Probabilistic Multi-agent Only-believing

Qihui Feng
RWTH Aachen University
Aachen, Germany
feng@kbsg.rwth-aachen.de

Gerhard Lakemeyer
RWTH Aachen University
Aachen, Germany
gerhard@kbsg.rwth-aachen.de

ABSTRACT

Levesque introduced the notion of only-knowing to precisely capture the beliefs of a knowledge base. While numerous studies of only-knowing have emerged, such as the representation of probabilistic beliefs or reasoning about beliefs in an uncertain dynamical system, most remain confined to single-agent contexts. This limitation predominantly stems from an absence of a logical framework, which faithfully extends Levesque's intuition of only-knowing to multi-agent, probabilistic scenarios.

In this paper, we introduce a first-order logical account with probabilistic beliefs and only-believing of many agents. We demonstrate that the categorical fragment of our account forms a $KD45_n$ modal system, and the notion of belief has behavior following the laws of probability. We also show how an agent's beliefs and non-beliefs about the environment or other agents' beliefs are precisely captured through the modalities of only-believing, which paves the way to generalize tools for interfacing with symbolic, probabilistic knowledge bases. By way of example, we demonstrate how non-monotonic conclusions including default reasoning can be handled by our account.

KEYWORDS

knowledge representation; modal epistemic logic; only-believing; multi-agent systems; reasoning about probability

ACM Reference Format:
Qihui Feng and Gerhard Lakemeyer. 2024. Probabilistic Multi-agent Only-believing. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 9 pages.

1 INTRODUCTION

In multi-agent systems, where agents constantly interact or coordinate, reasoning about knowledge and belief is of interest for many applications. By gathering detailed information, not only from the environment but also from other agents' mental states, an agent can determine when communication with other agents is necessary, enhance collaboration on tasks, and optimize its strategy against other agents in scenarios where agents operate concurrently.

Among many symbolic approaches that represent an agent's knowledge and belief, the notion of only-knowing[23, 24] is particularly valuable: An agent's beliefs of a knowledge base (**KB**) are modelled in terms of only-knowing a collection of sentences, and sentences that are not logically entailed by the **KB** are taken to be precisely those not believed. In Levesque's logic of only-knowing $\mathcal{OL}$ [23, 24], the classical epistemic operator $\boldsymbol{K}$ is used for knowledge, and in addition, a modality $\boldsymbol{O}$ is used for only-knowing. To illustrate, $\boldsymbol{O}p$ means that p is all the agent knows. $\boldsymbol{O}p$ entails $\boldsymbol{K}p$, but it also entails $\neg\boldsymbol{K}q$ and $\neg\boldsymbol{K}\neg q$ for proposition q different from p. This is different from classical epistemic logic where $\boldsymbol{K}p$ entails neither $\neg\boldsymbol{K}q$ nor $\neg\boldsymbol{K}\neg q$. Furthermore, only-knowing also shows a close connection to autoepistemic logic[27] and can be applied to autoepistemic reasoning or default reasoning[8]. With some simple augmentations, non-monotonic conclusions can be reached without using meta-logical notions such as fixpoints or partial orders[20, 31]. For instance, let **KB** be a single sentence $\{Bird(tweety)\}$, namely "tweety is a bird", δ a sentence to express the default $\forall x.[Bird(x) \wedge \neg\boldsymbol{K}\neg Fly(x) \supset Fly(x)]$, i.e. any bird is assumed to be able to fly unless we know the opposite. Then we have the following non-monotonic properties in $\mathcal{OL}$: $\boldsymbol{O}(\mathbf{KB} \wedge \delta)$ entails $\boldsymbol{K}Fly(tweety)$, but $\boldsymbol{O}(\mathbf{KB} \wedge \delta \wedge \neg Fly(tweety))$ entails $\boldsymbol{K}\neg Fly(tweety)$. In other words, the initial belief "tweety can fly" is retracted when a new fact is added.

Numerous researches on single-agent only-knowing emerged: Lakemeyer and Levesque extended the notion of only-knowing to capture different forms of default reasoning [20], Belle et al. proposed the logic $\mathcal{OBL}$ to describe *only-believing* and admit knowledge bases with incomplete, probabilistic specifications [6]. There is also work on reasoning in dynamical domains [5, 19, 21]. It has also been shown how to capture an agent's belief after actions via only-believing and how to perform projection reasoning [25, 26].

Naturally, extending these works into the multi-agent scenario would result in an expressive account to represent and reason about the mental states of agents, and about how their minds change as a result of actions. As for non-monotonic reasoning, via multi-agent only-believing one could expect to express default assumption not only on the facts (e.g. the aforementioned "Birds can fly" example) but also on other agents' beliefs, like "It is assumed that agent 2 believes that the coin is fair". The expressiveness of such an account should be of interest for planning and decision-making in cooperative or concurrent games. However, research into multi-agent only-knowing has faced more obstacles than expected, and extensions in terms of probabilistic belief or belief after actions are rarely considered yet. Most studies are confined to characterizing only-knowing of categorical knowledge. Halpern and Lakemeyer attempted to handle the extension independently [13, 18]. However, these accounts make use of arbitrary Kripke structures and lose the simplicity of Levesque's semantics. Furthermore, each account has some undesirable properties respectively. These properties are avoided in a joint work from Halpern and Lakemeyer [14], but it forces us to have the semantic notion of validity directly in the language. For this reason, that proposal is not natural, and it is

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

matched with a proof theory that has a set of new axioms to deal with these new notions. Belle and Lakemeyer revisited the construction and proposed an account $\mathcal{ONL}_n$ with a natural possible-world semantics[3]. However, their notion of only-knowing didn't precisely capture the belief and non-belief for sentences of higher depth of nesting. For example, let $\boldsymbol{O}_1$ denotes the only-knowing of agent 1, $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ stand for the belief of agent 1 and 2 respectively, then $\boldsymbol{O}_1 p \supset \boldsymbol{K}_1(\boldsymbol{K}_2 q)$ and $\boldsymbol{O}_1 p \supset \neg\boldsymbol{K}_1(\boldsymbol{K}_2 q)$ are both satisfiable in their semantics.

The major difficulty is as follows: If we faithfully follow Levesque's principle of only-knowing, that the beliefs of an agent are precisely those following from its knowledge base, then in the multi-agent cases, only-knowing a sentence involves the belief and non-belief of sentences with an arbitrary depth of nesting beliefs. In a single-agent scenario, e.g. $\mathcal{OL}$, every formula specifies the agent's belief about the fact, or about its own belief, which we say to be of depth 1. In a multi-agent scenario, however, there are sentences describing an agent's belief about another agent's belief, which are of depth 2 at least. A sentence describing agent 1's belief about agent 2's belief about agent 3's belief is at least of depth 3. Such a belief nesting can be arbitrarily deep. If the depth is not restricted, constructing a structure which describes only-knowing up to all depths will be semantically difficult and will easily lead to a circular definition.

The study of belief with arbitrary depths of nesting is indeed meaningful, in particular for reasoning about common knowledge. However, we argue that it suffices to consider finite depths in many applications, where common knowledge or beliefs on an infinite set of sentences are not considered. We adapt the idea from Aucher and Belle [1] and use modalities of the form $\boldsymbol{O}_i^{(k)}$ to describe agent i's only-believing up to depth k. We focus on extending the notion of only-knowing to multi-agent, probabilistic cases and spare us from the troubles handling infinite nesting of beliefs.

The rest of the paper is organized as follows. We begin with introducing a new logic of multi-agent only-believing $\mathcal{OBL}_m$. Then we analyze the properties of the logic, in particular those exclusive for only-believing. Relation to other logical accounts will also be discussed. Finally, we briefly conclude our work.

2 THE LOGIC $\mathcal{OBL}_m$

We introduce the logic $\mathcal{OBL}_m$. It can be considered as a multi-agent extension of $\mathcal{OBL}$ [6] and a probabilistic extension of $\mathcal{ONL}_n$ [3].

2.1 The Language

The logic $\mathcal{OBL}_m$ is a first-order modal logic with equality. Let $Ag = \{1, \ldots, m\}$ denote a set of m agents. The vocabulary consists of (FO) *variables* and *predicate symbols*. For simplicity, function symbols are excluded. The language includes a countable set of *standard names* $\mathcal{N}$, which are syntactically treated like constants. This can be viewed as having fixed infinite domain closure axioms with the *unique name assumption*, which further allows FO quantification to be understood substitutionally. The set of rational numbers $\mathbb{Q}$ is included as a sub-sort of standard names. We call a predicate other than =, applied to first-order variables or standard names, an *atomic* formula. An atomic formula without variables is called a *ground atom*.

Standard FO connectives $\wedge, \neg, \forall$ and modalities $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_m$ are used to construct formulae. For each agent index i, $\boldsymbol{B}_i(\alpha\colon r)$ is read as "agent i believes α with degree r". To illustrate, a sentence $\boldsymbol{B}_1(Rain\colon 0.8)$ means agent 1 believes that there is an 80% chance of raining. In particular, we write $\boldsymbol{K}_i\alpha$ as an abbreviation of $\boldsymbol{B}_i(\alpha\colon 1)$ and read it as "agent i knows α".

