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 intuitions 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 KD45n 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:


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 $OL$ [23, 24], the classical epistemic operator $K$ is used for knowledge, and in addition, a modality $O$ is used for only-knowing. To illustrate, $Op$ means that $p$ is all the agent knows. $Op$ entails $Kp$, but it also entails $\neg Kq$ and $\neg K\neg q$ for proposition $q$ different from $p$. This is different from classical epistemic logic where $Kp$ entails neither $\neg Kq$ nor $\neg K\neg q$. 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 $(\text{Bird(tweety)})$, namely “tweety is a bird”, $\delta$ a sentence to express the default $\forall x.[\text{Bird}(x) \land \neg K\neg \text{Fly}(x) \supset \text{Fly}(x)]$, 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 $OL$: $O(\text{KB} \land \delta)$ entails $\text{KFly(tweety)}$, but $O(\text{KB} \land \delta \land \neg \text{KFly(tweety)})$ entails $\neg \text{KFly(tweety)}$. 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 $OBL$ 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
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 of \( ONL_n \) 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 \( O_i \) denotes the only-knowing of agent 1, \( K_1 \) and \( K_2 \) stand for the belief of agent 1 and 2 respectively, then \( O_1 p \supset K_1(K_2 q) \) and \( O_1 p \supset \neg K_1(K_2 q) \) 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., \( OL \), 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 \( O^{(k)} \) 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 \( OBL_m \). 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 \( OBL_m \)

We introduce the logic \( OBL_m \). It can be considered as a multi-agent extension of \( OBL [6] \) and a probabilistic extension of \( ONL_n [3] \).

2.1 The Language

The logic \( OBL_m \) 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 \( N \), 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 \( Q \) 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 \( \land, \neg, \lor, \forall \) and modalities \( B_1, \ldots, B_m \) are used to construct formulae. For each agent index \( i, B_i(\alpha; r) \) is read as “agent i believes \( \alpha \) with degree \( r \)”. To illustrate, a sentence \( B_i(\text{Rain}; 0.8) \) means agent i believes that there is an 80% chance of raining. In particular, we write \( K_1 \alpha \) as an abbreviation of \( B_i(\alpha; 1) \) and read it as “agent i knows \( \alpha \)”.

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 \( O^{(1)} \) 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 \( O^{(1)} \)-formula is \( O^{(1)}_i(\exists x. [K_j(P(x)) \land \neg K_j(P(x))] \), 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 \( O^{(2)} \) 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

\[
O^{(2)}_i(q \land \exists x. [K_j(P(x)) \land \neg K_j(P(x))])
\]

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

We include modalities \( O^{(k)}_1, \ldots, O^{(k)}_m \) with natural number \( k > 0 \). For each \( k \), \( O^{(k+i)} \) specifies agent i’s only-believing about other agents’ beliefs up to depth \( k \), and formula \( O^{(k)}_i(\alpha; r_1, \ldots, r_i; r_j) \) is read as “All agent i believes up to depth \( k \) is: \( \alpha \) with degree \( r_1, \ldots, \alpha_1 \) with degree \( r_1 \)”. We write \( O^{(k)}_i(\alpha; 1) \) to mean \( O^{(k)}_i(\alpha; 1) \).

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

\[
\begin{align*}
dep[\alpha; i] = 0 &\text{ for atomic formula } \alpha, \text{ dep}[t_1 = t_2; i] = 0 \\
dep[\neg \alpha; i] = dep[\alpha; i] \\
dep[\alpha \land \beta; i] = \max\{dep[\alpha; i], dep[\beta; i]\} \\
dep[v \times \alpha; i] = dep[\alpha; i] \\
dep[\forall x. \alpha; i] = dep[\alpha; i] \\
dep[B_i(\alpha; r; i)] = \max\{\max\{dep[\alpha; j]; j \neq i\} + 1, dep[\alpha; i]\} \\
dep[B_i(\alpha; r; i)] = 0 &\text{ for } i \neq j \\
dep[O^{(k)}_i(\alpha; r_1, \ldots, r_i; r_j; i)] = k \\
dep[O^{(k)}_i(\alpha; r_1, \ldots, r_i; r_j; i)] = 0 &\text{ for } i \neq j
\end{align*}
\]

For a \( B \)-formula, the depth increases only when a nesting of different agents’ beliefs occurs. For an \( O \)-formula, the depth is given by the superscript \( k \).

