Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020)
Main Track

A First-Order Logic of Limited Belief Based on Possible Worlds

Gerhard Lakemeyer[1] and **Hector J. Levesque**[2]
[1] Dept. of Computer Science, RWTH Aachen University, Germany
[2] Dept. of Computer Science, University of Toronto, Canada
gerhard@cs.rwth-aachen.de, hector@cs.toronto.edu

Abstract

In a recent paper Lakemeyer and Levesque proposed a first-order logic of limited belief to characterize the beliefs of a knowledge base (KB). Among other things, they show that their model of belief is expressive, eventually complete, and tractable. This means, roughly, that a KB may consist of arbitrary first-order sentences, that any sentence which is logically entailed by the KB is eventually believed, given enough reasoning effort, and that reasoning is tractable under reasonable assumptions. One downside of the proposal is that epistemic states are defined in terms of sets of clauses, possibly containing variables, giving the logic a distinct syntactic flavour compared to the more traditional possible-world semantics found in the literature on epistemic logic. In this paper we show that the same properties as above can be obtained by defining epistemic states as sets of three-valued possible worlds. This way we are able to shed new light on those properties by recasting them using the more familiar notion of truth over possible worlds.

1 Introduction

Ever since the idea of a knowledge-based system was first introduced by McCarthy (1963), it has been a challenge to come up with a model of belief that characterizes the conclusions such systems should be able to draw from a given knowledge base (KB). Perhaps the main issue has been to combine a high degree of expressiveness with a notion of tractable reasoning. In a recent paper Lakemeyer and Levesque (Lakemeyer and Levesque 2019) (LL) proposed a model of belief, which is perhaps the most successful to date in addressing these and other concerns. They achieve expressiveness in the sense that knowledge bases may contain arbitrary sentences in first-order logic. Tractability is achieved by considering a notion of mental effort, characterized by a single parameter $k = 0, 1, \dots$ and corresponding belief operators $\boldsymbol{B}_k$. The idea is that $\boldsymbol{B}_0$ captures obvious beliefs, which include all the sentences in the KB and some easy logical consequences. When k increases, $\boldsymbol{B}_k$ would include more and more logical consequences of the KB, requiring more and more effort to be computed. Nevertheless, LL show that, for any fixed k and under some reasonable assumptions, the reasoning task remains tractable. Moreover, LL show that any logical consequence of the KB is believed at some level k.

In their work LL also consider a notion of *only-believing* using a modal operator $\boldsymbol{O}$, where $\boldsymbol{O}\phi$ is intended to mean that all that is believed at level 0 is ϕ. They then characterize the beliefs at any level k in terms of the valid sentences in the logic of the form $\boldsymbol{O}\text{KB} \supset \boldsymbol{B}_k\phi$. The main contributions of LL can be summarized as follows:

- expressiveness: for any sentence ϕ, $\boldsymbol{O}\phi$ is satisfiable, and moreover $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_0\phi)$;
- cumulativity: for any k and any ϕ $\models (\boldsymbol{B}_k\phi \supset \boldsymbol{B}_{k+1}\phi)$.;
- soundness: for any k, any KB and ϕ, if $\models (\boldsymbol{O}\text{KB} \supset \boldsymbol{B}_k\phi)$, then $\models (\text{KB} \supset \phi)$;
- eventual completeness: for any KB and any ϕ, if $\models (\text{KB} \supset \phi)$, then for some k $\models (\boldsymbol{O}\text{KB} \supset \boldsymbol{B}_k\phi)$;
- tractability: for any k, KB, and α, the question as to whether $\models (\boldsymbol{O}\text{KB} \supset \boldsymbol{B}_k\alpha)$ is decidable (and has polynomial data complexity in cases of interest).

While these are all very desirable properties, which no other model of limited belief exhibits (see the related-work section for more details), a downside of LL's logic is that it has a distinct syntactic flavour. In particular, an epistemic state is defined as a (finite) set of clauses C, which may mention free variables. Moreover, believing an arbitrary sentence at level k is defined, roughly, by reducing it to believing clauses, which in turn are obtained by k resolution steps from C. While there is nothing wrong with this technically, we find this approach not very compelling from a semantic point of view. Beginning with Hintikka's seminal work on modelling belief using Kripke structures (Kripke 1959; Hintikka 1962), the best understood models of belief have been those based on possible worlds. In the simplest case, when an epistemic state consists of a set of possible worlds, belief is defined in terms of truth in all those worlds.