To extend the notion of "only-knowing" to multi-agent systems, we start from the single-agent cases and inductively extend the language to represent only-believing of any depth. We use modality $\boldsymbol{O}_i^{(1)}$ to specify agent i's only-believing up to depth 1, which captures the beliefs and non-beliefs about facts and the agent's own beliefs. An example of $\boldsymbol{O}_i^{(1)}$-formula is $\boldsymbol{O}_i^{(1)}(\exists x.[P(x) \wedge \neg\boldsymbol{K}_i P(x)])$, which means "All agent i believes up to depth 1 is the existence of an unknown object x s.t. $P(x)$ holds." Now we go one step deeper: We use modality $\boldsymbol{O}_i^{(2)}$ to specify agent i's only-believing up to depth 2, which stands for the beliefs and non-beliefs about fact, about the agent's own beliefs and other agents' beliefs up to depth 1. An example is

$$\boldsymbol{O}_i^{(2)}(q \wedge \exists x.[\boldsymbol{K}_j(P(x)) \wedge \neg\boldsymbol{K}_i\boldsymbol{K}_j(P(x))])$$

with $i \neq j$. The difference between $\boldsymbol{O}_i^{(2)}$ and $\boldsymbol{O}_i^{(1)}$ is: While the former specifies agent i's beliefs about other agents' beliefs, the latter does not. For instance, $\boldsymbol{O}_i^{(2)}(p\colon 1)$ implies that agent i knows nothing about what other agents know or believe, e.g. $\boldsymbol{O}_i^{(2)}(p\colon 1)$ entails $\neg\boldsymbol{K}_i\boldsymbol{K}_j q$ for any proposition q, but $\boldsymbol{O}_i^{(1)}(p\colon 1)$ entails neither $\boldsymbol{K}_i\boldsymbol{K}_j q$ nor $\neg\boldsymbol{K}_i\boldsymbol{K}_j q$.

We include modalities $\boldsymbol{O}_1^{(k)}, \ldots \boldsymbol{O}_m^{(k)}$ with natural number $k > 0$. For each k, $\boldsymbol{O}_i^{(k+1)}$ specifies agent i's only-believing about other agents' beliefs up to depth k, and formula $\boldsymbol{O}_i^{(k)}(\alpha_1\colon r_1, \ldots, \alpha_l\colon r_l)$ is read as "All agent i believes up to depth k is: α_1 with degree r_1, ..., and α_l with degree r_l". We write $\boldsymbol{O}_i^{(k)}\alpha$ to mean $\boldsymbol{O}_i^{(k)}(\alpha\colon 1)$.

Definition 1 (i-depth). *For $i \in Ag$, the i-depth of a formula α, written as $dep[\alpha, i]$, is inductively defined as*

- $dep[\alpha, i] = 0$ *for atomic formula* α, $dep[t_1 = t_2, i] = 0$
- $dep[\neg\alpha, i] = dep[\alpha, i]$
- $dep[\alpha \wedge \beta, i] = max(dep[\alpha, i], dep[\beta, i])$
- $dep[\forall x.\alpha, i] = dep[\alpha, i]$
- $dep[\boldsymbol{B}_i(\alpha\colon r), i] = \max\{\max\{dep[\alpha, j] \mid j \neq i\}+1, dep[\alpha, i]\}$
- $dep[\boldsymbol{B}_j(\alpha\colon r), i] = 0$ *for* $j \neq i$
- $dep[\boldsymbol{O}_i^{(k)}(\alpha_1\colon r_1, \ldots, \alpha_l\colon r_l), i] = k$
- $dep[\boldsymbol{O}_j^{(k)}(\alpha_1\colon r_1, \ldots, \alpha_l\colon r_l), i] = 0$ *for* $j \neq i$

For a $\boldsymbol{B}$-formula, the depth increases only when a nesting of different agents' beliefs occurs. For an $\boldsymbol{O}$-formula, the depth is given by the superscript k.

Example 1. *Let α denote the formula $(\exists r.r > 0.5 \wedge \boldsymbol{B}_2(p\colon r))$. Consider the formula $\boldsymbol{K}_1\alpha$, namely, agent 1 knows that p is believed by agent 2 with a degree greater than 0.5. Then*

$$\begin{aligned} dep[\boldsymbol{K}_1\alpha, 1] &= \max\{dep[\alpha, 1], 1 + dep[\alpha, 2]\} \\ &= \max\{0, 2\} = 2 \\ dep[\boldsymbol{K}_1\alpha, 2] &= 0 \end{aligned}$$

i.e. $\boldsymbol{K}_1\alpha$ has 1-depth 2 and 2-depth 0.

The set of well-formed formulae (wffs) is the smallest set including:

- any atomic formulae,
- $t = t'$ where t and t' are variables or standard names,
- if α, β are formulae, then $\alpha \wedge \beta$, $\neg\alpha$, $\forall x.\alpha$ are formulae
- if α is a formula, then for any $i \in \{1, \ldots, m\}$ and number r, $\boldsymbol{B}_i(\alpha\colon r)$ is a formula.
- for $j \in \{1, \ldots, l\}$, $\boldsymbol{B}_i(\alpha_j\colon r_j)$ is a formula with i-depth k_j, $k \geq \max\{k_1, \ldots, k_l\}$, then $\boldsymbol{O}_i^{(k)}(\alpha_1\colon r_1, \ldots, \alpha_l\colon r_l)$ is a formula.

An example of illegal formulae is $\boldsymbol{O}_1^{(1)}(p \wedge \boldsymbol{K}_2 q\colon 0.8)$, where the sentence being believed is of 1-depth 2. Only-believing this sentence requires at least $k = 2$. For the rest of the paper, we only consider wffs and proper superscript k.

A formula with no free variables is called a *sentence*. A formula not mentioning modalities is said to be *objective*. A formula, where all predicate symbols appear within the scope of a modal operator is called *subjective*. Given $i \in Ag$, a formula is called *i-objective* if modal operators $\boldsymbol{B}_i$ and $\boldsymbol{O}_i^{(k)}$ only appear within the scope of $\boldsymbol{B}_j$ or $\boldsymbol{O}_j^{(k')}$ where $j \neq i$. For instance, $p \wedge \boldsymbol{O}_2^{(3)}(q \wedge \boldsymbol{K}_1 q)$ is 1-objective since $\boldsymbol{K}_1$ only appears in the scope of $\boldsymbol{O}_2^{(3)}$, namely, knowing p is agent 2's conjecture of agent 1's mental state and not necessarily to be the actual belief of agent 1. Inversely, a formula is called *i-subjective* if any predicates or modal operators $\boldsymbol{B}_j$ and $\boldsymbol{O}_j^{(k)}$ s.t. $j \neq i$ only appear within the scope of operators $\boldsymbol{B}_i$ or $\boldsymbol{O}_i^{(k')}$, e.g. $\neg\boldsymbol{K}_1 p \wedge \boldsymbol{B}_1(p \wedge \boldsymbol{K}_2 q\colon 0.5)$. We use True to denote an objective tautology $\forall x.(x = x)$ and use False to denote its negation.

2.2 The Semantics

The semantics of $\mathcal{OBL}_m$ is given in terms of possible worlds, where a *world* is a set of ground atoms considered as true. The set of all worlds is denoted as $\mathcal{W}$.

In the logic $\mathcal{OBL}$, a distribution d is a function from $\mathcal{W}$ to the set of non-negative real numbers $\mathbb{R}^{\geq 0}$ and an epistemic state is defined as a set of distributions.[1] We extend the notions into multi-agent cases: A distribution d defined above describes an agent's weighting to each possible world. For an agent i, to describe agent i's belief about other agents' beliefs on the actual world, the weight assigned to world w, namely $d(w)$, is partitioned and reassigned to tuples of form $(w, e_1, \ldots, e_{m-1})$ where each e_j is an $\mathcal{OBL}$ epistemic state describing the belief of an agent other than i. Formally, we have the following definition:

Definition 2 (k-distribution). *For $k \geq 1$, a k-distribution, written as d^k, is inductively defined as:*

- $d^1 : (\mathcal{W} \times \underbrace{\{\varnothing\} \times \cdots \times \{\varnothing\}}_{m-1}) \to \mathbb{R}^{\geq 0}$,
- $d^{k+1} : (\mathcal{W} \times \underbrace{\mathcal{E}^k \times \cdots \times \mathcal{E}^k}_{m-1}) \to \mathbb{R}^{\geq 0}$,

where $\mathcal{E}^k = 2^{\mathcal{D}^k}$, and $\mathcal{D}^k$ is the set of all k-distributions.

[1] The idea to incorporate a set of distributions instead of a single distribution in the epistemic state derives from the philosophical stance that *de re* knowledge about degrees of belief should not be valid. Namely, if epistemic states only contain single distributions, formulae such as $\exists r.\boldsymbol{K}(\boldsymbol{B}(\phi\colon r))$ will be valid for any ϕ, which is counter-intuitive.

When the context is clear, we omit the superscript and write d instead of d^k. We say e is an *epistemic state*, specifically a k-epistemic state, if e is a set of k-distributions, i.e. $e \in \mathcal{E}^k$. The number k is said to be the depth of e, written as $dep[e] = k$.

Similar to $\mathcal{OBL}$, we appeal to three conditions Bnd, Eq, Norm to obtain probability distributions.

