The central idea is to introduce several sorts of formulas[2]: *world formulas* are to be interpreted in a world of the model, and *agent formulas* are interpreted in a point of view of a particular agent (i.e. in a vertex, for hypergraph models). We usually denote world formulas in capital letters Φ, Ψ, and agent formulas in lowercase, with a subscript indicating the name of an agent φ_a, ψ_a. Thus, our logic has $|\mathcal{A}| + 1$ sorts, one for each agent $a \in \mathcal{A}$, and an extra one for the world formulas.

A key observation is the following. Let us look again at the semantics of the knowledge operator for simplicial models, condition (2). It says that an agent a knows φ in a world w, precisely when φ holds in every world w' in which a has the same point of view. In fact, this definition does not refer to the "real world" w; it only refers to the point of view of a in this

---

[2] A familiar example of a many-sorted modal logic is CTL*, where the syntax is divided between state formulas and path formulas.

world! This suggests that the knowledge operator K_a should really be interpreted not in a world, but in a point of view of agent a. This gives us the following syntax for agent a's formulas, where p_a ranges over atomic propositions concerning agent a:

$$\varphi_a ::= p_a \mid \neg\varphi \mid \varphi \wedge \psi \mid \mathsf{K}_a\Phi$$

This gives us the first $|\mathcal{A}|$ sorts of our logic, with one sort of agent formulas for each $a \in \mathcal{A}$. Note that agent formulas can only talk about one particular agent. For instance, the expression $\mathsf{K}_a\Phi \wedge \mathsf{K}_b\Psi$ is not a syntactically valid formula. Also note that Φ is a world formula. We now explain the syntax of world formulas.

Since agents can be alive or dead, we cannot talk about the knowledge of a specific agent in world formulas, or we will run into the issue of how to define the knowledge of a dead agent. Instead, we introduce two new modal operators, E_a and A_a, that can test whether an agent exists in this world. As the names indicate, E_a has an existential flavor, while A_a is its universal counterpart. However, note that they are not binders: E_a should **not** be read as "there exists an agent a such that". Rather, the intuitive meaning of those operators is the following:

- $\mathsf{E}_a\varphi_a$: "there exists a point of view for agent a such that φ_a holds".
- $\mathsf{A}_a\varphi_a$: "for every point of view of agent a, φ_a holds".

The syntax of world formulas is as follows, where p_e denotes atomic propositions that do not talk about specific agents ('e' stands for *environment* here).

$$\Phi ::= p_e \mid \neg\Phi \mid \Phi \wedge \Psi \mid \mathsf{E}_a\varphi_a \mid \mathsf{A}_a\varphi_a \qquad \text{where } a \in \mathcal{A}$$

Note that, unlike with agent formulas, we have $|\mathcal{A}|$-many modal operators to choose from. So, for example, the following world formula is syntactically valid: $\mathsf{E}_a\mathsf{K}_a\Phi \wedge \mathsf{A}_b\mathsf{K}_b\Psi$. It is read "there exists a point of view of agent a where a knows Ψ, and for every point of view of agent b, b knows Ψ". Observe how whenever we want to talk about the knowledge of an agent, we are forced to explicitly quantify over the points of view of that agent. This avoids entirely the question of "undefined formulas", where we had to make an arbitrary decision about the meaning of knowledge for dead agents.

▶ Remark 4. It might seem strange to have a modality A_a quantifying over "all points of view of agent a", when there can be at most one point of view per agent in a given world. First, note that when agent a is absent in a world, the operator A_a is vacuously true, while E_a is false. So these two operators do behave differently in hypergraph models. Secondly, one could consider an extension of hypergraph models where an agent can have multiple points of view about the world. We briefly explore this idea in Section 4.3, where we relate it to the neighborhood semantics of epistemic logic [17].

In some sense, our logic can be viewed as a *refinement* of the usual knowledge operator K_a into two distinct operators: $\mathsf{A}_a\mathsf{K}_a\Phi$ is the 2-valued semantics of [10, 8], which is vacuously true when agent a is dead; while $\mathsf{E}_a\mathsf{K}_a\Phi$ is closer to the 3-valued semantics of [3, 18]

Related work. As we already explained, this work is directly related to the line of work on simplicial models [9], especially those that deal with impure simplicial complexes [3, 10, 18, 8]. Recently, a single-sorted epistemic logic on hypergraphs was considered in [2] to study weakly aggregative logics. There, vertices of hypergraphs are not colored as they are interpreted as worlds, so these models lie in-between epistemic frames and neighborhood frames. A framework that uses adjoint modalities was studied in [19] in the context of epistemic

modalities "agent is uncertain about" and "agent has information that". In interpreted systems [6], epistemic frames are generated by explicitly modeling the local states of agents and global states of the environment. However, at the level of syntax, no difference is made between local properties of the agents, and global properties of the environment.

Plan of the paper.   In Section 2, we start by describing the syntax of the logic 2CH and its semantics in chromatic hypergraphs. We give an axiomatization of the logic in Section 3, where we prove the completeness result in Theorem 23. In Section 4, we relate hypergraph models with partial epistemic models by showing an isomorphism of categories (Theorem 26). We use this equivalence to formulate a translation from KB4$_n$-formulas into 2CH-formulas, showing that the latter is a conservative extension of the former (Theorem 30).

2   Two-level chromatic hypergraph logic 2CH

We introduce a many-sorted refinement of multi-agent epistemic logic KB4$_n$ studied in [10]. Our logic has $|\mathcal{A}| + 1$ sorts of formulas: $|\mathcal{A}|$ sorts of *agent* formulas, one per agent, and one sort of *world* formulas. Since this logic is intended to be interpreted on chromatic hypergraphs (see Section 2.2), we name it the *2-level Chromatic Hypergraph Logic*, 2CH. The intended interpretation of agent formulas is to describe local information that belongs to a point of view of a specific agent. On the other hand, world formulas talk about properties of the environment, or world. As we will see in Section 3, this many-sorted logic, or two-level logic, embeds faithfully the logic KB4$_n$; but it also makes explicit (and not up to model interpretation as in [10]) the various choices about how much agents can observe each other's presence or absence.

2.1   Syntax

Fix a finite set $\mathcal{A}$ of agents. For each $a \in \mathcal{A}$, we have a set Ap_a of atomic propositions about agent a. We also have a set Ap_e of atomic propositions for the environment. We use lowercase letters with subscripts $\varphi_a, \psi_a, \ldots$ to denote agent formulas, and uppercase letters $\Phi, \Psi, \ldots$ for world formulas.