Example 1. Let \( \alpha \) denote the formula \((\exists r. r > 0.5 \land B_2(p; r)) \). Consider the formula \( K_1 \alpha \), namely, agent 1 knows that \( p \) is believed by agent 2 with a degree greater than 0.5. Then

\[
\begin{align*}
dep[K_1 \alpha; 1] = \max\{\max\{dep[\alpha; 1], 1 + dep[\alpha; 2]\}\} \\
&= \max\{0, 2\} = 2 \\
dep[K_1 \alpha; 2] = 0
\end{align*}
\]

i.e. \( K_1 \alpha \) 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 \( \alpha, \beta \) are formulae, then \( \alpha \land \beta, \neg \alpha \) are formulae
- if \( \varphi \) is a formula, then for any \( i \in \{1, \ldots, m\} \) and number \( r \), \( B_j(\alpha: r) \) is a formula.
- for \( j \in \{1, \ldots, l\} \), \( B_j(\alpha_j: r_j) \) is a formula with \( i \)-depth \( k_j \), \( k \geq \max\{k_1, \ldots, k_l\} \), then \( O^i_j( k \lfloor a_1, \ldots, a_r \rceil ) \) is a formula.

An example of illegal formulae is \( O^1_j( p \land K_2 q: 0.8) \), 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 \( i \in \mathbb{N} \), a formula is called \( i \)-objective if modal operators \( B_j \) and \( O^i_j \) only appear within the scope of \( B_j \) or \( O^i_j \) when \( j \neq i \). For instance, \( p \land O^2_2(q \land K_2 p) \) is 1-objective since \( K_2 \) only appears in the scope of \( O^3_2 \), 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. Conversely, a formula is called \( i \)-subjective if any predicates or modal operators \( B_j \) or \( O^i_j( k \lfloor a \rceil ) \) s.t. \( j \neq i \) only appear within the scope of operators \( B_j \) or \( O^i_j( k \lfloor a \rceil ) \), e.g. \( \neg K_1 p \land B_1(p \land K_2 q: 0.5) \). We use \texttt{TRUE} to denote an objective tautology \( \forall x. (x = x) \) and use \texttt{FALSE} to denote its negation.

### 2.2 The Semantics

The semantics of \( \mathcal{OB}L_m \) is given in terms of possible worlds, where a \textit{world} is a set of ground atoms considered as true. The set of all worlds is denoted as \( \mathcal{W} \).

In the logic \( \mathcal{OB}L \), 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 \( \mathcal{OB}L \) 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:

- \( d^1 : (\mathcal{W} \times \{ \emptyset \} \times \cdots \times \{ \emptyset \}) \rightarrow \mathbb{R}^{\geq 0} \),
- \( d^{k+1} : (\mathcal{W} \times d^k \times \cdots \times d^k) \rightarrow \mathbb{R}^{\geq 0} \),

where \( d^k = 2^{d^k} \), and \( d^k \) 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 \textit{de re} knowledge about degrees of belief should not be valid. Namely, if epistemic states only contain single distributions, formulae such as \( \exists \phi. K(\phi: r) \) will be valid for any \( \phi \), 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 \textit{epistemic state}, specifically a \( k \)-epistemic state, if \( e \) is a set of \( k \)-distributions, i.e. \( e \in d^k \). The number \( k \) is said to be the depth of \( e \), written as \( \text{dep}(e) = k \).

Similar to \( \mathcal{OB}L \), we appeal to three conditions \( \mathcal{B} \), Eq, Norm to obtain probability distributions.

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

- \( \mathcal{B}(d, \mathcal{U}, r) \) iff there is no \( l \), \( \{(w_1, e_1), \ldots, (w_l, e_l)\} \subseteq \mathcal{U} \) s.t.\(^2\)
  \[ \sum_{j=1}^{l} d(w_j, e_j) > r \]
- \( \text{Eq}(d, \mathcal{U}, r) \) iff \( \mathcal{B}(d, \mathcal{U}, r) \) and no \( r' < r \) s.t. \( \mathcal{B}(d, \mathcal{U}, r') \);
- \( \text{Norm}(d, \mathcal{U}, \mathcal{V}, r) \) iff there is a number \( b \neq 0 \) such that \( \text{Eq}(d, (\mathcal{U}, b \times r)) \) and \( \text{Eq}(d, (\mathcal{V}, b)) \).