Definition 3. *Let d be a $(k+1)$-distribution for some $k \geq 0$, $\mathcal{U}, \mathcal{V} \subseteq (\mathcal{W} \times \mathcal{E}^k \times \cdots \times \mathcal{E}^k)$ and $\mathcal{U} \subseteq \mathcal{V}$, r any real number. We define conditions Bnd, Eq and Norm as follows:*

- *Bnd$(d, \mathcal{U}, r)$ iff there is no l, $\{(w_1, \vec{e}_1), \ldots, (w_l, \vec{e}_l)\} \subseteq \mathcal{U}$ s.t.*[2]
$$\sum_{j=1}^{l} d(w_j, \vec{e}_j) > r$$
- *Eq$(d, \mathcal{U}, r)$ iff Bnd$(d, \mathcal{U}, r)$ and no $r' < r$ s.t. Bnd$(d, \mathcal{U}, r')$;*
- *Norm$(d, \mathcal{U}, \mathcal{V}, r)$ iff there is a number $b \neq 0$ such that Eq$(d, \mathcal{U}, b \times r)$ and Eq$(d, \mathcal{V}, b)$.*

Intuitively, Bnd$(d, \mathcal{U}, r)$ ensures the weight of tuples $(w, \vec{e})$ in $\mathcal{U}$ wrt d is bounded by r. Eq$(d, \mathcal{U}, r)$ ensures that the weight is bounded and r is the supremum. Given Norm$(d, \mathcal{U}, \mathcal{V}, r)$, r can be viewed as the normalized sum of the weight of worlds in $\mathcal{U}$ in relation to $\mathcal{V}$. Essentially, although the distribution d is defined over an uncountable domain, these conditions on d admit a well-defined summation, and the weights on worlds can indeed be interpreted as probabilities:

Theorem 1. *For $k \geq 0$, suppose that d is a $(k+1)$-distribution. Let $\mathcal{V} = (\mathcal{W} \times \mathcal{E}^k \times \cdots \times \mathcal{E}^k)$ and $\mathcal{U} = \{(w, \vec{e}) \mid d(w, \vec{e}) \neq 0\}$. For any $b \geq 0$, if Bnd$(d, \mathcal{V}, b)$, then $\mathcal{U}$ is countable.*

Proof. Let $\mathcal{U}_j = \{(w, \vec{e}) \in \mathcal{U} \mid d(w, \vec{e}) \geq 1/j\}$ for $j \in \mathbb{N}^+$. It is easy to see that $\mathcal{U} = \bigcup \mathcal{U}_i$. Suppose that $\mathcal{U}$ is uncountable, then there is some $\epsilon > 0$ such that $\mathcal{U}_\epsilon = \{w \in \mathcal{U} \mid d(w, \vec{e}) \geq \epsilon\}$ is infinite (Otherwise we could enumerate $\mathcal{U}$ by enumerating $\mathcal{U}_j$ starting at $j = 1$). Consider any countably infinite sequence $(w_l, \vec{e}_l)$ taken from $\mathcal{U}_\epsilon$. Since $d(w_l, \vec{e}_l) \geq \epsilon$ for all l, the sum $\sum_{l=1}^{\infty} d(w_l, \vec{e}_l)$ is unbounded, contradicting the assumption that $\mathcal{V}$ is bounded. □

By satisfying the Bnd conditions, a distribution assigns non-zero values to a countable support set $Supp(d) = \{(w, \vec{e}) \mid d(w, \vec{e}) \neq 0\}$. For the rest of the paper, by summation of weights of a bounded distribution d, we always mean summation in its support set, i.e.

$$\sum_{(w,\vec{e}) \in E} d(w, \vec{e}) \doteq \sum_{(w,\vec{e}) \in E \cap Supp(d)} d(w, \vec{e})$$

In $\mathcal{OBL}$, the necessitation rule does not hold for modality $\boldsymbol{B}$ (also $\boldsymbol{K}$). For instance, True is valid in $\mathcal{OBL}$ yet $\boldsymbol{K}$True is not. Besides, an empty epistemic state is legal in $\mathcal{OBL}$, which breaks the laws of probability (e.g. $\boldsymbol{B}(p\colon 0.1) \wedge \boldsymbol{B}(p\colon 0.2)$ will be satisfied by the empty state). Fortunately, these undesirable properties can be easily avoided by ruling out improper distributions and epistemic states.

Definition 4 (regularity). *We inductively define the sets of **regular** k-distributions $\mathcal{D}^k$ and **regular** k-epistemic states $\mathcal{E}^k$: Let $\mathcal{E}^0 = \{\varnothing\}$. For any $k > 0$, let $\mathcal{V} = (\mathcal{W} \times \mathcal{E}^{k-1} \times \cdots \times \mathcal{E}^{k-1})$,*

[2] We write the arguments of a k-distribution as $(w, \vec{e})$ instead of $(w, e_1, \ldots, e_{m-1})$.

- $\mathcal{D}^k = \{d \mid \text{Eq}(d, \mathcal{V}, 1)\}$;
- $\mathcal{E}^k = 2^{\mathcal{D}^k} \setminus \{\varnothing\}$.

For the rest of the paper, we consider only regular distributions and epistemic states. Now we define the truth of sentences. By a model, we mean a tuple $(w, \vec{e})$, where $\vec{e} = (e_1, \ldots, e_m)$ and each e_i is an epistemic state (e_i denotes the i-th argument of $\vec{e}$). We say a formula α and $\vec{e}$ are *compatible* if $dep[\alpha, i] \leq dep[e_i]$ f.a. $i \in \{1, \ldots, m\}$. The truth value of objective sentences is assigned as follows:

- $w, \vec{e} \models P(n)$ iff $P(n) \in w$;
- $w, \vec{e} \models t_1 = t_2$ iff t_1 and t_2 are identical standard name;
- $w, \vec{e} \models \neg\alpha$ iff not $w, \vec{e} \models \alpha$;
- $w, \vec{e} \models \alpha \wedge \beta$ iff $w, \vec{e} \models \alpha$ and $w, \vec{e} \models \beta$;
- $w, \vec{e} \models \forall x.\alpha$ iff $w, \vec{e} \models \alpha^x_n$ for any $n \in \mathcal{N}$.

Here α^x_n means the formula obtained by substituting each appearance of free variable x in α by a standard name n. The semantics for the objective fragment is identical to language $\mathcal{L}$ [24]. Supposing that $\vec{e}$ is compatible with $\boldsymbol{B}_i(\alpha\colon r)$, then

- $w, \vec{e} \models \boldsymbol{B}_i(\alpha\colon r)$ iff f.a. $d \in e_i$, $\text{Norm}(d, \mathcal{W}^{i,e_i}_\alpha, \mathcal{W}^{i,e_i}_{\text{True}}, r)$,

where for $i \in Ag$, $e \in \mathcal{E}^k$ and formula φ, $\mathcal{W}^{i,e}_\varphi$ is defined as

$$\mathcal{W}^{i,e}_\varphi = \{(w, \vec{e}_{-i}) \mid w, e_1, \ldots, e_{i-1}, e, e_{i+1}, \ldots, e_m \models \varphi, \quad e_j \in \mathcal{E}^{k-1} \text{ for } j \neq i\}$$

and $\vec{e}_{-i} = (e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_m)$. Apparently, for any i and $e \in \mathcal{E}^{k+1}$, $\mathcal{W}^{i,e}_{\text{True}} = (\mathcal{W} \times \mathcal{E}^k \times \cdots \times \mathcal{E}^k)$. We drop the superscript and write $\mathcal{W}_{\text{True}}$ for simplicity.

Given an epistemic state of depth k, we can define the semantics of only-believing up to depth k:

- For $e_i \in \mathcal{E}^k$, $w, \vec{e} \models \boldsymbol{O}^{(k)}_i(\alpha_1\colon r_1, \ldots, \alpha_l\colon r_l)$ iff f.a. $d \in \mathcal{D}^k$,

  $d \in e_i$ **iff** $\text{Norm}(d, \mathcal{W}^{i,e_i}_{\alpha_j}, \mathcal{W}_{\text{True}}, r_j)$ for $j \in \{1, \ldots, l\}$.

We still need to assign truth values for only-believing up to depth k, while the model is of a greater depth (otherwise sentences like $\boldsymbol{O}^{(1)}_1 p \wedge \boldsymbol{O}^{(2)}_1 (p \wedge \boldsymbol{K}_2 p)$ are unsatisfiable). To do so, we need to capture a model's belief of lower depths and omit the depths greater than needed. This is achieved via the notion which we call "*regression*":

Definition 5 (Regression). *We inductively define the regression of any distributions and epistemic states:*

- *For $d \in \mathcal{D}^2$, $d' \in \mathcal{D}^1$, d' is the regression of d, written as $d' = \boldsymbol{R}[d]$ if for any w,*
  $$\sum_{e^*_1, \ldots, e^*_{m-1}} d(w, e^*_1, \ldots e^*_{m-1}) = d'(w, \varnothing, \ldots, \varnothing)$$
- *For $e \in \mathcal{E}^2$, $e' \in \mathcal{E}^1$, we say e' is the regression of e, written as $e' = \boldsymbol{R}[e]$, iff $e' = \{\boldsymbol{R}[d] \mid d \in e\}$*
- *For $d \in \mathcal{D}^{k+1}$, $d' \in \mathcal{D}^k$, d' is the regression of d, written as $d' = \boldsymbol{R}[d]$, if for any $w, e'_1, \ldots, e'_{m-1}$*
  $$\sum_{\tilde{e}_1, \ldots, \tilde{e}_{m-1} \in E} d(w, \tilde{e}_1, \ldots, \tilde{e}_{m-1}) = d'(w, e'_1, \ldots, e'_{m-1})$$
  where $E = \{\tilde{e}_1, \ldots, \tilde{e}_{m-1} \mid e'_1 = \boldsymbol{R}[\tilde{e}_1], \ldots, e'_{m-1} = \boldsymbol{R}[\tilde{e}_{m-1}]\}$
- *For $e \in \mathcal{E}^{k+1}$, $e' \in \mathcal{E}^k$, we say e' is the regression of e, written as $e' = \boldsymbol{R}[e]$, iff $e' = \{\boldsymbol{R}[d] \mid d \in e\}$*

By definition, for each $e \in \mathcal{E}^{k+1}$, there exists a unique $e' \in \mathcal{E}^k$ s.t. $e' = \boldsymbol{R}[e]$. We write $e' = \boldsymbol{R}^{(2)}[e]$ to mean $e' = \boldsymbol{R}[\boldsymbol{R}[e]]$. Analogously, $e' = \boldsymbol{R}^{(k)}[e]$ means $e' = \boldsymbol{R}[\cdots \boldsymbol{R}[e] \cdots]$ with k nested $\boldsymbol{R}$. The following lemma indicates that $\boldsymbol{R}[e]$ faithfully reflects the properties of e in lower depths. Therefore, for $e \in \mathcal{E}^{k+1}$, it is reasonable to assign truth values for only-believing up to depth k based on the truth assignment of $\boldsymbol{R}[e]$.