▶ **Definition 5.** *The language of the logic* 2CH *is defined as follows. For each agent $a \in \mathcal{A}$, there is a sort of agent formulas generated by the following grammar:*

$$\varphi_a ::= p_a \mid \neg\varphi \mid \varphi \wedge \psi \mid \widehat{\mathsf{K}}_a\Phi \qquad \text{where } p_a \in \mathsf{Ap}_a$$

The sort of world formulas is generated by the following grammar:

$$\Phi ::= p_e \mid \neg\Phi \mid \Phi \wedge \Psi \mid \mathsf{E}_a\varphi_a \qquad \text{where } a \in \mathcal{A} \text{ and } p_e \in \mathsf{Ap}_e$$

▶ Remark 6. We emphasize that this is truly an $(|\mathcal{A}| + 1)$-sorted logic, not a 2-sorted one. In particular, it is important to notice that the indexes in the operators $\widehat{\mathsf{K}}_a$ and E_a play a very different role. To illustrate that, let us fix two distinct agents $a, b \in \mathcal{A}$. To write an a-sorted agent formula φ_a, we are only allowed to refer to agent a's atomic propositions; and we are only allowed to use the modal operator $\widehat{\mathsf{K}}_a$, with index a. No reference to agent b is allowed. Similarly, an agent formula φ_b of sort b is not allowed to refer to agent a. Crucially, the following formula is ill-formed: $\widehat{\mathsf{K}}_a\Phi \wedge \widehat{\mathsf{K}}_b\Psi$. It cannot be generated by the grammar φ_a of a-sorted formulas, nor by the grammar φ_b of b-sorted formulas. Thus, it is not a valid formula in our logic. On the other hand, the sort of world formulas behaves much closer to the usual language of multi-agent epistemic logic. In a world formula, we are allowed to choose between $|\mathcal{A}|$-many modal operators E_a. Thus, a world formula of the form $\mathsf{E}_a\varphi_a \wedge \mathsf{E}_b\varphi_b$ is allowed.

We will use standard propositional connectives like true, $\vee$, $\Rightarrow$, defined as usual. There are also dual modalities: $\mathsf{K}_a\Phi := \neg\widehat{\mathsf{K}}_a\neg\Phi$, and $\mathsf{A}_a\varphi_a := \neg\mathsf{E}_a\neg\varphi_a$. Modalities are read as follows: $\widehat{\mathsf{K}}_a$ means "agent a considers possible that", K_a means "agent a knows that", E_a means "there exists a point of view of agent a such that", and A_a means "for all points of view of agent a". We call $\widehat{\mathsf{K}}_a$ and E_a *existential* modalities and K_a and A_a *universal* modalities.

2.2   Semantics

A *hypergraph* is a generalization of a graph, where instead of just edges between pairs of vertices, one has *hyperedges* that can connect multiple vertices at once. In the introduction, we defined *simple* hypergraphs (Definition 2), to explain the proximity with simplicial complexes. In fact, we will be slightly more general than that, and allow multiple hyperedges to have the same set of vertices; this will allow us to model non-proper behavior. Namely, a (non-simple) hypergraph H is a triple (V, E, P), where V is the set of vertices, E is the set of hyperedges, and $P : E \to 2^V$ assigns to each hyperedge a set of vertices.

In the context of multi-agent systems, we need moreover to consider *chromatic* hypergraphs, where each vertex is assigned an agent name. This could be done by adding an extra piece of data $\chi : V \to \mathcal{A}$, as in chromatic simplicial complexes [9]. Instead, we tweak the definition a little bit in order to make explicit the set of vertices assigned to each individual agent. This is similar to the definition of chromatic semi-simplicial sets in [8].

▶ **Definition 7.** *A chromatic hypergraph H is a tuple $(E, \{V_a\}_{a\in\mathcal{A}}, \{\mathsf{pr}_a\}_{a\in\mathcal{A}})$, where:*
- E *is a set of hyperedges,*
- *for all $a \in \mathcal{A}$, V_a is the set of* views *of agent a,*
- *for each agent $a \in \mathcal{A}$, $\mathsf{pr}_a : E \to V_a$ is a surjective partial function. Additionally, we require that for each $e \in E$, $\mathsf{pr}_a(e)$ is defined for at least one $a \in \mathcal{A}$.*

Indeed, defining $V = \bigcup_{a\in\mathcal{A}} V_a$ as the total set of vertices of a chromatic hypergraph, and $P(e) = \{\mathsf{pr}_a(e) \mid a \in \mathcal{A}\} \subseteq 2^V$, we can view H as a regular hypergraph. We will use the words *view* and *vertex* interchangeably, as well as *world* and *hyperedge*. Given a world $e \in E$, when $\mathsf{pr}_a(e)$ is undefined, we say that agent a is *dead* in e. Otherwise, a is *alive* in e, and we call $\mathsf{pr}_a(e)$ the view of a in e. Moreover, when $\mathsf{pr}_a(e) = v$, we say that v belongs to e, or that e contains v, and occasionally write $v \in_a e$. If a hyperedge consists of views $v_0, \ldots v_{n-1}$ then we say that these views are *compatible*. Let us explain each condition imposed on pr_a.
- pr_a is surjective: every view belongs to at least one world.
- pr_a is partial: not all agents are required to be alive in a world.
- pr_a is functional: every world contains at most one view of each agent.
- $\mathsf{pr}_a(e)$ is defined for at least one $a \in \mathcal{A}$: every world contains at least one alive agent.

▶ Remark 8. In chromatic hypergraphs, views and worlds are of equal importance: they are both described explicitly in the sets E and $(V_a)_{a\in\mathcal{A}}$. This contrasts with Kripke structures, where the set of worlds is explicit, but the views are implicit in the indistinguishability relations. Recent work on simplicial models has highlighted the importance of considering the views as a first-class notion. Here, we take this idea one step further with our two-level syntax, where world formulas talk about properties of the world, and agent formulas talk about properties of an agent's point of view. We also avoid choosing between *local* vs. *global* atomic propositions (in the sense of [9]): agent formulas can talk about (local) properties of agent $a \in \mathcal{A}$ using atoms in Ap_a, and world formulas can talk about (global) properties of the world using atoms in Ap_e.

▶ **Example 9.** Three examples of chromatic hypergraphs are depicted in Figure 2. The three agents a, b, c are represented as colored shapes ●, ■, ▲, respectively. In all three hypergraphs, there is only one view per agent.

On the leftmost figure, all possible combinations of views are compatible, giving 7 hyperedges. Intuitively, in this situation, the agents do not know whether other agents exist. A scenario like this is standard in distributed computing, where agents are processes, and they do not know whether other processes are concurrently running.

The hypergraph depicted in the middle has a single hyperedge containing the three views. In this situation, all agents know that everyone is alive, as it is the only possible world. This represents a scenario where every agent has guarantees that the other two agents are running.

The rightmost figure is a hypergraph with three hyperedges, each of which contains two views. It represents a situation where the points of views are pairwise compatible, but not all three of them are compatible. That is, there is no possible world that realizes all three of them at once. This could model a scenario where each agent receives a message from one of the other two agents, but they do not know who sent the message. Another interpretation might be in a quantum setting, as an example of *contextuality* [1].
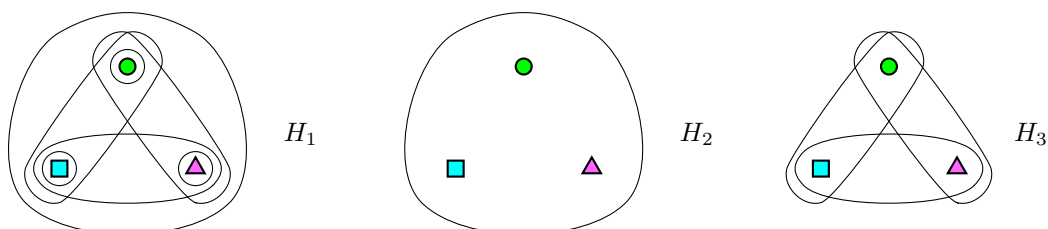
Figure 2 Three examples of chromatic hypergraphs H_1, H_2, H_3.