Intuitively, \( \mathcal{B}(d, \mathcal{U}, r) \) ensures the weight of tuples \((w, e)\) in \( \mathcal{U} \) is bounded by \( r \). \( \text{Eq}(d, \mathcal{U}, r) \) ensures that the weight is bounded and \( r \) is the supremum. Given \( \text{Norm}(d, \mathcal{U}, \mathcal{V}, r) \), \( r \) 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 \( \mathcal{V} = (\mathcal{W} \times d^k \times \cdots \times d^k) \) and \( \mathcal{U} = \{(w, e) \mid d(w, e) \neq 0\} \). For any \( b \geq 0 \), if \( \mathcal{B}(d, \mathcal{V}, b) \), then \( \mathcal{U} \) is countable.

**Proof.** Let \( U_j = \{(w, e) \in \mathcal{U} \mid d(w, e) \geq 1/j\} \) for \( j \in \mathbb{N}^+ \). It is easy to see that \( \mathcal{U} = \bigcup U_j \). Suppose that \( \mathcal{U} \) is uncountable, then there is some \( e > 0 \) such that \( U_e = \{(w \in \mathcal{U}) \mid d(w, e) \geq e\} \) is infinite (Otherwise we could enumerate \( \mathcal{U} \) by enumerating \( U_j \) starting at \( j = 1 \)). Consider any countably infinite sequence \((w_1, e_1)\) taken from \( U_e \). Since \( d(w_j, e_j) \geq e \) for all \( l \), the sum \( \sum_{j=1}^{\infty} d(w_j, e_j) \) is unbounded, contradicting the assumption that \( \mathcal{V} \) is bounded.

By satisfying the \( \mathcal{B} \) conditions, a distribution assigns non-zero values to a countable support set \( \text{Supp}(d) = \{(w, e) \mid d(w, e) \neq 0\} \). 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.

\[ \sum_{(w, e) \in \text{Supp}(d)} d(w, e) \]

In \( \mathcal{OB}L \), the necessitation rule does not hold for modality \( B \) (also \( K \)). For instance, \texttt{True} is valid in \( \mathcal{OB}L \) yet \( K \texttt{True} \) is not. Similarly, an empty epistemic state is legal in \( \mathcal{OB}L \), which breaks the laws of probability (e.g. \( B(p: 0.1) \land B(p: 0.2) \) 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 \textit{regular} \( k \)-distributions \( d^k \) and \textit{regular} \( k \)-epistemic states \( d^k \). Let \( d^0 = \{ \emptyset \} \). For any \( k > 0 \), let \( \mathcal{V} = (\mathcal{W} \times d^{k-1} \times \cdots \times d^0) \),

\(^2\)We write the arguments of a \( k \)-distribution as \((w, e)\) instead of \((w, e_1, \ldots, e_{m-1})\).
• $D^k = \{ d \mid E(d, \chi, 1) \}$; 
• $\mathcal{E}^k = 2^{D^k} \setminus \{ \emptyset \}$.

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 $(w, \tilde{e})$, where $\tilde{e} = (e_1, \ldots, e_n)$ and each $e_i$ is an epistemic state ($e_i$ denotes the $i$-th argument of $\tilde{e}$). We say a formula $\alpha$ and $\tilde{e}$ are compatible if $\text{dep}(\alpha, i) \leq \text{dep}(e_i)$ f.a. $i \in \{ 1, \ldots, m \}$. The truth value of objective sentences is assigned as follows:

• $w, \tilde{e} \models P(n)$ iff $P(n) \in w$; 
• $w, \tilde{e} \models t_1 = t_2$ iff $t_1$ and $t_2$ are identical standard name; 
• $w, \tilde{e} \models \neg \alpha$ iff not $w, \tilde{e} \models \alpha$; 
• $w, \tilde{e} \models \alpha \lor \beta$ iff $w, \tilde{e} \models \alpha$ or $w, \tilde{e} \models \beta$; 
• $w, \tilde{e} \models \forall \alpha$ iff $w, \tilde{e} \models \alpha_n$ for any $n \in N$.