Lemma 1. *Given $\vec{e}$ and $\vec{e'}$ s.t. $e'_i = \boldsymbol{R}[e_i]$ f.a. $i \in \{1, \ldots, m\}$. For any formula α compatible with $\vec{e'}$ and mentioning no $\boldsymbol{O}$-operators, $w, \vec{e} \models \alpha$ iff $w, \vec{e'} \models \alpha$.*

Proof. We prove the lemma via induction.

Basis:

For objective α, the proof is trivial since the truth value is irrelevant to $\vec{e}$ and $\vec{e'}$.

Induction hypothesis:

Suppose that the statement holds for any $\vec{e}, \vec{e'}, \alpha'$ s.t. $e'_i = \boldsymbol{R}[e_i]$, $e'_i \in \mathcal{E}^l$ for some $l < k_i$ and α' compatible with $\vec{e'}$.

- Induction on $\wedge$, $\neg$ or $\forall$ is trivial.
- For $e_i \in \mathcal{E}^{k_i+1}$, $e_i \models \boldsymbol{B}_i(\alpha\colon r)$ iff f.a. $d \in e_i$,[3]
  $$\sum_{(w, \vec{e^*}_{-i}) \in \mathcal{W}^{i,e_i}_\alpha} d(w, \vec{e^*}_{-i}) = r \tag{\#}$$

  We introduce an auxiliary function for the proof:
  For $\vec{e'}_{-i} = (e'_1, \ldots, e'_{i-1}, e'_{i+1}, \ldots, e'_m)$, $e'_j \in \mathcal{E}^k$, and $\vec{e''}_{-i} = (e''_1, \ldots, e''_{i-1}, e''_{i+1}, \ldots, e''_m)$, $e''_j \in \mathcal{E}^{k-1}$, we define
  $$\mathbb{I}(\vec{e'}_{-i}, \vec{e''}_{-i}) = \begin{cases} 1 & e''_j = \boldsymbol{R}[e'_j] \text{ f.a. } j \neq i; \\ 0 & \text{otherwise.} \end{cases}$$

  Since the regression is unique, given $\vec{e'}_{-i}$, there's only one tuple $\vec{e''}_{-i}$ with $\mathbb{I}(\vec{e'}_{-i}, \vec{e''}_{-i}) = 1$. Thus when $\vec{e'}_{-i}$ is fixed,
  $$\sum_{\vec{e''}_{-i}} \mathbb{I}(\vec{e'}_{-i}, \vec{e''}_{-i}) = 1 \tag{*}$$

  Back to (#), let $d' = \boldsymbol{R}[d]$, we have
  $$\begin{aligned} r &= \sum_{(w, \vec{e^*}_{-i}) \in \mathcal{W}^{i,e_i}_\alpha} d(w, \vec{e^*}_{-i}) \cdot 1 \\ &= \sum_{(w, \vec{e^*}_{-i}) \in \mathcal{W}^{i,e_i}_\alpha} \Big(d(w, \vec{e^*}_{-i}) \cdot \sum_{\vec{e''}_{-i}} \mathbb{I}(\vec{e^*}_{-i}, \vec{e''}_{-i})\Big) && \text{(Eq. *)} \\ &= \sum_{\{(w, \vec{e^*}_{-i}, \vec{e''}_{-i}) \mid (w, \vec{e^*}_{-i}) \in \mathcal{W}^{i,e_i}_\alpha\}} d(w, \vec{e^*}_{-i}) \cdot \mathbb{I}(\vec{e^*}_{-i}, \vec{e''}_{-i}) \\ &= \sum_{\{(w, \vec{e^*}_{-i}, \vec{e''}_{-i}) \mid (w, \vec{e^*}_{-i}) \in \mathcal{W}^{i,e_i}_\alpha, e''_j = \boldsymbol{R}[e^*_j] \text{ f.a. } j \neq i\}} d(w, \vec{e^*}_{-i}) \end{aligned}$$
  (Def. $\mathbb{I}$)

[3] By definition, the truth of $\boldsymbol{B}_i(\alpha\colon r)$ is irrelevant to w and epistemic states of other agents except i. Thus we write $e_i \models \boldsymbol{B}_i(\alpha\colon r)$ instead of $w, \vec{e} \models \boldsymbol{B}_i(\alpha\colon r)$.

For $(w, \vec{e^*}_{-i}) \in \mathcal{W}_\alpha^{i,e_i}$, $e'_i = \boldsymbol{R}[e_i]$ and $e''_j = \boldsymbol{R}[e^*_j]$ f.a. $j \neq i$, by induction hypothesis we have $w, e''_1, \ldots, e'_i, \ldots, e''_m \models \alpha$, i.e. $(w, \vec{e''}_{-i}) \in \mathcal{W}_\alpha^{i,e'_i}$. Therefore,

$$
\begin{aligned}
r &= \sum_{\{(w,\vec{e^*}_{-i},\vec{e''}_{-i}) \mid (w,\vec{e''}_{-i}) \in \mathcal{W}_\alpha^{i,e'_i}, e''_j = \boldsymbol{R}[e^*_j] \text{ f.a. } j \neq i\}} d(w, \vec{e^*}_{-i}) \quad \text{(IH)} \\
&= \sum_{(w,\vec{e''}_{-i}) \in \mathcal{W}_\alpha^{i,e'_i}} \Big(\sum_{\{\vec{e^*}_{-i} \mid e''_j = \boldsymbol{R}[e^*_j] \text{ f.a. } j \neq i\}} d(w, \vec{e^*}_{-i}) \Big) \\
&= \sum_{(w,\vec{e''}_{-i}) \in \mathcal{W}_\alpha^{i,e'_i}} d'(w, \vec{e''}_{-i}) \quad \text{(Def. 5)}
\end{aligned}
$$

Thus $\textsc{Norm}(d', \mathcal{W}_\alpha^{i,e'_i}, \mathcal{W}_{\textsc{True}}^{i,e'_i}, r)$ f.a. $d' \in e'_i$, i.e. $e'_i \models \boldsymbol{B}_i(\alpha : r)$ □

Now we can complete the semantics:

- For $e_i \in \mathcal{E}^{k'}$ s.t. $k' > k$, $w, \vec{e} \models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l)$ iff $w, e_1, \ldots, R[e_i], \ldots, e_m \models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l)$.

The result of Lem.1 can be extended to any wffs, including those with only-believing:

Theorem 2. *Given $\vec{e}, \vec{e'}$ s.t. $e'_i = \boldsymbol{R}[e_i]$. For any formula α compatible with both $\vec{e}, \vec{e'}$, then $w, \vec{e} \models \alpha$ iff $w, \vec{e'} \models \alpha$*

The proof is similar to Lem 1. The only difference is the induction for $\boldsymbol{O}$-formulae, which directly follows the definition of the semantics.

For a sentence α and a set of sentences Σ, we write $\Sigma \models \alpha$ (read as Σ logically entails α) to mean that for every model $(w, e_1, \ldots, e_m)$ compatible with α and all sentences in Σ, if $w, e_1, \ldots, e_m \models \alpha'$ f.a. $\alpha' \in \Sigma$, then $w, e_1, \ldots, e_m \models \alpha$. We say α is valid (written as $\models \alpha$) if $\{\} \models \alpha$. When α is objective, we write $w \models \alpha$ instead of $w, e_1, \ldots, e_m \models \alpha$. When α is subjective, we write $e_1, \ldots, e_m \models \alpha$ (or $\vec{e} \models \alpha$). For i-subjective α, we write $e_i \models \alpha$.

3 PROPERTIES OF THE LOGIC

In this section, we study the properties of modalities $\boldsymbol{B}_i$ and $\boldsymbol{K}_i$, For $\boldsymbol{K}_i$, where the probabilistic belief is reduced to a categorical one, $\mathcal{OBL}_m$ satisfies the $KD45_n$ properties. The Barcan formulae[16] are also valid. For $\boldsymbol{B}_i$, we show that the degree of belief follows the laws of probability, the properties of introspection are also extended to more general cases. We provide proofs for some of the properties. The rest can be proved similarly.