▶ **Definition 10.** *A chromatic hypergraph model is a tuple $(H, \{\ell_a\}_{a \in \mathcal{A}}, \ell_e)$, where H is a chromatic hypergraph, and $\ell_a : \mathsf{Ap}_a \to P(V_a)$, $\ell_e : \mathsf{Ap}_e \to P(E)$ are valuation functions.*

From now on, we sometimes omit the adjective "chromatic" when clear from context.

We can now define the semantics of 2CH formulas with respect to chromatic hypergraph models. Given a hypergraph model H, the satisfaction relations are defined for every sort by mutual induction. As expected, world formulas are interpreted in a world $e \in E$, and agent formulas are interpreted in a point of view $v \in V_a$ of that agent.

$H, v \models_a p_a$	iff	$v \in \ell_a(p_a)$		$H, e \models_e p_e$	iff	$e \in \ell_e(p_e)$
$H, v \models_a \neg\varphi$	iff	$H, v \not\models_a \varphi$		$H, e \models_e \neg\Phi$	iff	$H, e \not\models_e \Phi$
$H, v \models_a \varphi \wedge \psi$	iff	$H, v \models_a \varphi$ and $H, v \models_a \psi$		$H, e \models_e \Phi \wedge \Psi$	iff	$H, e \models_e \Phi$ and $H, e \models_e \Psi$
$H, v \models_a \mathsf{K}_a\Phi$	iff	$H, e \models_e \Phi$ for some $e \in E$		$H, e \models_e \mathsf{E}_a\varphi$	iff	$H, v \models_a \varphi$ for some $v \in V_a$
		such that $\mathsf{pr}_a(e) = v$				such that $\mathsf{pr}_a(e) = v$

2.3 Examples

Let us illustrate the semantics of 2CH on a few examples.

▶ **Example 11.** Consider again the three hypergraphs from Example 9, which we denote by H_1, H_2, H_3, from left to right. For now, we do not worry about atomic propositions: we just need to assume that the sets Ap_* are non-empty, in order to get the constant $\mathsf{true} = p \vee \neg p$. Also, recall that the agents are depicted as $a = ●$, $b = ■$, and $c = ▲$.

In the model H_1, let us consider the hyperedge $e = \{●, ■\}$. Then we have $H_1, e \models \mathsf{E}_a\mathsf{true}$ and $H_1, e \models \mathsf{E}_b\mathsf{true}$, but $H_1, e \not\models \mathsf{E}_c\mathsf{true}$. Indeed, the edge e does not contain a point of view of agent c, i.e., $\mathsf{pr}_c(e)$ is undefined. In fact, the world formula "$\mathsf{E}_a\mathsf{true}$" is satisfied exactly in the worlds where agent a is alive. So let us write $\mathsf{alive}(a) := \mathsf{E}_a\mathsf{true}$, so that $H, e \models \mathsf{alive}(a)$ iff e contains a point of view of a, that is, iff $\mathsf{pr}_a(e)$ is defined.

We can now talk about whether agents know that other agents are alive. Let $v = ●$ be the (unique) point of view of a in the model. In model H_1, all combinations of alive and dead agents are possible. So as expected, $H_1, v \models \neg\mathsf{K}_a\mathsf{alive}(b)$. In model H_2 however, a knows that all agents are alive, since this is the only possible world: $H_2, v \models \mathsf{K}_a(\mathsf{alive}(b) \wedge \mathsf{alive}(c))$. Finally in model H_3 the situation is more complicated: a knows that another agent is alive, but does not know which one. Thus, $H_3, v \models \mathsf{K}_a(\mathsf{alive}(b) \vee \mathsf{alive}(c)) \wedge \neg\mathsf{K}_a\mathsf{alive}(b) \wedge \neg\mathsf{K}_a\mathsf{alive}(c)$.

▶ **Example 12** (2-agent binary input model with solo executions). As an example where atomic propositions play a role, we consider the situation where two agents are given a binary input value, either 0 or 1. Moreover, there can be solo executions (a.k.a. initial crash failures, in distributed computing), so that agents do not know if they are running alone or not. As before, the agents are depicted as $a = ●$ and $b = ■$. The sets of atomic propositions are $\mathsf{Ap}_a = \{0_a, 1_a\}$, $\mathsf{Ap}_b = \{0_b, 1_b\}$, and $\mathsf{Ap}_e = \{\mathsf{solo}\}$. The agent atomic propositions hold in the points of view indicated on the picture below. The environment atomic proposition solo holds in the four singleton hyperedges where only one agent is alive.
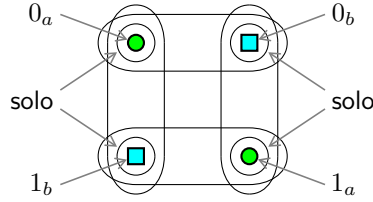
Let v be the top-left vertex, where agent a has input value 0. Then by definition $H, v \models 0_a$. Moreover, a does not know whether agent b is alive: $H, v \models \neg\mathsf{K}_a\mathsf{E}_b\mathsf{true}$, which we could also reformulate as $H, v \models \neg\mathsf{K}_a\mathsf{solo}$, that is, a does not know whether this is a solo execution. We can also say that a considers possible that b is alive with value 1: $H, v \models \widehat{\mathsf{K}}_a\mathsf{E}_b 1_b$. As a last example, we can express that a knows that if b is alive, its value is either 0 or 1: $H, v \models \mathsf{K}_a(\neg\mathsf{solo} \Rightarrow \mathsf{E}_b(0_b \vee 1_b))$. Alternatively, we could have used the operator A_b to express the same fact without a conditional: $H, v \models \mathsf{K}_a\mathsf{A}_b(0_b \vee 1_b)$.

2.4   Safe and unsafe knowledge

In traditional epistemic logics, formulas are interpreted in a world of the model. In the logic 2CH, to talk about the knowledge of an agent in a world, we first need to quantify over the points of view of this agent. Since there are two quantifiers E_a and A_a, we obtain two different knowledge operators on worlds, which we call *safe* and *unsafe* knowledge.

$$K_a^{\mathrm{safe}}\Phi := \mathsf{E}_a\mathsf{K}_a\Phi \qquad\qquad K_a^{\mathrm{unsafe}}\Phi := \mathsf{A}_a\mathsf{K}_a\Phi$$

These two operators only differ in the knowledge of dead agents. Indeed, given a world e and an agent a which is dead in e (i.e., $\mathsf{pr}_a(e)$ is undefined), we have $H, e \not\models K_a^{\mathrm{safe}}\Phi$ (dead agents know nothing), whereas $H, e \models K_a^{\mathrm{unsafe}}\Phi$ (dead agents know everything). However, when the agent a is alive in e, the two notions agree: $H, e \models K_a^{\mathrm{safe}}\Phi \iff H, e \models K_a^{\mathrm{unsafe}}\Phi$.