Here $\alpha_n$ means the formula obtained by substituting each appearance of free variable $x$ in $\alpha$ by a standard name $n$. The semantics for the objective fragment is identical to language $L$ [24]. Supposing that $\tilde{e}$ is compatible with $B_1(\alpha: r)$, then

• $w, \tilde{e} \models B_1(\alpha: r)$ iff f.a. $d \in e_i$, $\text{Norm}(d, W^{\alpha}_d, W^{\alpha}_{\text{True}} \cdot r)$,

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

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

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

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

• For $e_i \in \mathcal{E}_k$, $w, \tilde{e} \models O^k_d(\alpha_1: r_1, \ldots, \alpha_i: r_i)$ iff f.a. $d \in D^k$, $d \models \alpha_i$,

we still need to assign truth values for only-believing up to depth $k$, only if the model is of a greater depth (otherwise sentences like $O^1_d(p \land O^2_d(p \land K_e p))$ 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 $d \in D^2$, $d' \in D^1$, $d'$ is the regression of $d$, written as $d' = R[d]$ if for any $w$,

$$ \sum_{e_1, \ldots, e_m} d(w, e_1, \ldots, e_m) = d'(w, \emptyset, \ldots, \emptyset) $$

• For $e \in \mathcal{E}^2$, $e', e'' \in \mathcal{E}^1$, we say $e'$ is the regression of $e$, written as $e' = R[e]$, iff $e' = \{ R[d] \mid d \in e \}$.

• For $d \in D^{k+1}$, $d' \in D^k$, $d'$ is the regression of $d$, written as $d' = R[d]$, if for any $w, e_1', \ldots, e_{m-1}'$,

$$ \sum_{e_1' \ldots, e_m - 1} d(w, e_1', \ldots, e_m) = d'(w, e_1', \ldots, e_{m-1}') $$

where $E = \{ e_1, \ldots, e_m \mid e'_i = R[\tilde{e}^i], \ldots, e'_{m-1} = R[\tilde{e}^m] \}$.

• For $e \in \mathcal{E}^{k+1}$, $e' \in \mathcal{E}^k$, we say $e'$ is the regression of $e$, written as $e' = R[e]$, iff $e' = \{ R[d] \mid d \in e \}$.

By definition, for each $e \in \mathcal{E}^{k+1}$, there exists a unique $e' \in \mathcal{E}^k$ s.t. $e' = R[e]$. We write $e' = R^{(\geq)}[e]$ to mean $e' = R[R[e]]$. Analogously, $e' = R^{(k)}[e]$ means $e' = R[R[R[\cdots R[e]]]]$ with $k$ nested $R$. The following lemma indicates that $R[e]$ 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 $R[e]$.

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

**Proof.** We prove the lemma via induction.

**Basis:**

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

**Induction hypothesis:**

Suppose that the statement holds for any $\tilde{e}, \tilde{e}', \alpha'$ s.t. $e' = R[e]$,

• $e' \in \mathcal{E}^k$ for some $l < k$ and $\alpha'$ compatible with $e'$.

• Induction on $\land$, $\lor$ or $\forall$ is trivial.

• For $e_i \in \mathcal{E}^{k+1}$, $e_i \models B_1(\alpha: r)$ iff f.a. $d \in D^3$,

$$ \sum_{(w, \tilde{e} - i) \in W^{\alpha}_{e_i}} d(w, \tilde{e} - i) = r \tag*{(#)} $$

We introduce an auxiliary function for the proof:

For $\tilde{e} - i = (e_1', \ldots, e_{i-1}', e_{i+1}', \ldots, e_m')$, $e_i' \in \mathcal{E}^k$, and $\tilde{e}'' - i = (e_1'', \ldots, e_{i-1}'', e_{i+1}'', \ldots, e_m'')$, $e_i'' \in \mathcal{E}^{k-1}$, we define

$$ I(\tilde{e} - i, \tilde{e}'' - i) = \begin{cases} 1 & \text{if } e_i'' = R[e_i'] \text{ f.a. } i \neq i; \\ 0 & \text{otherwise}. \end{cases} \tag*{(*)} $$

Since the regression is unique, given $\tilde{e} - i$, there’s only one tuple $\tilde{e}'' - i$ with $I(\tilde{e} - i, \tilde{e}'' - i) = 1$. Thus when $\tilde{e} - i$ is fixed,

$$ \sum_{\tilde{e}'' - i} I(\tilde{e} - i, \tilde{e}'' - i) = 1 \tag*{(*)} $$

Back to (#), let $d'' = R[d]$, we have

$$ r = \sum_{(w, \tilde{e} - i) \in W^{\alpha}_{e_i}} d(w, \tilde{e} - i) \cdot 1 \tag*{(*)} $$