3.1 Knowledge

$\mathcal{OBL}_m$ satisfies the $KD45_n$ properties:

- (Necessitation) If $\models \alpha$, then $\models \boldsymbol{K}_i\alpha$.
- (Consistency) $\models \boldsymbol{K}_i\alpha \supset \neg\boldsymbol{K}_i\neg\alpha$

Proof. Suppose that $e_i \models \boldsymbol{K}_i\alpha$, then for any $d \in e_i$,

$$
\begin{aligned}
\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\neg\alpha}^{i,e_i}} d(w, \vec{e'}_{-i}) &= \sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\textsc{True}} \setminus \mathcal{W}_\alpha^{i,e_i}} d(w, \vec{e'}_{-i}) \\
&= 1 - \sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_\alpha^{i,e_i}} d(w, \vec{e'}_{-i}) = 0
\end{aligned}
$$

Thus $e_i \models \neg\boldsymbol{K}_i\neg\alpha$. □

- (Distribution) $\models \boldsymbol{K}_i\alpha \wedge \boldsymbol{K}_i(\alpha \supset \beta) \supset \boldsymbol{K}_i\beta$

Proof. For any e_i compatible with $\boldsymbol{K}_i\alpha$ and $\boldsymbol{K}_i(\alpha \supset \beta)$, if $e_i \models \boldsymbol{K}_i\alpha \wedge \boldsymbol{K}_i(\alpha \supset \beta)$, then for $d \in e_i$,

$$
\begin{aligned}
\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_\beta^{i,e_i}} d(w, \vec{e'}_{-i}) = &\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_\alpha^{i,e_i}} d(w, \vec{e'}_{-i}) \\
&+ \sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\alpha\supset\beta}^{i,e_i}} d(w, \vec{e'}_{-i}) \\
&- \sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\alpha\vee\alpha\supset\beta}^{i,e_i}} d(w, \vec{e'}_{-i}) \\
&= 1 + 1 - 1 = 1
\end{aligned}
$$

i.e. $e_i \models \boldsymbol{K}_i\beta$. □

- (Pos. Introspection) $\models \boldsymbol{K}_i\alpha \supset \boldsymbol{K}_i\boldsymbol{K}_i\alpha$
- (Neg. Introspection) $\models \neg\boldsymbol{K}_i\alpha \supset \boldsymbol{K}_i\neg\boldsymbol{K}_i\alpha$

Proof. Suppose that $e_i \models \neg\boldsymbol{K}_i\alpha$, then f.a. w and $\vec{e'}_{-i}$, $w, e'_1, \ldots, e'_{i-1}, e_i, e'_{i+1}, \ldots, e'_m \models \neg\boldsymbol{K}_i\alpha$. Thus f.a. $d \in e_i$,

$$
\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\neg\boldsymbol{K}_i\alpha}^{i,e_i}} d(w, \vec{e'}_{-i}) = \sum_{(w,\vec{e'}_{-i})} d(w, \vec{e'}_{-i}) = 1
$$

i.e. $e_i \models \boldsymbol{K}_i\neg\boldsymbol{K}_i\alpha$. □

Barcan formulae (both universal and existential versions):

- $\models \forall x.\boldsymbol{K}_i\alpha \supset \boldsymbol{K}_i\forall x.\alpha$

  Proof. Suppose that $e_i \models \forall x.\boldsymbol{K}_i\alpha$. By definition we have $e_i \models \boldsymbol{K}_i\alpha_n^x$ for any $n \in \mathcal{N}$. Then f.a. $d \in e_i$,

$$
\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\alpha_n^x}^{i,e_i}} d(w, \vec{e'}_{-i}) = 1 \text{ f.a. } n \in \mathcal{N}
$$

  Namely,

$$
\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\neg\alpha_n^x}^{i,e_i}} d(w, \vec{e'}_{-i}) = 0 \text{ f.a. } n \in \mathcal{N}
$$

  Therefore

$$
\begin{aligned}
\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\forall x.\alpha}^{i,e_i}} d(w, \vec{e'}_{-i}) &= 1 - \sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\exists x.\neg\alpha}^{i,e_i}} d(w, \vec{e'}_{-i}) \\
&\geq 1 - \sum_{n \in \mathcal{N}} \sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}_{\neg\alpha_n^x}^{i,e_i}} d(w, \vec{e'}_{-i}) = 1
\end{aligned}
$$

  Thus $e_i \models \boldsymbol{K}_i\forall x.\alpha$ □

  It is worth mentioning that the converse result also holds, i.e. $\models \boldsymbol{K}_i\forall x.\alpha \supset \forall x.\boldsymbol{K}_i\alpha$.
- $\models \exists x.\boldsymbol{K}_i\alpha \supset \boldsymbol{K}_i\exists x.\alpha$
  The proof is of the same spirit as the universal one.
- $\not\models \boldsymbol{K}_i\exists x.\alpha \supset \exists x.\boldsymbol{K}_i\alpha$.
  Let $w, w' \in \mathcal{W}$ satisfy $P(n_1) \in w$, $P(n') \notin w$ f.a. $n' \neq n_1$, $P(n_2) \in w'$, $P(n') \notin w'$ f.a. $n' \neq n_2$. Let $d_i \in \mathcal{D}^1$ assign weight 0.5 to both w and w'. Other worlds are assigned 0. Then $\{d_i\}$ satisfies $\boldsymbol{K}_i\exists x.\alpha$ but not $\exists x.\boldsymbol{K}_i\alpha$.

3.2 Degree of Belief

For any $i \in Ag$ and formula α, β,

- if α is valid, then $\models \boldsymbol{B}_i(\alpha\colon 1)$
- $\models \boldsymbol{B}_i(\alpha\colon r) \supset \neg\boldsymbol{B}_i(\alpha\colon r')$ for $r' \neq r$
- if $\models \alpha \equiv \beta$, then $\models \boldsymbol{B}_i(\alpha\colon r) \equiv \boldsymbol{B}_i(\beta\colon r)$ for any r.
- $\models \boldsymbol{B}_i(\alpha\colon r) \supset \boldsymbol{B}_i(\neg\alpha\colon 1-r)$
- $\models \boldsymbol{B}_i(\alpha \wedge \beta\colon r) \wedge \boldsymbol{B}_i(\alpha \wedge \neg\beta\colon r') \supset \boldsymbol{B}_i(\alpha\colon r+r')$
- $\models \boldsymbol{B}_i(\alpha\colon r) \wedge \boldsymbol{B}_i(\beta\colon r') \wedge \boldsymbol{B}_i(\alpha \wedge \beta\colon r'') \supset \boldsymbol{B}_i(\alpha \vee \beta\colon n)$ where n is a standard name of sort number and $n = r+r'-r''$.

Most of the properties can be proved in a way similar to those of the previous subsection. Here we only provide the proof of the last one:

Proof. Suppose that $e_i \models \boldsymbol{B}_i(\alpha\colon r) \wedge \boldsymbol{B}_i(\beta\colon r') \wedge \boldsymbol{B}_i(\alpha \wedge \beta\colon r'')$, then f.a. $d \in e_i$, it holds $\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha}} (w, \vec{e'}_{-i}) = r$, $\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\beta}} d(w, \vec{e'}_{-i}) = r'$, $\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha\wedge\beta}} d(w, \vec{e'}_{-i}) = r''$. By applying the laws of set,

$$\mathcal{W}^{i,e_i}_{\alpha\vee\beta} = \mathcal{W}^{i,e_i}_{\alpha} + \mathcal{W}^{i,e_i}_{\neg\alpha\wedge\beta} = \mathcal{W}^{i,e_i}_{\alpha} + \mathcal{W}^{i,e_i}_{\beta} - \mathcal{W}^{i,e_i}_{\alpha\wedge\beta}$$

Thus $\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha\vee\beta}} d(w, \vec{e'}_{-i}) = r + r' - r''$ □

Results of introspection also hold for degrees of belief. Furthermore, it can be extended to arbitrary i-subjective sentences.

- $\models \boldsymbol{B}_i(\alpha\colon r) \supset \boldsymbol{K}_i\boldsymbol{B}_i(\alpha\colon r)$
- $\models \neg\boldsymbol{B}_i(\alpha\colon r) \supset \boldsymbol{K}_i\neg\boldsymbol{B}_i(\alpha\colon r)$
- For any i-subjective formula α, $\models \alpha \supset \boldsymbol{K}_i\alpha$

Proof. Suppose that α is i-subjective and $e_i \models \alpha$. Since α is i-subjective, given a model $(w, e_1, \ldots, e_m)$, the truth value is irrelevant to w or epistemic states other than e_i. For any $d \in e_i$,

$$\sum_{\{w,\vec{e'}_{-i} \mid w,\vec{e'} \models \alpha\}} d(w, \vec{e'}_{-i}) = \sum_{\{w,\vec{e'}_{-i}\}} d(w, \vec{e'}_{-i}) = 1$$

Thus $e_i \models \boldsymbol{K}_i\alpha$ □

4 ONLY-BELIEVING

We discuss the properties of only-believing in this section. First, we examine the relation between only-believing up to different depths and demonstrate how a hierarchy of only-believing is built. For cases of only-believing when the argument is a group of i-objective formulae, we show the uniqueness of the model and the nice properties it brings. For sentences beyond i-objective, we demonstrate how certain types of autoepistemic reasoning can be modelled, and how the specification of only-believing contributes to the expressiveness.

Intuitively, what being only-believed should be believed at first:

Proposition 1. $\models \boldsymbol{O}^{(k)}_i(\alpha_1\colon r_1, \ldots \alpha_l\colon r_l) \supset \bigwedge_{j=1}^{l} \boldsymbol{B}_i(\alpha_j\colon r_j)$

If an agent only believes α with degree r up to a certain depth, then she should also only believe it up to a lower (but compatible) depth. The converse result does not necessarily hold. Formally,

Proposition 2. $\models \boldsymbol{O}^{(k+1)}_i(\alpha\colon r) \supset \boldsymbol{O}^{(k)}_i(\alpha\colon r)$.