3　Axiomatics

The logic 2CH has all the usual inference rules of classical propositional logic (such as modus ponens) together with classical tautologies. It also has the following rules for modalities. We annotate the $\vdash$ symbol with the sort of the corresponding formula, $a \in \mathcal{A}$ for agent sorts and e for the world sort.

$$\frac{\vdash_e \Phi}{\vdash_a \mathsf{K}_a\Phi} \text{ Nec-a} \qquad \frac{\vdash_a \varphi}{\vdash_e \mathsf{A}_a\varphi} \text{ Nec-e} \qquad \frac{\vdash_e \Phi \to \Psi}{\vdash_a \heartsuit\Phi \to \heartsuit\Psi} \text{ RM} \qquad \frac{\vdash_a \varphi \to \psi}{\vdash_e \heartsuit\varphi \to \heartsuit\psi} \text{ RM'}$$

$$\frac{\vdash_e \Phi \to \mathsf{A}_a\psi}{\vdash_a \widehat{\mathsf{K}}_a\Phi \to \psi} \text{ Adj-1} \qquad\qquad \frac{\vdash_a \varphi \to \mathsf{K}_a\Psi}{\vdash_e \mathsf{E}_a\varphi \to \Psi} \text{ Adj-2}$$

where $\heartsuit \in \{\mathsf{E}_a, \mathsf{A}_a\}$ for rule RM, and $\heartsuit \in \{\widehat{\mathsf{K}}_a, \mathsf{K}_a\}$ for rule RM'. The first two rules are necessitation rules. The next two rules are monotonicity rules. The last two rules are called adjunction rules, and the double horizontal bar indicates that they go in both directions: top-to-bottom and bottom-to-top. They describe the interaction between the two pairs of modalities. Finally, we have the following axiom schemes for modalities:

- $\vdash_a \varphi \to \widehat{\mathsf{K}}_a\mathsf{E}_a\varphi$: every point of view belongs to some world;
- $\vdash_a \widehat{\mathsf{K}}_a\mathsf{E}_a\varphi \to \varphi$: every world has at most one point of view of a given agent;
- $\vdash_e \bigvee_{a\in\mathcal{A}} \mathsf{E}_a\mathsf{true}$: every world contains at least one point of view.

As we will see, the axiom schemes for modalities correspond to the defining properties of chromatic hypergraphs. Therefore, we will call the first axiom scheme *surjecitivity*, the second one *functionality*, and the third one *non-emptiness*.

▶ **Proposition 13.** *The logic* 2CH *is sound with respect to chromatic hypergraphs.*

3.1　Playing with the logic 2CH

In this logic, we can show that the universal modalities satisfy axiom K:

▶ **Proposition 14.** *For* $\heartsuit \in \{\mathsf{A}_a, \mathsf{K}_a\}$*, the axiom* $\mathsf{K}_\heartsuit$ *holds:* $\heartsuit(\varphi \to \psi) \to (\heartsuit\varphi \to \heartsuit\psi)$*.*

We give a list of useful formulas that are derivable in the logic 2CH.

▶ **Proposition 15.** *The following statements are derivable in the two-level logic* 2CH*:*

1. $\mathsf{E}_a\varphi \to \mathsf{A}_a\varphi$
2. $\mathsf{K}_a\Phi \to \mathsf{K}_a\mathsf{E}_a\mathsf{K}_a\Phi$
3. $\mathsf{E}_a\mathsf{K}_a\Phi \to \Phi$
4. $\Phi \to \mathsf{A}_a\widehat{\mathsf{K}}_a\Phi$
5. $\varphi \to \mathsf{K}_a\underline{\mathsf{E}}_a\varphi$
6. $\mathsf{K}_a\Phi \to \widehat{\mathsf{K}}_a\Phi$

Here is an intuitive explanation of these formulas:
1. If φ holds in a point of view of a, then it holds in all points of view of a. That is, there can be at most one point of view per world.
2. This is a form of positive introspection: if an agent knows something, then he knows that he knows it. However, we can put stress on the last "he", that is, "if he knows something, then he knows that *he* knows it".
3. This is a form of veracity: if a fact about a world is known by someone, then it is true.
4. If a certain fact holds in a world, then the agents in this world consider this fact possible. This is related to negative introspection.
5. Agents know the local facts about themselves.
6. This is the usual modal axiom D. It reflects the fact that every point of view belongs to at least one world.

In [9], the use of local atomic propositions leads to a so-called assumption of locality, $K_a(p_{a,x}) \vee K_a(\neg p_{a,x})$. In hypergraph models, valuations are local by construction:

▶ **Proposition 16.** *For any $p_a \in \mathsf{Ap}_a$, the formula $\mathsf{K}_a\mathsf{E}_a p_a \vee \mathsf{K}_a\mathsf{E}_a \neg p_a$ is derivable in* 2CH.

Proof. By adjunction rules, we have $\widehat{\mathsf{K}}_a\mathsf{A}_a p_a \to p_a$ and $p_a \to \mathsf{K}_a\mathsf{E}_a p_a$. By cut rule, we have $\widehat{\mathsf{K}}_a\mathsf{A}_a p_a \to \mathsf{K}_a\mathsf{E}_a p_a$. This is equivalent to $\mathsf{K}_a\mathsf{E}_a p_a \vee \mathsf{K}_a\mathsf{E}_a \neg p_a$ by propositional logic.   ◀

Note that the above proof does not use the fact that p_a is atomic, so in fact, an agent can decide *any* formula about itself. That is, for any φ, $\mathsf{K}_a\mathsf{E}_a\varphi \vee \mathsf{K}_a\mathsf{E}_a\neg\varphi$ is derivable.

3.2   Completeness

The proof of completeness uses the standard canonicity argument, extended to the many-sorted case. There is nothing surprising: the canonical model consists of the maximal consistent sets of formulas, now of several sorts. These sets satisfy standard properties, and together form a chromatic hypergraph.

▶ **Definition 17.** *A set of formulas S_* of sort $*$ is* inconsistent *if* false *can be derived from it. Otherwise, it is called consistent. A consistent set of formulas is* maximal *if it is not a proper subset of any other consistent set of formulas.*

▶ **Proposition 18.** *Maximal consistent sets (MCS) of formulas of sort $*$ are closed under modus ponens: for any formula ξ of sort $*$, either $\xi \in S_*$ or $\neg\xi \in S_*$, and every (non-maximal) consistent set of formulas of sort $*$ is contained in a maximal consistent set of formulas of sort $*$.*

▶ **Definition 19.** *The canonical hypergraph model consists of the following:*
- *the set of hyperedges is $E = \{S_e \mid S_e$ is a MCS of sort $e\}$;*
- *for each agent a, the set of vertices is $V_a = \{S_a \mid S_a$ is a MCS of sort $a\}$;*

together with relations $R_a \subseteq E \times V_a$ for every agent a, which are defined as follows: $S_E R_a S_a$ iff for all formulas Φ, if $\Phi \in S_E$, then $\widehat{\mathsf{K}}_a\Phi \in S_a$. The valuation function is defined by: $\ell_*(p_*) = \{S_* \mid p_* \in S_*\}$.