$$ = \sum_{(w, \tilde{e} - i) \in W^{\alpha}_{e_i}} \left( d(w, \tilde{e} - i) \cdot \sum_{\tilde{e}'' - i} I(\tilde{e} - i, \tilde{e}'' - i) \right) \tag*{Eq. (*)} $$

$$ = \sum_{(w, \tilde{e} - i) \in W^{\alpha}_{e_i}} \sum_{\tilde{e}'' - i} d(w, \tilde{e}'' - i) \cdot I(\tilde{e} - i, \tilde{e}'' - i) \tag*{(*)} $$

$$ = \sum_{(w, \tilde{e} - i)} d(w, \tilde{e} - i) $$

$$ = \sum_{(w, \tilde{e} - i) \in W^{\alpha}_{e_i}} d(w, \tilde{e} - i) \tag*{(*)} $$

$$ \sum_{\tilde{e}'' - i} I(\tilde{e} - i, \tilde{e}'' - i) $$

\(^{3}\text{By definition, the truth of } B_1(\alpha: r) \text{ is irrelevant to } w \text{ and epistemic states of other agents except } \iota. \text{ Thus we write } e_i \models B_1(\alpha: r) \text{ instead of } w, \tilde{e} \models B_1(\alpha: r). \)
In this section, we study the properties of modalities $OBL$.

Proof. $∀ w, e_i’ = R[e_i]$ if and only if $w \models e_i’$. By induction hypothesis, we have $w, e_i’ \vdash e_i’ \models α$, i.e. $(w, e_i’ \vdash ∈ W_α^{e_i’})$. Therefore,

$$r = \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash)$$ (IH)

$$= \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} \left( \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) \right)$$ (Def. 5)

$$= \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d'(w, e_i’ \vdash)$$

Thus $\text{Norm} (d', W_α^{e_i’}, W_α^{e_i’}; r)$ f.a. $d’ ∈ e_i’$ i.e. $e_i’ \models B_1 (x; r) \Box$

Now we can complete the semantics:

- For $e_i ∈ O^{(k)}$ st. $k’ > k$, $w, e_i \models O^{(k)}_i (α_1, . . . , α_k, α_{k+1})$ iff $w, e_1, . . . , R[e_1], . . . , e_m \models O^{(k)}_i (α_1, . . . , α_{k}, α_{k+1})$.

The result of Lemma 1 can be extended to any wfs, including those with only-believing:

**Theorem 2.** Given $\vec{e}, \vec{e’}$ s.t. $e_i’ = R[e_i]$. For any formula $α$ compatible with both $\vec{e}, \vec{e’}$, then $\vec{e} ⊨ α$ if $w, e’ ⊨ α$.

The proof is similar to Lemma 1. The only difference is the induction for $O^{(k)}$-formulae, which directly follows from the definition of the semantics.

For a sentence $α$ and a set of sentences $Σ$, we write $Σ \models α$ (read as $Σ$ logically entails $α$) to mean that for every model $w, e_1, . . . , e_m$ compatible with $α$ and all sentences in $Σ$, when $w, e_1, . . . , e_m \models α$.

We say $α$ is valid (written as $\models α$) if $\{\} \models α$. When $α$ is objective, we write $w \models α$ instead of $w, e_1, . . . , e_m \models α$. When $α$ is subjective, we write $e_1, . . . , e_m \models α$ (or $\vec{e} \models α$). For $i$-subjective $α$, we write $e_i \models α$.

3 PROPERTIES OF THE LOGIC

In this section, we study the properties of modalities $B_1$ and $K_i$. For $K_i$, where the probabilistic belief is reduced to a categorical one, $\mathcal{OBL}_m$ satisfies the $KD45$ properties. The Barcan formulae[16] are also valid. For $B_1$, 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 $KD45$ properties:

- (Necessitation) If $\models α$, then $\models K_i α$.
- (Consistency) $\models K_i α → ¬ K_i ¬ α$.

Proof. Suppose that $e_i \models K_i α$, then for any $d ∈ e_i$:

$$\sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 1 \implies \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 0$$

Thus $e_i \models ¬ K_i ¬ α$.

- (Distribution) $\models K_i α → K_i (α → β) → K_i β$.

Proof. For any $e_i$ compatible with $K_i α$ and $K_i (α → β)$, if $e_i \models K_i α$ and $K_i (α → β)$, then for $d ∈ e_i$:

$$\sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) + \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 1 + 1 = 1$$

Thus $e_i \models K_i β$.