Proof. The proof is similar to Lem. 1 with function $\mathbb{I}$ as used in the lemma. Suppose that $e_i \models \boldsymbol{O}^{(k+1)}_i(\alpha\colon r)$, w.l.o.g. we assume that $e_i \in \mathcal{E}^{(k+1)}$. By Thm 2 it suffices to prove that $e'_i \models \boldsymbol{O}^{(k)}_i(\alpha\colon r)$ for $e'_i = \boldsymbol{R}[e_i]$. For $d \in \mathcal{D}^{(k+1)}$, let d' be the k-distribution s.t. $d' = \boldsymbol{R}[d]$. By the semantics, $d \in e_i$ iff

$$\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha}} d(w, \vec{e'}_{-i}) = r$$

$$\begin{aligned}
&\sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha}} d(w, \vec{e'}_{-i}) \cdot 1 \\
=& \sum_{(w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha}} \Big(d(w, \vec{e'}_{-i}) \cdot \sum_{\vec{e''}_{-i}} \mathbb{I}(\vec{e'}_{-i}, \vec{e''}_{-i})\Big) && \text{(Eq. } *) \\
=& \sum_{\{(w,\vec{e'}_{-i},\vec{e''}_{-i}) \mid (w,\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha}, e''_j = \boldsymbol{R}[e'_j] \text{ f.a. } j\neq i\}} d(w, \vec{e'}_{-i}) \\
=& \sum_{\{(w,\vec{e'}_{-i},\vec{e''}_{-i}) \mid (w,\vec{e''}_{-i}) \in \mathcal{W}^{i,e'_i}_{\alpha}, e''_j = \boldsymbol{R}[e'_j] \text{ f.a. } j\neq i\}} d(w, \vec{e'}_{-i}) && \text{(Thm. 2)} \\
=& \sum_{(w,\vec{e''}_{-i}) \in \mathcal{W}^{i,e'_i}_{\alpha}} \Big(\sum_{\{\vec{e'}_{-i} \mid e''_j = \boldsymbol{R}[e'_j] \text{ f.a. } j\neq i\}} d(w, \vec{e'}_{-i})\Big) \\
=& \sum_{(w,\vec{e''}_{-i}) \in \mathcal{W}^{i,e'_i}_{\alpha}} d'(w, \vec{e''}_{-i}) = r && \text{(Def. 5)}
\end{aligned}$$

By Def. 5, $d' \in e'_i$ iff $d \in e_i$. Thus $e'_i \models \boldsymbol{O}^{(k)}_i(\alpha\colon r)$. □

There exists α s.t. $\not\models \boldsymbol{O}^{(k)}_i(\alpha\colon r) \supset \boldsymbol{O}^{(k+1)}_i(\alpha\colon r)$. An example can be easily constructed: Let $e_i \in \mathcal{E}^2$ and $e_i \models \boldsymbol{O}^{(2)}_1(p \wedge \boldsymbol{K}_2 p)$, then $e_i \models \boldsymbol{O}^{(1)}_1 p$ but $e_i \not\models \boldsymbol{O}^{(2)}_1 p$.

The proposition demonstrates that from $\boldsymbol{O}^{(k)}_i$ to $\boldsymbol{O}^{(k+1)}_i$ it is indeed a more precise specification of an agent's only-believing.

There are unsatisfiable $\boldsymbol{O}$-formulae, for example $\boldsymbol{O}^{(k)}_i(p \wedge \boldsymbol{K}_i \neg p)$. Fortunately, for a large fragment of the language, an $\boldsymbol{O}$-formula is satisfiable when the corresponding belief formulae are satisfiable:

Proposition 3. *Let $\alpha_1, \ldots, \alpha_l$, be i-objective. If $\bigwedge_{j=1}^{l} \boldsymbol{B}_i(\alpha_j\colon r_j)$ is satisfiable, then $\boldsymbol{O}^{(k)}_i(\alpha_1\colon r_1, \ldots \alpha_l\colon r_l)$ is satisfiable.*

4.1 Unique Model and Properties

In general, an $\boldsymbol{O}^{(k)}_i$-formula could be satisfied by more than one epistemic state. For example $\boldsymbol{O}^{(1)}_1\big((p \wedge \boldsymbol{K}_1 p) \vee (\neg p \wedge \boldsymbol{K}_1 \neg p)\big)$. For i-objective sentence α, however, there is a unique k-epistemic state $e_i \in \mathcal{E}^k$ which satisfies $\boldsymbol{O}^{(k)}_i \alpha$:

Lemma 2. *Given i-objective α and k s.t. $\boldsymbol{K}_i\alpha$ is satisfiable and $\boldsymbol{O}^{(k)}_i\alpha$ is a wff, there is precisely one $e_i \in \mathcal{E}^k$ s.t. $e_i \models \boldsymbol{O}^{(k)}_i\alpha$.*

Proof. Prop. 3 proves the existence of e_i. Given $e_i \in \mathcal{E}^k$, by the definition $e_i \models \boldsymbol{O}^{(k)}_i\alpha$ iff f.a. $d \in \mathcal{D}^k$, $d \in e_i$ iff

$$\sum_{(w',\vec{e'}_{-i}) \in \mathcal{W}^{i,e_i}_{\alpha}} d(w', \vec{e'}) = 1$$

Since α is i-objective, if $w', e'_1, \ldots, e_i, \ldots, e'_m \models \alpha$, then f.a. $\tilde{e}_i$, $w', e'_1, \ldots, \tilde{e}_i, \ldots, e'_m \models \alpha$. For any $\tilde{e}_i \in \mathcal{E}^k$ s.t. $\tilde{e}_i \models \boldsymbol{O}_i^{(k)}\alpha$, it is trivial that f.a. $d \in \mathcal{D}^k$, $d \in e_i$ iff $d \in \tilde{e}_i$, i.e. $e_i = \tilde{e}_i$. □

For a model of a depth greater than needed, the uniqueness of the model can be interpreted as follows: Suppose that $e_i \in \mathcal{E}^k$ is the unique model satisfying $\boldsymbol{O}_i^{(k)}\alpha$. Then for all $e_i^* \in \mathcal{E}^{k'}$ s.t. $k' > k$ and $e_i^* \models \boldsymbol{O}_i^{(k)}\alpha$, e_i^* will reduce to e_i after finite steps of regression:

Lemma 3. *Given i-objective α and k s.t. $\boldsymbol{K}_i\alpha$ is satisfiable and $\boldsymbol{O}_i^{(k)}\alpha$ is a wff. Let $e_i \in \mathcal{E}^k$ be the model s.t. $e_i \models \boldsymbol{O}_i^{(k)}\alpha$. For any e_i^* such that $e_i^* \models \boldsymbol{O}_i^{(k)}\alpha$, $e_i = \boldsymbol{R}^{(l)}[e_i^*]$ for some l.*

Proof. For any $k' > k$ and $e_i^* \in \mathcal{E}^{k'}$, e_i^* reduces to $\tilde{e}_i \in \mathcal{E}^k$ after $k' - k$ steps of regression. Given $e_i^* \models \boldsymbol{O}_i^{(k)}\alpha$, by applying Thm. 2 we have $\tilde{e}_i \models \boldsymbol{O}_i^{(k)}\alpha$. By Lem. 2, $\tilde{e}_i$ is unique, i.e. $\tilde{e}_i = e_i$. □

In general, the results can be extended to the case with multiple beliefs:

Theorem 3. *Given i-objective $\alpha_1 \ldots \alpha_l$ s.t. $\bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j)$ is satisfiable and $\boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots \alpha_l : r_l)$ is a wff, there is precisely one $e_i \in \mathcal{E}^k$ such that $e_i \models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots \alpha_l : r_l)$. For any e_i^* such that $e_i^* \models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots \alpha_l : r_l)$, $e_i = \boldsymbol{R}^{(l)}[e_i^*]$ for some l.*

The proof is similar to Lem.2 and Lem.3.

A nice property of the uniqueness is that, given what is only-believed by an agent, everything else being believed or not believed by the agent is logically implied. Formally, we have the following results:

Theorem 4. *Given i-objective formulae $\alpha_1, \ldots, \alpha_l$ and arbitrary β s.t. $\boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l)$ is satisfiable, $dep[\boldsymbol{B}_i(\beta : r'), i] \le k$. Then for any $0 < r' < 1$, either $\boldsymbol{B}_i(\beta : r')$ or $\neg\boldsymbol{B}_i(\beta : r')$ is entailed by $\boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l)$, and*

- $\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \boldsymbol{B}_i(\beta : r')$ *iff* $\models \bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j) \supset \boldsymbol{B}_i(\beta : r')$
- $\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \neg\boldsymbol{B}_i(\beta : r')$ *iff* $\not\models \bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j) \supset \boldsymbol{B}_i(\beta : r')$

Proof. By Thm.3, it exists a unique $e_i \in \mathcal{E}^k$ such that $e_i \models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l)$. If $\bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j) \supset \boldsymbol{B}_i(\beta : r')$ is valid, by Prop.1, $\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \boldsymbol{B}_i(\beta : r')$. When $\bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j) \supset \boldsymbol{B}_i(\beta : r')$ is invalid, there exists $d_i \in \mathcal{D}^k$ s.t. $\{d_i\} \models \bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j) \wedge \neg\boldsymbol{B}_i(\beta : r')$. By the semantics $d_i \in e_i$. Hence $e_i \models \neg\boldsymbol{B}_i(\beta : r')$ and $\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \neg\boldsymbol{B}_i(\beta : r')$. For any $e'_i \in \mathcal{E}^{k'}$ s.t. $k' > k$, $e'_i \models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l)$, By Thm. 2 and 3, $e'_i \models \boldsymbol{B}_i(\beta : r')$ iff $e_i \models \boldsymbol{B}_i(\beta : r')$. □

Theorem 5. *Given i-objective formulae $\alpha_1, \ldots, \alpha_l$ and arbitrary β, s.t. $\boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l)$ is satisfiable, $dep[\boldsymbol{K}_i\beta, i] \le k$, then either $\boldsymbol{O}_i^{(k)}(\alpha : r) \supset \boldsymbol{K}_i\beta$ or $\boldsymbol{O}_i^{(k)}(\alpha : r) \supset \neg\boldsymbol{K}_i\beta$ is valid, and*

- $\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \boldsymbol{K}_i\beta$ *iff* $\models \bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j) \supset \boldsymbol{K}_i\beta$ *or* $\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \beta$
- $\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \neg\boldsymbol{K}_i\beta$ *iff* $\not\models \bigwedge_j \boldsymbol{B}_i(\alpha_j : r_j) \supset \boldsymbol{K}_i\beta$ *and* $\not\models \boldsymbol{O}_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \beta$

Given distinct propositions p, q, to determine whether $\boldsymbol{O}_i^{(k)}p$ entails $\boldsymbol{K}_i\neg\boldsymbol{K}_iq$, it suffices to check the validity of $\boldsymbol{K}_ip \supset \boldsymbol{K}_i\neg\boldsymbol{K}_iq$ and $\boldsymbol{O}_i^{(k)}p \supset \neg\boldsymbol{K}_iq$. The latter is valid since $\not\models \boldsymbol{K}_ip \supset \boldsymbol{K}_iq$ and $\not\models \boldsymbol{O}_i^{(k)}p \supset q$. Thus $\models \boldsymbol{O}_i^{(k)}p \supset \boldsymbol{K}_i\neg\boldsymbol{K}_iq$.