▶ **Proposition 20.** *In the canonical model, if $S_E R_a S_a$, if $\mathsf{K}_a\Phi \in S_a$, then $\Phi \in S_E$.*

▶ **Lemma 21.** *The canonical model is a chromatic hypergraph, that is, R_a is surjective, functional, and for any S_E there is S_a such that $S_E R_a S_a$ for at least one a.*

Proof. First, suppose that $S_E R_a S_a$, $S_E R_a S_a'$, and $S_a \neq S_a'$. It means that there is a formula φ which is in S_a, but is not in S_a'. By adjoint axiom and modus ponens, $\mathsf{K}_a\mathsf{E}_a\varphi$ is in S_a. By Proposition 20, $\mathsf{E}_a\varphi$ is in S_E. Using the definition of the canonical model, $\widehat{\mathsf{K}}_a\mathsf{E}_a\varphi$ belongs to S_a'. From there, by functionality axiom and modus ponens, $\varphi \in S_a'$, which is a contradiction. Thus, $S_a = S_a'$.

Second, we need to show that in the canonical model every vertex belongs to a hyperedge. Assume this is not the case, that is, there is a vertex S_a that does not belong to any hyperedge. It means that there is no maximal consistent set of formulas that contains $S = \{\Phi \mid \mathsf{K}_a\Phi \in S_a\}$. In particular, it means that S is itself not consistent, that is, there is a finite set of formulas $\{\Phi_i\}$ such that $\bigwedge_i \Phi_i \to$ false is derivable. By applying necessitation, we get that $\mathsf{K}_a(\bigwedge_i \Phi_i \to$ false$)$ is derivable, thus belongs to S_a. As K_a distributes over conjunction, and every $\mathsf{A}_a\Phi_i$ is in S_a, we get that $\mathsf{K}_a \bigwedge_i \Phi_i$ is in S_a. Applying modus ponens, we get that $\mathsf{K}_a$false is in S_a. Using surjectivity axiom, we get that false is in S_a, that is S_a is not consistent, which is a contradiction.

Lastly, we need to show that every hyperedge contains some vertex. Suppose it is not the case, that is there is a hyperedge S_E that does not contain any vertex. It means that for every a, the set $\widehat{\mathsf{K}}_a S_E$ is not consistent. Thus, for all a, there is a finite set of formulas $\{\Phi_i^a\} \subset S_E$, such that $\bigwedge_i \widehat{\mathsf{K}}_a \Phi_i^a \to \mathsf{false}$ is derivable. We now show that this implies that S_E is not consistent. By applying necessitation, we get that $\mathsf{A}_a(\bigwedge_i \widehat{\mathsf{K}}_a \Phi_i^a \to \mathsf{false})$ is derivable for every a. By (K) and modus ponens, we derive $\mathsf{A}_a(\bigwedge \widehat{\mathsf{K}}_a \Phi_i^a) \to \mathsf{A}_a \mathsf{false}$. Combining them all together, we have that $\bigwedge_a \bigwedge_i (\mathsf{A}_a \widehat{\mathsf{K}}_a \Phi_i^a) \to \bigwedge_a \mathsf{A}_a \mathsf{false}$ is derivable too. The antecedent is in S_E because every Φ_i^a is in S_E and $\Phi \to \mathsf{A}_a \widehat{\mathsf{K}}_a \Phi$ is an axiom. Thus, $\bigwedge_a \mathsf{A}_a \mathsf{false}$ is in S_E, which means that S_E is not consistent since $\bigvee_a \mathsf{E}_a \mathsf{true}_a$ is an axiom, which is its negation. We have a contradiction, which means that every hyperedge contains some vertex. ◀

▶ **Lemma 22.** *In the canonical model, $S_* \models_* \xi$ iff $\xi \in S_*$.*

▶ **Theorem 23.** *The logic* 2CH *is complete with respect to chromatic hypergraph models.*

4 Links to related work

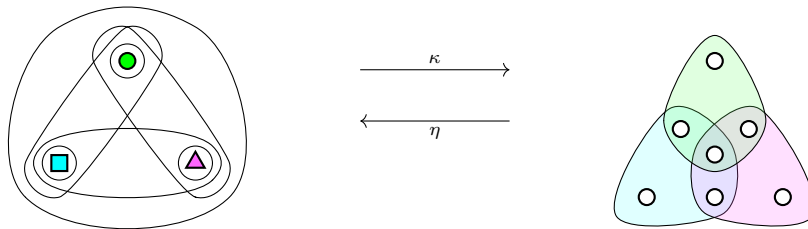### 4.1 Equivalence with partial epistemic frames

Let us recall first the definition of a partial epistemic frame, which has been one of the main models used in the study of epistemic logics such as $\mathsf{KB4}_n$ in [10]:

▶ **Definition 24.** *Given the set of agents $\mathcal{A}$, a partial epistemic frame $\mathcal{M}$ consists of a set of worlds M together with a family of partial equivalence relations $\{\sim_a\}_{a \in \mathcal{A}}$, such that for every $w \in M$, $w \sim_a w$ for at least one $a \in \mathcal{A}$. A morphism of partial epistemic frames is a function $f : M \to M'$ such that for every $a \in \mathcal{A}$ and $w, w' \in M$, $w \sim_a w'$ implies $f(w) \sim_a f(w')$.*

We can transform a partial epistemic frame into a chromatic hypergraph, and vice versa, using the following construction. Suppose we are given a partial epistemic frame $\mathcal{M}$. We construct a chromatic hypergraph $\eta(\mathcal{M})$ by setting $E = M$ and $V_a = M/_{\sim_a}$, that is the set of hyperedges is exactly the set of worlds, and the set of vertices of color a is the set of equivalence classes of $\sim_a$. We then set $\mathsf{pr}_a(w)$ to be $[w]_a$, that is the equivalence class of w under $\sim_a$. It is easy to check that this indeed defines a chromatic hypergraph.

Conversely, given a chromatic hypergraph H, we can construct a partial epistemic frame $\kappa(H)$. We set the set of worlds M to be equal to the set of hyperedges of H, and $e \sim_a e'$ if and only if e and e' share an a-colored vertex. This yields a partial equivalence relation.

These maps can be seen as the dual hypergraph construction: if $H = (V, E)$ is a (non-chromatic) hypergraph, then H^* is the hypergraph (E, V) where the hyperedges are the vertices of H and the vertices are the hyperedges of H. Partial epistemic frames can be seen as hypergraphs that have colored hyperedges which are defined by equivalence classes. The correspondence is exemplified in Figure 3.

Figure 3 Example of correspondence between chromatic hypergraphs and frames.