- (Pos. Introspection) $\models K_i α → K_i K_i α$.
- (Neg. Introspection) $\models ¬ K_i α → K_i ¬ K_i α$.

Proof. Suppose that $e_i \models ¬ K_i α$, then f.a. $w$ and $e_i’ \vdash$, $w, e_i, . . . , e_i’ \vdash$, $e_i, . . . , e_i’ \vdash, . . . , e_m \models ¬ K_i α$. Thus f.a. $d ∈ e_i$, $\sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 1$ i.e. $e_i \models K_i ¬ K_i α$.

Barcan formulae (both universal and existential versions):

- $\models ∀ x K_i x → K_i ∀ x α$.

Proof. Suppose that $e_i \models ∀ x K_i x$. By definition we have $e_i \models K_i α_a$ for any $a ∈ N$. Then f.a. $d ∈ e_i$:

$$\sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 1 \forall a ∈ N$$

Namely,

$$\sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 0 \forall a ∈ N$$

Therefore

$$\sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 1 = \sum_{(w, e_i’ \vdash ∈ W_α^{e_i’})} d(w, e_i’ \vdash) = 1$$

Thus $e_i \models K_i ∀ x α$.

It is worth mentioning that the converse result also holds, i.e. $\models K_i ∀ x α → ∀ x K_i α$.

- $\models ∀ x K_i x → K_i ∀ x α$.

Proof. The proof is of the same spirit as the universal one. $\models K_i x → K_i ∃ x α$.

Let $w, w’ ∈ W$ satisfy $P(n_1) ∈ w, P(n’1) ∈ w$ f.a. $n’1 ≠ n_1$, $P(n_2) ∈ w’, P(n’2) ∈ w’$ f.a. $n’2 ≠ n_2$. Let $d_i ∈ D^3$ assign weight 0.5 to both $w$ and $w’$. Other worlds are assigned 0. Then $d_i$ satisfies $K_i ∃ x α$ but not $∃ x K_i α$. 


3.2 Degree of Belief

For any $i \in Ag$ and formula $\alpha, \beta$,
- if $\alpha$ is valid, then $\models B_i(\alpha : 1)$
- $\models B_i(\alpha : r) \supset \neg B_i(\alpha : r')$ for $r' \neq r$
- if $\alpha \equiv \beta$, then $\models B_i(\alpha : r) \equiv B_i(\beta : r)$ for any $r$. 
- $\models B_i(\alpha : r) \supset B_i(\beta : r)$ for any $\beta$.
- $\models B_i(\alpha \land \beta : r) \supset B_i(\alpha \land \beta : r')$ for \( r' > r \)
- $\models B_i(\alpha : r) \land B_i(\beta : r') \supset B_i(\alpha \lor \beta : r')$
- 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 $e_i = B_i(\alpha : r) \land B_i(\beta : r') \land B_i(\alpha \land \beta : r'')$, then f.a. $d \in e_i$, it holds $\sum_{(w, e'_{-i}) \in W_{α}\beta} d(w, e'_{-i}) = r$, $\sum_{(w, e'_{-i}) \in W_{αi}} d(w, e'_{-i}) = r'$, $\sum_{(w, e'_{-i}) \in W_{αi\beta}} d(w, e'_{-i}) = r''$.

By applying the laws of set,
\[
W_{αi\beta} = W_{αi} + W_{αi\beta} - W_{αiβ}
\]

Thus $\sum_{(w, e'_{-i}) \in W_{αi\beta}} d(w, e'_{-i}) = r + r' - r''$. 

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

- $\models B_i(\alpha : r) \supset K_i B_i(\alpha : r)$
- $\models \neg B_i(\alpha : r) \supset K_i \neg B_i(\alpha : r)$
- For any i-subjective formula $\alpha, \models \alpha \equiv K_i \alpha$

**Proof.** Suppose that $\alpha$ is i-subjective and $e_i = \alpha$. Since $\alpha$ 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 $d \in e_i$,
\[
\sum_{(w, e'_{-i}) \in W_{αi\beta}} d(w, e'_{-i}) = \sum_{(w, e'_{-i}) \in W_{αi}} d(w, e'_{-i}) = 1
\]

Thus $e_i \models K_i \alpha$.

\[\square\]

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.** $\models O_i^{(k)}(\alpha_1 : r_1, \ldots, \alpha_l : r_l) \supset \bigwedge_{j=1}^l B_i(\alpha_j : r_j)$

If an agent only believes $\alpha$ 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.** $\models O_i^{(k+1)}(\alpha : r) \supset O_i^{(k)}(\alpha : r)$.