Note that this is not true for $\boldsymbol{K}_i$, e.g. neither $\boldsymbol{K}_ip \supset \boldsymbol{K}_i\neg\boldsymbol{K}_iq$ nor $\boldsymbol{K}_ip \supset \neg\boldsymbol{K}_i\neg\boldsymbol{K}_iq$ is valid. Also, we would like to reiterate that the results of Thm.4 and Thm.5 hold only under the restriction of depth and do not hold when $dep[\boldsymbol{K}_i\beta, i] > k$. For example, both $\boldsymbol{O}_1^{(1)}p \wedge \boldsymbol{K}_1(p \wedge \boldsymbol{K}_2p)$ and $\boldsymbol{O}_1^{(1)}p \wedge \boldsymbol{K}_1(p \wedge \neg\boldsymbol{K}_2p)$ are satisfiable.

If β is also i-objective, we can reduce the use of the outermost modalities:

Corollary 1. *Given i-objective α, β, s.t. $\boldsymbol{O}_i^{(k)}\alpha$ is satisfiable, then*

- $\models \boldsymbol{O}_i^{(k)}\alpha \supset \boldsymbol{K}_i\beta$ *iff* $\models \alpha \supset \beta$
- $\models \boldsymbol{O}_i^{(k)}\alpha \supset \neg\boldsymbol{K}_i\beta$ *iff* $\not\models \alpha \supset \beta$

4.2 Default Reasoning

Levesque and Lakemeyer have already shown how defaults can be encoded in terms of only-knowing [24]. We can go beyond that: Besides the well-known "birds can fly" default, our account can also express defaults about what another agent believes.

Example 2. *Let predicate $Fair(x)$ mean "(The coin) x is fair". Let δ denote the sentence*

$$\forall x.[Coin(x) \wedge \neg\exists r.(r > 0 \wedge \boldsymbol{B}_1(\neg\boldsymbol{K}_2Fair(x) : r)) \supset \boldsymbol{K}_2Fair(x)]$$

i.e. for any coin, agent 2 believes that the coin is fair unless it is believed to a positive degree that she doesn't believe it. Let KB *be the sentence $\{Coin(silver)\}$, then the following sentences are valid:*

- $\boldsymbol{O}_1^{(2)}(\mathbf{KB} \wedge \delta) \supset \boldsymbol{K}_1\boldsymbol{K}_2Fair(silver)$
- $\boldsymbol{O}_1^{(2)}(\mathbf{KB} \wedge \delta \wedge \boldsymbol{K}_2Fair(silver)) \supset \boldsymbol{K}_1\boldsymbol{K}_2Fair(silver)$
- $\boldsymbol{O}_1^{(2)}(\mathbf{KB} \wedge \delta \wedge \neg\boldsymbol{K}_2Fair(silver)) \supset \boldsymbol{K}_1\neg\boldsymbol{K}_2Fair(silver)$

We show the validity of the first sentence. For simplicity, only two agents are considered.

Proof. Let $e_1 \models \boldsymbol{O}_1^{(2)}(\mathbf{KB} \wedge \delta)$, w.l.o.g. $e_1 \in \mathcal{E}^2$. We first prove $e_1 \models \neg\exists r.r > 0 \wedge \boldsymbol{B}_1(\neg\boldsymbol{K}_2Fair(silver) : r)$ by contradiction. Assuming that the opposite holds, i.e. there is a number $n > 0$ s.t. $e_1 \models \boldsymbol{B}_1(\neg\boldsymbol{K}_2Fair(silver) : n)$. By the definition of $\boldsymbol{O}_1^{(2)}$, f.a. $d \in \mathcal{D}^2$, $d \in e_1$ iff $\textsc{Norm}(d, \mathcal{W}^{1,e_1}_{\mathbf{KB}\wedge\delta}, \mathcal{W}_{\textsc{True}}, 1)$. Namely, $d \in e_1$ iff f.a. $w \in \mathcal{W}$ and $e_2 \in \mathcal{E}^1$, if $w, e_1, e_2 \not\models \mathbf{KB} \wedge \delta$ then $d(w, e_2) = 0$. We select two states $e'_2, e''_2 \in \mathcal{E}^1$ s.t.

$$e'_2 \models \neg\boldsymbol{K}_2Fair(silver) \wedge \forall x.[x \neq silver \supset \boldsymbol{K}_2Fair(x)]$$
$$e''_2 \models \forall x.\boldsymbol{K}_2Fair(x)$$

The choice of e'_2, e''_2 is arbitrary as long as the conditions are satisfied. Let w' be a world s.t. $w' \models \mathbf{KB}$, n' a number in $[0, 1]$ and $n' \neq n$. We define a 2-distribution d_1 as follows:

$$d_1(w, e_2) = \begin{cases} n' & w = w', e_2 = e'_2; \\ 1 - n' & w = w', e_2 = e''_2; \\ 0 & otherwise. \end{cases}$$

With the assumption that $e_1 \models \boldsymbol{B}_1(\neg \boldsymbol{K}_2 Fair(silver) : n)$, we conclude that $w' e_1, e_2' \models \mathbf{KB} \wedge \delta$ and $w' e_1, e_2'' \models \mathbf{KB} \wedge \delta$. Thus f.a. w, e_2, $d_1(w, e_2) = 0$ if $w, e_1, e_2 \not\models \mathbf{KB} \wedge \delta$, i.e. $d_1 \in e_1$. Base on the semantics $e_1 \models \neg \boldsymbol{B}_1(\neg \boldsymbol{K}_2 Fair(silver) : n)$, which contradicts the assumption.

Since sentence $\neg \exists r. r > 0 \wedge \boldsymbol{B}_1(\neg \boldsymbol{K}_2 Fair(silver) : r)$ is 1-subjective, $e_1 \models \boldsymbol{K}_1(\neg \exists r. r > 0 \wedge \boldsymbol{B}_1(\neg \boldsymbol{K}_2 Fair(silver) : r))$. With the distribution rule, we prove that $e_1 \models \boldsymbol{K}_1 \boldsymbol{K}_2 Fair(silver)$.

We also check the existence of such a model. Let e_1 be precisely the set of all distributions d_1 such that $d_1(w, e_2) = 0$ f.a. (w, e_2) which satisfies at least one of the conditions:

- $w \models \neg Coin(silver)$ or $e_2 \models \neg \boldsymbol{K}_2 Fair(silver)$;
- $w \models Coin(n)$ and $e_2 \models \neg \boldsymbol{K}_2 Fair(n)$ for $n \neq silver$.

By the semantics, it can be proved that $e_1 \models \boldsymbol{O}_1^{(2)}(\mathbf{KB} \wedge \delta)$. □

5 RELATION TO OTHER LOGICS

5.1 $\mathcal{OBL}$ is part of $\mathcal{OBL}_m$

$\mathcal{OBL}$ is a first-order, single-agent account of subjective probability and only-believing[6], which is an extension of the logic $\mathcal{OL}$ proposed by Levesque [23, 24]. For cases where at most one agent is involved, $\mathcal{OBL}_m$ reduces to $\mathcal{OBL}$—with exceptions: In $\mathcal{OBL}$, the empty epistemic state is legal, and distributions, where the weightings of worlds do not form a proper probability distribution, are not excluded. The former setting breaks the consistency, e.g. $\boldsymbol{K}\textsc{True} \wedge \boldsymbol{K}\textsc{False}$ is satisfiable in $\mathcal{OBL}$, and the latter breaks the necessitation rules, e.g. $\textsc{True}$ is valid in $\mathcal{OBL}$ but $\boldsymbol{K}\textsc{True}$ is not. We argue that omitting these structures in the semantics has no loss of expressiveness or generality. On the other hand, by the exclusion of these structures, the logic will obtain the complete $KD45_n$ properties, while only simple, straightforward restrictions are required. Formally, we have the following result:

Theorem 6. *Let $\models_{\mathcal{OBL}}$ denote the satisfaction relation in $\mathcal{OBL}$, for any well-formed $\mathcal{OBL}_m$ sentence α where no modalities appear except $\boldsymbol{B}_i$ and $\boldsymbol{O}_i^{(1)}$ for an agent i. Let α' be the sentence to replace every occurrence of $\boldsymbol{B}_i$ in α with $\boldsymbol{B}$, and replace $\boldsymbol{O}_i^{(1)}$ with $\boldsymbol{O}$, then*

$$\models \alpha \text{ iff } \boldsymbol{K}\textsc{True} \wedge \neg \boldsymbol{K}\textsc{False} \models_{\mathcal{OBL}} \alpha'$$

Examples include the properties discussed in Sections 3 and 4.