In fact, η and κ can be extended to morphisms of partial epistemic frames and chromatic hypergraphs, giving an equivalence of categories. First, we need to define the corresponding morphisms of chromatic hypergraphs:

▶ **Definition 25.** *A morphism of hypergraphs $f : H \to H'$ is a family of functions $f_a : V_a \to V'_a$ for each agent a, together with a function $f_e : E \to E'$ such that for all $a \in \mathcal{A}$, if $\mathsf{pr}_a(e) = v$ then $\mathsf{pr}'_a(f_e(e)) = f_a(v)$.*

▶ **Theorem 26.** *The category of partial epistemic frames is isomorphic to the category of chromatic hypergraphs. In particular, for any chromatic hypergraph H, $\eta(\kappa(H))$ is isomorphic to H, and for any partial epistemic frame $\mathcal{M}$, $\kappa(\eta(\mathcal{M}))$ is isomorphic to $\mathcal{M}$.*

In light of Theorem 26, chromatic hypergraphs and partial epistemic frames contain exactly the same information. So, in theory, we could have defined the semantics of 2CH in partial epistemic frames. However, this would be quite unnatural to do, since epistemic frames do not have a tangible notion of point of view: we would have to attach atomic propositions, and interpret agent formulas, in the equivalence classes of $\sim_a$.

Instead, we can still embed partial epistemic models (with only world atomic propositions) into a subclass of chromatic hypergraphs models, such that $\mathsf{Ap}_a = \varnothing$ for every agent. From this, we can extend Theorem 26 to work at the level of models:

▶ **Corollary 27.** *The category of partial epistemic models is isomorphic to the category of chromatic hypergraph models with empty sets of atomic propositions for agents.*

4.2   Translation from KB4$_n$ to 2CH

We can use the equivalence of epistemic frames and hypergraphs for showing how the logics KB4$_n$ and 2CH are related: we will show that 2CH is a conservative extension of KB4$_n$ + NE, where axiom NE ensures that there is an alive agent in each world (see [10] for further details). From semantics side, the worlds of epistemic frames are the hyperedges of hypergraphs, thus the formulas of KB4$_n$ are to be translated to the *world* formulas of 2CH. In particular, when defining the translation, we set the set of world atomic propositions to be the set of atomic propositions of KB4$_n$. The translation of formulas is defined recursively as follows:

$$\ulcorner p \urcorner := p \qquad \ulcorner \neg \Phi \urcorner := \neg \ulcorner \Phi \urcorner \qquad \ulcorner \Phi \wedge \Psi \urcorner := \ulcorner \Phi \urcorner \wedge \ulcorner \Psi \urcorner \qquad \ulcorner K_a \Phi \urcorner := \mathsf{A}_a \mathsf{K}_a \ulcorner \Phi \urcorner$$

Essentially, this translation interprets the knowledge operator of KB4$_n$ using the *unsafe knowledge* operator described in Section 2.4. So, if in a given world agent a is dead, $\ulcorner K_a \Phi \urcorner$ will be vacuously true.

▶ **Proposition 28.** *For a partial epistemic frame $\mathcal{M}$ and a formula Φ of KB4$_n$, $\mathcal{M}, w \models \Phi$ iff $\eta(\mathcal{M}), \eta(w) \models_e \ulcorner \Phi \urcorner$.*

Proof. We show the statement by induction on the structure of Φ. For atomic propositions, as well as boolean connectives, the proof is trivial. For the modality, we have: $\mathcal{M}, w \models K_a \Phi$ if and only if for all $w' \in M$ such that $w \sim_a w'$, $\mathcal{M}, w' \models \Phi$. By induction, this is equivalent to for all $w' \in M$ such that $w \sim_a w'$, $\eta(\mathcal{M}), \eta(w') \models \ulcorner \Phi \urcorner$. By definition of η, it is the same as for all hyperedges $e \in \eta(\mathcal{M})_E$ that share an a vertex with $\eta(w)$, $\eta(\mathcal{M}), e \models \ulcorner \Phi \urcorner$. This is equivalent to $\eta(\mathcal{M}), \eta(w) \models \ulcorner K_a \Phi \urcorner$. ◀

▶ **Corollary 29.** *Φ is valid in a partial epistemic frame $\mathcal{M}$ iff $\ulcorner \Phi \urcorner$ is valid in $\eta(\mathcal{M})$.*

Using this and Theorem 23, we can show that 2CH is a conservative extension of KB4$_n$:

▶ **Theorem 30.** *For every* $\mathsf{KB4}_n$*-formula* Φ, $\vdash_{\mathsf{KB4}_n+\mathsf{NE}} \Phi$ *if and only if* $\vdash_e \ulcorner\Phi\urcorner$.

Proof. By completeness, for $\mathsf{KB4}_n$ we have $\vdash_{\mathsf{KB4}_n+\mathsf{NE}} \Phi \Leftrightarrow \models \Phi$. By Corollary 29, we have $\models \Phi \Leftrightarrow \models_e \ulcorner\Phi\urcorner$. And by completeness for 2CH, we have $\models_e \ulcorner\Phi\urcorner \Leftrightarrow \vdash_e \ulcorner\Phi\urcorner$. ◀

As a corollary, we get that the combined modality $\mathsf{A}_a\mathsf{K}_a$ satisfies axioms K, B, and 4.

▶ **Remark 31.** Similarly, one can wonder which logic we would get if we translate formulas using instead the *safe knowledge* operator $K_a^{\mathrm{safe}}\Phi = \mathsf{E}_a\mathsf{K}_a\Phi$. First thing to note is that this does *not* yield the three-valued logic $\mathsf{S5}_n^{\bowtie}$ of [18][3]. We can show that the safe knowledge modality satisfies axioms K and T. However, this modality is not normal as the necessitation rule is not admissible: true is valid in every world, but it is not the case that $\mathsf{E}_a\mathsf{K}_a$true is valid in every world. Thus, we cannot derive $\mathsf{E}_a\mathsf{K}_a$true from true.

4.3 Correspondence with neighborhood frames

Chromatic hypergraphs require that every hyperedge contains at most one vertex of each color. So in every world, an agent can have either 0 or 1 point of view. But what happens if we drop this condition and allow agents to have multiple points of view about a given world? Technically, this can be achieved by replacing, in Definition 7, the partial function $\mathsf{pr}_a : E \to V_a$ by a relation $\mathsf{pr}_a \subseteq E \times V_a$, i.e., get rid of the *functionality* requirement.

This leads to an intriguing connection with neighborhood frames [17]. In this subsection we will not prove formal results, but rather give an intuition of the connection between the two notions, applying a similar construction as in Section 4.1.

Neighborhood frames generalize epistemic frames by allowing agents to have multiple points of view on the same world, which on the side of hypergraphs corresponds exactly to the situation when we allow hyperedges to contain multiple vertices of the same color:

▶ **Definition 32** ([17]). *A neighborhood frame is a pair* $\mathsf{M} = (S, \{N_a\}_{a\in\mathcal{A}})$, *where* S *is a set of states, and for every agent* $a \in \mathcal{A}$, N_a *is a function that assigns to every state* $s \in S$, *a set* $N_a(s) \subseteq 2^S$ *called the* a*-neighborhoods of* s.