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

\[\square\]

576
Since $\alpha$ is $i$-objective, if $w_i, e'_1, \ldots, e'_m \models \alpha$, then $w_i, e'_1, \ldots, e'_m \models \alpha$. For any $e_i \in E_k$ s.t. $e_i \models O^1(\alpha)$, it is trivial that $d \in D_k$ if $d \models e_i$, i.e. $e_i = e_i$.

For a model of a depth greater than needed, the uniqueness of the model can be interpreted as follows: Suppose $e_i \in \mathcal{E}_k$ is the unique model satisfying $O^i_1(\alpha)$. Then for all $e'_i \in \mathcal{E}_k$ s.t. $k' > k$ and $e'_i \models O^i_1(\alpha)$, $e_i$ will reduce to $e_i$ after finite steps of regression.

**Lemma 3.** Given $i$-objective $\alpha$ and $k$ s.t. $K_i \alpha$ is satisfiable and $O^i_1(\alpha)$ is a wff. Let $e_i \in \mathcal{E}_k$ be the model s.t. $e_i \models O^i_1(\alpha)$. For any $e'_i$ such that $e'_i \models O^i_1(\alpha)$, $e_i = R^1(\mathbb{R})$ for some $l$.

**Proof.** For any $k' > k$ and $e'_i \in \mathcal{E}_k$, $e'_i$ reduces to $e_i \in \mathcal{E}_k$ after $k' - k$ steps of regression. Given $e'_i \models O^i_1(\alpha)$, by applying Thm. 2 we have $e_i \models O^i_1(\alpha)$. By Lem. 2, $e_i$ is unique, i.e. $e_i = e_i$.

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

**Theorem 5.** Given $i$-objective formulae $\alpha_1, \ldots, \alpha_l$ s.t. $\bigwedge_j B_1(\alpha_j; r_j)$ is satisfiable and $O^i_1(\alpha_1)$ is a wff, there is precisely one $e_i \in \mathcal{E}_k$ such that $e_i \models O^i_1(\alpha_1) \land \ldots \land O^i_1(\alpha_l)$.

**Proof.** By Thm.3, it exists a unique $e_i \in E_k$ such that $e_i \models O^i_1(\alpha_1, \ldots, \alpha_l, r_i)$. If $\bigwedge_j B_1(\alpha_j; r_j) \models B_i(\beta; r')$, it is valid, by Prop.1, $\models O^i_1(\alpha_1, \ldots, \alpha_l, r_i) \models B_i(\beta; r)$.

For any $e'_i \in \mathcal{E}_k$ s.t. $k' > k$, $e'_i \models O^i_1(\alpha_1, \ldots, \alpha_l, r_i)$. By Thm. 2 and 3, $e'_i \models B_i(\beta; r')$.

**Corollary 1.** Given $i$-objective $\alpha, \beta$, s.t. $O^i_1(\alpha)$ is satisfiable, then

- $\models O^i_1(\alpha) \models \mathbb{K}_1 \beta$ if $\alpha \models \beta$
- $\models O^i_1(\alpha) \models \lnot \mathbb{K}_1 \beta$ if $\alpha \not\models \beta$

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 $Fair$ ($\text{Fair}(x)$) mean "(The coin) $x$ is fair". Let $\delta$ denote the sentence

$$\forall x. [\text{Coin}(x) \land \lnot \exists r. (r > 0 \land B_1(\lnot K_2 \text{Fair}(x): r) \models K_2 \text{Fair}(x))]$$

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

- $O^1(\text{KB} \land \delta) \models K_1 K_2 \text{Fair(silver)}$
- $O^1(\text{KB} \land \delta \land K_2 \text{Fair(silver)}) \models K_1 K_2 \text{Fair(silver)}$
- $O^1(\text{KB} \land \delta \land \lnot K_2 \text{Fair(silver)}) \models K_1 \lnot K_2 \text{Fair(silver)}$