5.2 Relation to $\mathcal{ONL}_n$

$\mathcal{ONL}_n$ proposed by Belle and Lakemeyer[3] is a first-order modal language for multi-agent only-knowing, where operators $\boldsymbol{L}_i, \boldsymbol{N}_i$ are used to describing what agent i "at least knows" and "at most knows". Only-knowing is considered as an abbreviation, e.g. $\boldsymbol{O}_i \alpha \equiv \boldsymbol{L}_i \alpha \wedge \boldsymbol{N}_i \neg \alpha$. The k-distributions or k-epistemic states in this paper are inspired by their work. One might notice that in $\mathcal{ONL}_n$, the superscript (k) is not required for only-knowing. In their account, only-knowing can not precisely capture beliefs and non-beliefs of sentences with greater depth. For instance, neither $\boldsymbol{O}_1 p \supset \boldsymbol{K}_1 \boldsymbol{K}_2 q$ nor $\boldsymbol{O}_1 p \supset \neg \boldsymbol{K}_1 \boldsymbol{K}_2 q$ is valid in $\mathcal{ONL}_n$. We resolve this problem by specifying the depth of only-knowing with the superscript. For $k \geq 2$, $\boldsymbol{O}_1^{(k)} p$ means that p is all agent 1 knows up to depth k, where the depth for $\boldsymbol{K}_1 \boldsymbol{K}_2 q$ is also included. We have $\models \boldsymbol{O}_1^{(k)} p \supset \neg \boldsymbol{K}_1 \boldsymbol{K}_2 q$.

6 OTHER RELATED WORK

In the introduction, we have discussed related work about only-knowing and default reasoning. We provide more about notions related to only-knowing as well as unifying logic and probability.

The notion of only-knowing did not emerge in isolation. The idea of involving accessible and inaccessible worlds is also adopted by Humberstone [17], Ben-David and Gafni [7]. Characteristics of only-knowing are shared by similar concepts such as *total knowledge* from Pratt-Hartmann [30] and *minimal knowledge* from Halpern and Moses [15].

Numerous pieces of literature can be found in reasoning about knowledge and probabilities. For example the work from Nilsson [28]. First-order accounts of probabilities are discussed by both Bacchus [2] and Halpern [12]. For multi-agent scenarios, Fagin and Halpern proposed a framework based on Kripke semantics to reason about higher-order probabilities [11]. Their scheme was later extended to express common beliefs [32]. For the interest of AI, there are approaches combining first-order logic and probabilistic graphical models, such as first-order Belief network [29], Markov Logic Network [10], etc.

7 CONCLUSION

The main results of this work are as follows: We proposed a first-order modal logic for multi-agent only-believing, which for the first time generalizes Levesque's semantics to a multi-agent, probabilistic scenario. While $\mathcal{OBL}_m$ is downward compatible with $\mathcal{OBL}$ and $\mathcal{OL}$, the ability to capture the beliefs and non-beliefs of a knowledge base is lifted so that meta-beliefs involving many agents can also be considered. We also explored the capability of default reasoning. Compared with previous works, $\mathcal{OBL}_m$ is strictly more expressive.

Several aspects of future work are considered: We plan to augment the account to express common belief. Many of the previous works with similar designs on semantics support the possibility of such an extension. In particular, recent work from Cramer et. al.[9] shows that for a not fully-introspective account of categorical knowledge, combining common knowledge and only-knowing up to arbitrary depth can be achieved by at most $\omega^2 + 1$ layers of nesting in models.[4] Since the language includes first-order logic as a fragment, reasoning in $\mathcal{OBL}_m$ is in general undecidable. A feasible solution is to restrict the types of knowledge to be represented, or reconsider the form of reasoning [22]. We also intend to consider multi-agent only-believing in a dynamic system. Belle and Lakemeyer proposed an account for categorical knowledge [4], yet it was restricted to actions with fully observable, deterministic effects. For a multi-agent scenario with incomplete information, action taken by an agent may not be observable for another. Uncertainty in action should be modelled in a probabilistic form.

ACKNOWLEDGMENTS

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 2236/2 'UnRAVeL' and the EU ICT-48 2020 project TAILOR (No. 952215).

[4] ω is a *limit ordinal* number greater than all natural numbers. The notion of ordinal numbers is a generalization of natural numbers in set theory.

REFERENCES

[1] Guillaume Aucher and Vaishak Belle. 2015. Multi-agent only knowing on planet Kripke. In *International Joint Conference on Artificial Intelligence*.

[2] Fahiem Ivor Bacchus. 1989. Representing and reasoning with probabilistic knowledge. (1989).

[3] Vaishak Belle and Gerhard Lakemeyer. 2010. Multi-agent only-knowing revisited. In *Twelfth International Conference on the Principles of Knowledge Representation and Reasoning*.

[4] Vaishak Belle and Gerhard Lakemeyer. 2014. Multiagent only knowing in dynamic systems. *Journal of Artificial Intelligence Research* 49 (2014), 363–402.

[5] Vaishak Belle and Gerhard Lakemeyer. 2017. Reasoning about Probabilities in Unbounded First-Order Dynamical Domains.. In *IJCAI*. 828–836.

[6] Vaishak Belle, Gerhard Lakemeyer, and Hector Levesque. 2016. A first-order logic of probability and only knowing in unbounded domains. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 30.

[7] Shai Ben-David and Yael Gafni. 1989. All we believe fails in impossible worlds.

[8] Jianhua Chen. 1994. The logic of only knowing as a unified framework for non-monotonic reasoning. *Fundamenta Informaticae* 21, 3 (1994), 205–220.

[9] Marcos Cramer, Samuele Pollaci, and Bart Bogaerts. 2023. Mathematical Foundations for Joining Only Knowing and Common Knowledge. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning*, Vol. 19. 167–177.

[10] Pedro Domingos and Daniel Lowd. 2009. Markov Logic. In *Markov Logic: An Interface Layer for Artificial Intelligence*. Springer, 9–22.

[11] Ronald Fagin and Joseph Y Halpern. 1994. Reasoning about knowledge and probability. *Journal of the ACM (JACM)* 41, 2 (1994), 340–367.

[12] Joseph Y Halpern. 1990. An analysis of first-order logics of probability. *Artificial intelligence* 46, 3 (1990), 311–350.

[13] Joseph Y Halpern. 1993. Reasoning about only knowing with many agents. In *Proceedings of the eleventh national conference on Artificial intelligence*. 655–661.

[14] Joseph Y Halpern and Gerhard Lakemeyer. 2001. Multi-agent Only Knowing. *Journal of Logic and Computation* 11, 1 (2001), 41–70.

[15] Joseph Y Halpern and Yoram Moses. 1985. *Towards a theory of knowledge and ignorance: Preliminary report.* Springer.

[16] Gr E Hughes and MJ Cresswell. 1968. An introduction to modal logic, Methuen and Co. *Ltd., London* 70 (1968).

[17] IL Humberstone. 1986. A more discriminating approach to modal logic. *Journal of Symbolic Logic* 51, 2 (1986), 503–504.

[18] Gerhard Lakemeyer. 1993. All they know: A study in multi-agent autoepistemic reasoning. In *ijcai*, Vol. 93. 376–381.

[19] Gerhard Lakemeyer and Hector J Levesque. 2004. Situations, si! Situation terms, no!. In *KR*. 516–526.

[20] Gerhard Lakemeyer and Hector J Levesque. 2005. Only-knowing: Taking it beyond autoepistemic reasoning. In *AAAI*, Vol. 5. 633–638.

[21] Gerhard Lakemeyer and Hector J Levesque. 2011. A semantic characterization of a useful fragment of the situation calculus with knowledge. *Artificial Intelligence* 175, 1 (2011), 142–164.

[22] Gerhard Lakemeyer and Hector J Levesque. 2019. A Tractable, Expressive, and Eventually Complete First-Order Logic of Limited Belief.. In *IJCAI*. 1764–1771.

[23] Hector J Levesque. 1990. All I know: a study in autoepistemic logic. *Artificial intelligence* 42, 2-3 (1990), 263–309.

[24] Hector J Levesque and Gerhard Lakemeyer. 2001. *The logic of knowledge bases*. MIT Press.

[25] Daxin Liu and Qihui Feng. 2023. On the progression of belief. *Artificial Intelligence* 322 (2023), 103947.

[26] Daxin Liu and Gerhard Lakemeyer. 2021. Reasoning about Beliefs and Meta-Beliefs by Regression in an Expressive Probabilistic Action Logic.. In *ijcai*.

[27] Robert C Moore. 1985. Semantical considerations on nonmonotonic logic. *Artificial intelligence* 25, 1 (1985), 75–94.

[28] Nils J Nilsson. 1986. Probabilistic logic. *Artificial intelligence* 28, 1 (1986), 71–87.

[29] David Poole. 2003. First-order probabilistic inference. In *IJCAI*, Vol. 3. 985–991.

[30] Ian Pratt-Hartmann. 2000. Total knowledge. In *AAAI/IAAI*. 423–428.

[31] Riccardo Rosati. 2000. On the decidability and complexity of reasoning about only knowing. *Artificial Intelligence* 116, 1-2 (2000), 193–215.

[32] Siniša Tomović, Zoran Ognjanović, and Dragan Doder. 2015. Probabilistic common knowledge among infinite number of agents. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty: 13th European Conference, ECSQARU 2015, Compiègne, France, July 15-17, 2015. Proceedings 13*. Springer, 496–505.