An example of a situation where neighborhood frames are required is as follows. Suppose we have two processes, communicating through shared memory. The memory has two cells, and each cell stores a bit of information: 0 or 1. Processes are given access to memory cells arbitrarily, and both can be assigned the same cell. They know which cell is assigned to them, and they know the value that is stored in this cell, that is, they read the value of the cell. Therefore, a process can have two points of view on the same situation, depending on which cell it is given access to. Assume for example that the shared memory stores values $(0, 1)$. Process a, when assigned the first cell, knows that the memory stores 0, and when assigned the second cell, knows that the memory stores 1. So the set of possible states is $S = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. In state $(0, 1)$, the two possible points of view of process a are described by neighborhoods: $N_a((0, 1)) = \{\{(0, 0), (0, 1)\}, \{(0, 1), (1, 1)\}\}$

We can also make sense of this example using generalized chromatic hypergraphs, where we allow hyperedges that contain multiple points of view of the same agent. In that case, we can model our example as follows. We have two agents, a and b. There are four hyperedges corresponding to four possible states of the memory: $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. Each agent has four possible points of view, depending on which memory cell is

---

[3] A translation of $\mathsf{S5}_n^{\bowtie}$ into 2CH is possible, but it is more involved. One must first translate the well-definedness judgment, then set $\ulcorner K_a\Phi\urcorner := K_a^{\mathrm{safe}}(\text{well-defined}(\Phi) \Rightarrow \ulcorner\Phi\urcorner)$.

assigned (left or right), and which bit is read (0 or 1). Let us denote the vertices of agent a by $\{(a, m, b) \mid m \in \{L, R\}, b \in \{0, 1\}\}$, and similarly for agent b. Then, a vertex (a, m, b) belongs to a hyperedge (x_L, x_R) if and only if $x_m = b$.

Recall the duality construction of Figure 3, switching the role of vertices and hyperedges. In our case, the set of hyperedges becomes the set of states, and the set of vertices defines the neighborhood function. In our example, the hyperedge/state $s = (0, a)$ contains two vertices of agent a: $(a, L, 0)$ and $(a, R, 1)$. The first vertex corresponds to the a-neighborhood $\{(0, 0), (0, 1)\}$, which is the set of hyperedges containing this vertex. Similarly, the second vertex corresponds to the neighborhood $\{(0, 1), (1, 1)\}$. This recovers the set $N_a(s)$ of the corresponding neighborhood frame. Note that we do not obtain all neighborhood frames in this way, but only those in which a world belongs to all of its neighborhoods.

5 Conclusion

In this paper, we proposed a many-sorted modal logic for reasoning about knowledge in multi-agent systems that treat as first-class citizens both participating agents and the environment. This allowed us to reconcile the numerous logics and models of the literature, which indeed struggled with expressing coherent general global properties of worlds and local properties of agents. There are two main extensions that we are currently studying based on this work. First, having points of view of agents as first-class citizens, a combination of epistemic logics with temporal modalities allows us to provide a framework with greater emphasis on local action of agents, compared to e.g., DEL [5] or interpreted systems [6]. Secondly, a natural question arises as to whether we can reconcile chromatic hypergraphs, with chromatic (semi-)simplicial sets as studied in e.g., [8]. This would allow us to naturally extend our logic with a distributed knowledge operator.

References

1 Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, cohomology and paradox. In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic, CSL 2015*, volume 41 of *LIPIcs*, pages 211–228. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015. `doi:10.4230/LIPIcs.CSL.2015.211`.

2 Yifeng Ding, Jixin Liu, and Yanjing Wang. Someone knows that local reasoning on hypergraphs is a weakly aggregative modal logic. *Synthese*, 201(2):1–27, 2023. `doi:10.1007/s11229-022-04032-y`.

3 Hans van Ditmarsch. Wanted Dead or Alive: Epistemic Logic for Impure Simplicial Complexes. In Alexandra Silva, Renata Wassermann, and Ruy J. G. B. de Queiroz, editors, *Logic, Language, Information, and Computation - 27th International Workshop, WoLLIC 2021, Proceedings*, volume 13038 of *Lecture Notes in Computer Science*, pages 31–46. Springer, 2021. `doi:10.1007/978-3-030-88853-4_3`.

4 Hans van Ditmarsch, Éric Goubault, Jérémy Ledent, and Sergio Rajsbaum. Knowledge and simplicial complexes. In Björn Lundgren and Nancy Abigail Nuñez Hernández, editors, *Philosophy of Computing*, volume 143, pages 1–50, Cham, 2022. Springer International Publishing. `doi:10.1007/978-3-030-75267-5_1`.

5 Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*, volume 337 of *Synthese Library*. Springer, 2007. `doi:10.1007/978-1-4020-5839-4`.

6 Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge*. MIT Press, Cambridge, MA, USA, 2003.

7 Jelle Gerbrandy and Willem Groeneveld. Reasoning about information change. *Journal of Logic, Language and Information*, 6:147–169, 1997.

8 Éric Goubault, Roman Kniazev, Jérémy Ledent, and Sergio Rajsbaum. Semi-simplicial set models for distributed knowledge. In *38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–13, 2023. `doi:10.1109/LICS56636.2023.10175737`.

9 Éric Goubault, Jérémy Ledent, and Sergio Rajsbaum. A simplicial complex model for dynamic epistemic logic to study distributed task computability. *Inf. Comput.*, 278:104597, 2021. `doi:10.1016/j.ic.2020.104597`.

10 Éric Goubault, Jérémy Ledent, and Sergio Rajsbaum. A simplicial model for KB4: Epistemic logic with agents that may die. In *39th International Symposium on Theoretical Aspects of Computer Science, STACS 2022*, pages 33:1–33:20, 2022. `doi:10.4230/LIPIcs.STACS.2022.33`.

11 Joseph Y. Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *J. ACM*, 37(3):549–587, 1990. `doi:10.1145/79147.79161`.

12 Joseph Y. Halpern and Moshe Y. Vardi. The complexity of reasoning about knowledge and time. I. lower bounds. *J. Comput. Syst. Sci.*, 38(1):195–237, 1989. `doi:10.1016/0022-0000(89)90039-1`.

13 Maurice Herlihy, Dmitry Kozlov, and Sergio Rajsbaum. *Distributed Computing Through Combinatorial Topology.* Morgan Kaufmann, San Francisco, CA, USA, 2013.

14 Maurice Herlihy and Sergio Rajsbaum. Algebraic topology and distributed computing: A primer. In Jan van Leeuwen, editor, *Computer Science Today: Recent Trends and Developments*, volume 1000 of *Lecture Notes in Computer Science*, pages 203–217. Springer, 1995. `doi:10.1007/BFb0015245`.

15 Maurice Herlihy, Sergio Rajsbaum, and Mark R. Tuttle. An overview of synchronous message-passing and topology. *Electronic Notes in Theoretical Computer Science*, 39(2):1–17, 2000. `doi:10.1016/S1571-0661(05)01148-5`.