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

**Proof.** Let $e_1 \models O^1(\text{KB} \land \delta)$, w.l.o.g. $e_1 \models E_2$. We first prove $e_1 \models \lnot \exists r. (r > 0 \land B_1(\lnot K_2 \text{Fair(silver)}: r))$ by contradiction. Assuming that the opposite holds, i.e. there is a number $n > 0$ s.t. $e_1 \models B_1(\lnot K_2 \text{Fair(silver)}: n)$. By the definition of $O^1(\text{KB})$, f.a. $d \in D^2$, $e_1 \models \text{Norm}^1(d, W_{K_1 \text{KB}} \circ W_{\text{Thres}^1})$. Namely, $d \in e_1$ iff $f.a. w \in W_{\text{KB}} \wedge e_2 \models E_1$. If $w \in W_{\text{KB}}$, we select two states $e'_2, e''_2 \in E_2$ if $e_1 \models B_i(\beta; r')$.
With the assumption that \( e_1 \models B_1(\neg K_2 Fair(silver) : n) \), we conclude that \( w', e'_1 \models KB \land \delta \) and \( w', e'_2 \models KB \land \delta \). Thus \( w, e_2, d_1|w, e_2| = 0 \) if \( w, e_1, e_2 \not\models KB \land \delta \). i.e., \( d_1 \in e_1 \). Base on the semantics \( e_1 \models \neg B_1(\neg K_2 Fair(silver) : n) \), which contradicts the assumption.

Since sentence \( \neg \exists r . r > 0 \land B_1(\neg K_2 Fair(silver) : r) \) is 1-subjective, \( e_1 \models K_1(\neg \exists r . r > 0 \land B_1(\neg K_2 Fair(silver) : r)) \). By the distribution rule, we prove that \( e_1 \models K_1 K_2 Fair(silver) \).

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 \( e_2 \models \neg K_2 Fair(silver) \);
- \( w \models Coin(n) \) and \( e_2 \models \neg K_2 Fair(n) \) for \( n \neq silver \).

By the semantics, it can be proved that \( e_1 \models O_1^{(k)}(KB \land \delta) \). □

5 RELATION TO OTHER LOGICS

5.1 \( OBL \) is part of \( OBL_m \)

\( OBL \) is a first-order, single-agent account of subjective probability and only-believing[6], which is an extension of the logic \( OBL \) proposed by Levesque [23, 24]. For cases where at most one agent is involved, \( OBL_m \) reduces to \( OBL \)—with exceptions: In \( OBL \), 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. \( K_1^{\text{True}} \land K_1^{\text{False}} \) is satisfiable in \( OBL \), and the latter breaks the necessitation rules, e.g. \( \Box r \) is valid in \( OBL \) but \( K_1^{\text{True}} \) 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 KD45n properties, while only simple, straightforward restrictions are required. Formally, we have the following result:

**Theorem 6.** Let \( \models_{OBL} \) denote the satisfaction relation in \( OBL \), for any well-formed \( OBL_m \) sentence \( \alpha \) where no modalities appear except \( B_i \) and \( O_1^{(i)} \) for an agent \( i \). Let \( a' \) be the sentence to replace every occurrence of \( B_j \) in \( a \) with \( B_j \), and replace \( O_1^{(i)} \) with \( O_1^{(j)} \), then

\[ a \ \text{iff} \ K_1^{\text{True}} \land \neg K_1^{\text{False}} \models_{OBL} \ a' \]

Examples include the properties discussed in Sections 3 and 4.

5.2 Relation to \( ONL_n \)

\( ONL_n \) proposed by Belle and Lakemeyer [3] is a first-order modal language for multi-agent only-knowing, where operators \( L_i, N_i \) are used to describing what agent \( i \) ”at least knows” and ”at most knows”. Only-knowing is considered as an abbreviation, e.g. \( O_1^a \equiv L_i a \land N_i \neg a \). The \( k \)-distributions or \( k \)-epistemic states in this paper are inspired by their work. One might notice that in \( ONL_n \), 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 \( O_1^p \supset K_1K_2q \) nor \( O_1^p \supset \neg K_1K_2q \) is valid in \( ONL_n \). We resolve this problem by specifying the depth of only-knowing with the superscript. For \( k \geq 2 \), \( O_1^{(k)}p \) means that \( p \) is all agent knows up to depth \( k \), where the depth for \( K_1K_2q \) is also included. We have \( \models O_1^{(k)}p \supset \neg K_1K_2q \).

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 \( OBL_m \) is downward compatible with \( OBL \) and \( OBL \), 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, \( OBL_m \) 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 \( o^3 + 1 \) layers of nesting in models.\(^4\) Since the language includes first-order logic as a fragment, reasoning in \( OBL_m \) 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).\(^4\)\text{\(\omega\) 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