16 Jaakko Hintikka. *Knowledge and Belief.* Cornell University Press, 1962.

17 E. Pacuit. *Neighborhood Semantics for Modal Logic.* Short Textbooks in Logic. Springer International Publishing, 2017. URL: `https://books.google.co.uk/books?id=WK4-DwAAQBAJ`.

18 Rojo Fanamperana Randrianomentsoa, Hans van Ditmarsch, and Roman Kuznets. Impure simplicial complexes: Complete axiomatization. *CoRR*, abs/2211.13543, 2022. `doi:10.48550/arXiv.2211.13543`.

19 Mehrnoosh Sadrzadeh and Roy Dyckhoff. Positive logic with adjoint modalities: Proof theory, semantics and reasoning about information. *Electronic Notes in Theoretical Computer Science*, 249:451–470, 2009. Proceedings of the 25th Conference on Mathematical Foundations of Programming Semantics (MFPS 2009). `doi:10.1016/j.entcs.2009.07.102`.

A Proofs

A.1 Proof of Proposition 14

Proof. First, we show that universal modalities distribute over conjunction, that is $\heartsuit(\xi \wedge \eta) \leftrightarrow (\heartsuit\xi \wedge \heartsuit\eta)$. Left-to-right direction: we have that $(\xi \wedge \eta) \to \xi$ and $(\xi \wedge \eta) \to \eta$. Applying the RM rule, we get that $\heartsuit(\xi \wedge \eta) \to \heartsuit\xi$ and $\heartsuit(\xi \wedge \eta) \to \heartsuit\eta$. From this, the left-to-right direction follows. Right-to-left direction: denote the modality adjoint to $\heartsuit$ by $\spadesuit$, that is, if $\heartsuit = \mathsf{A}_a$ then $\spadesuit = \widehat{\mathsf{K}}_a$ and if $\heartsuit = \mathsf{K}_a$ then $\spadesuit = \mathsf{E}_a$. We have that $\spadesuit\heartsuit\xi \to \xi$ from $\heartsuit\xi \to \heartsuit\xi$ and the corresponding adjunction rule, similarly for η. From this we have that $\spadesuit\heartsuit\xi \wedge \spadesuit\heartsuit\eta \to \xi \wedge \eta$. By the same proof as in the left-to-right direction, we have that $\spadesuit(\heartsuit\xi \wedge \heartsuit\eta) \to \spadesuit\heartsuit\xi \wedge \spadesuit\heartsuit\eta$. Thus, we have that $\spadesuit(\heartsuit\xi \wedge \heartsuit\eta) \to \xi \wedge \eta$. Applying the adjunction rule, we get that $\heartsuit(\xi \wedge \eta) \to \heartsuit\xi \wedge \heartsuit\eta$, which is the right-to-left direction.

The fact that K follows from the distribution of universal modalities over conjunction is a standard proof: From $((\xi \to \eta) \wedge \xi) \to \eta$ by RM we have $\heartsuit((\xi \to \eta) \wedge \xi) \to \heartsuit\eta$. By distribution, we have $(\heartsuit(\xi \to \eta) \wedge \heartsuit\xi) \to \heartsuit((\xi \to \eta) \wedge \xi)$. Combining these two, we get $(\heartsuit(\xi \to \eta) \wedge \heartsuit\xi) \to \heartsuit\eta$, and thus $\heartsuit(\xi \to \eta) \to (\heartsuit\xi \to \heartsuit\eta)$. ◀

A.2   Proof of Proposition 15

Proof. Recall the list of formulas:

1. $\mathsf{E}_a\varphi \to \mathsf{A}_a\varphi$;
2. $\mathsf{K}_a\Phi \to \mathsf{K}_a\mathsf{E}_a\mathsf{K}_a\Phi$;
3. $\mathsf{E}_a\mathsf{K}_a\Phi \to \Phi$;
4. $\Phi \to \mathsf{A}_a\widehat{\mathsf{K}}_a\Phi$;
5. $\varphi \to \mathsf{K}_a\mathsf{E}_a\varphi$;
6. $\mathsf{K}_a\Phi \to \widehat{\mathsf{K}}_a\Phi$.

For the first formula, we just apply the adjunction rule to $\widehat{\mathsf{K}}_a\mathsf{E}_a\varphi \to \varphi$, which is an axiom. In order to show the second formula, just apply the adjunction rule to $\mathsf{E}_a\mathsf{K}_a\Phi \to \mathsf{E}_a\mathsf{K}_a\Phi$, which is a tautology. Formulas 3, 4 and 5 are derived from $\mathsf{K}_a\Phi \to \mathsf{K}_a\Phi$, $\widehat{\mathsf{K}}_a\Phi \to \widehat{\mathsf{K}}_a\Phi$, and $\mathsf{E}_a\varphi \to \mathsf{E}_a\varphi$ respectively by applying the adjunction rule. The last formula is shown as follows: from formulas 3 and 4 we have $\mathsf{E}_a\mathsf{K}_a\Phi \to \mathsf{A}_a\widehat{\mathsf{K}}_a\Phi$. Applying the adjunction rule, we get $\widehat{\mathsf{K}}_a\mathsf{E}_a\mathsf{K}_a\Phi \to \widehat{\mathsf{K}}_a\Phi$. We also have $\mathsf{K}_a\Phi \to \widehat{\mathsf{K}}_a\mathsf{E}_a\mathsf{K}_a\Phi$, which is an axiom. From the last two formulas, we get $\mathsf{K}_a\Phi \to \widehat{\mathsf{K}}_a\Phi$. ◀

A.3   Proof of Theorem 26

Proof. We just need to show how η and κ are extended to morphisms. Functoriality is then straightforward, and checking that $\eta(\kappa(H)) \simeq H$ and $\kappa(\eta(\mathcal{M})) \simeq \mathcal{M}$ is also straightforward. Let $f : H \to H'$ be a morphism of chromatic hypergraphs. Then $\eta(f) : \eta(H) \to \eta(H')$ just sends a world e to $f_E(e)$. This is indeed a morphism of partial epistemic frames: suppose two worlds e and e' in $\eta(H)$ are $\sim_a$-equivalent. It means that in H these two hyperedges share a vertex, and thus in H' the two hyperedges $f_E(e)$ and $f_E(e')$ share a vertex, and thus $f_E(e) \sim_a f_E(e')$.

Now let $g : \mathcal{M} \to \mathcal{M}'$ be a morphism of partial epistemic frames. Then $\kappa(g) : \kappa(\mathcal{M}) \to \kappa(\mathcal{M}')$ sends a hyperedge w to $g(w)$, thus $\kappa(g)_E$ is defined. We need to show that it induces a map on vertices, and that the condition for morphsisms is satisfied. As g preserves $\sim_a$, it induces a map on equivalence classes, which is exactly $\kappa(g)_a$. Let w be a hyperedge and v be a vertex in $\kappa(\mathcal{M})$, such that $\mathsf{pr}_a(w) = v$. It means that w belongs to the equivalence class corresponding to v in $\mathcal{M}$. Thus, g(w) belongs to the equivalence class corresponding to $g(v)$ in $\mathcal{M}'$, and thus $\mathsf{pr}_a(g(w)) = g(v)$. ◀