In this paper, we will show that a close variant of LL's logic, satisfying all the desired properties mentioned above, can be obtained using sets of worlds as epistemic states. As we will see, many properties of belief such as belief equivalences will come out as semantic properties of worlds and not, as in the case of LL, as a result of normal-form transformations of sentences. The worlds themselves will have to be non-standard, as it is well known that classical (two-valued) worlds lead to so-called *logical omniscience* (Hintikka 1975), and hence undecidable reasoning in the first-order case. The main contribution of the paper is not a new logic, since we are mainly concerned with reestablishing the properties of LL's logic under a new semantics. Instead, by defining belief in terms of truth over possible worlds, which has a long tradition in epistemic logic, the main contribution is an alternative view of the valid sentences of the existing logic.

The rest of the paper is organized as follows. In the next section, we will go over our approach, mainly informally. This is followed by a brief review of the logic $\mathcal{L}$, a variant of first-order logic (Levesque and Lakemeyer 2001), which forms the basis of both LL's and our proposal. We then turn to the new logic in detail and prove that it satisfies the high-level properties listed above: expressiveness, cumulativity, tractability, soundness and eventual completeness, as well as certain desirable belief equivalences. The paper ends with related work and conclusions.

2 The Approach

In order to motivate our approach to modelling belief, it is instructive to consider how intractability comes about in classical logical reasoning. More precisely, if we think of how intractability arises in trying to determine whether or not KB logically entails ϕ, there are really two sources:

1. It can be too hard to make full use of the information provided by the KB.
2. It can be too hard to see if ϕ should be believed because of its own properties.

For the first item, consider, for example, a KB consisting of a set of ground clauses and where ϕ is some ground atom p. Determining whether KB $\models \phi$ in this case is the same as determining if KB $\cup \{\neg p\}$ is unsatisfiable. This task is co-NP-complete and the best known algorithms would perhaps use SAT-solvers to tackle this particular problem. In the first-order case one may need to resort to other methods like Resolution, but then the problem is already undecidable.

For the second item, consider the case where the KB is empty. Determining whether KB $\models \phi$ in this case is the same as determining if ϕ is logically valid. In the propositional case, this is not too hard when ϕ is small (relative to the size of the KB): we can convert ϕ to *CNF* and ensure that each resulting clause is a tautology. But for the full language with quantifiers, the task is undecidable in general.

To deal with these two items, we will be proposing a new model of belief with two separate mechanisms to keep the reasoning tractable. For the first item above, we will introduce epistemic states as sets of what we will call "extended" worlds; for the second item, we will preprocess the KB and the query ϕ using Skolemization and term substitution. The exact details will be presented beginning in the next section, but here is an informal outline of those ideas. The novelty of this paper compared to LL lies in the use of sets of extended worlds instead of sets of clauses. For the second item, we essentially follow what LL have done.

If we think of an epistemic state of a KB as the set of all worlds satisfying KB, it is easy to see that we need to go beyond classical two-valued worlds in order to obtain tractability. The problem, in essence, is that such epistemic states are too coarse in that they lump all logically equivalent knowledge bases together. For example, for KB $= \{p, (p \supset q)\}$, we might want an epistemic state where $\boldsymbol{B}_0(q)$ is false, but for the logically equivalent KB $= \{p, (p \supset q), q\}$, we want a different epistemic state where $\boldsymbol{B}_0(q)$ is true.

Here we will be using a finer-grained notion of epistemic state based on sets of *extended* worlds. An extended world will be defined as one where atomic sentences are mapped to one of three values, $\{0, 1, *\}$. A world that assigns p to $*$ supports both the truth and the falsity of p. Such a world will then be able to support the truth of both p and of $(p \supset q)$ without also supporting the truth of q. So the epistemic state e_1 made up of all extended worlds where p and $(p \supset q)$ are supported is a superset of an e_2 where q is also supported. In this way, in e_1 we can end up believing p and $(p \supset q)$ without believing q, whereas in e_2 we will end up believing these and q as well.

In going from belief at level k to belief at level $k+1$, we will end up moving from an epistemic state e to another one, $S(e)$, that has fewer extended worlds where we end up believing more sentences. As we will see, the idea is to eliminate some of the worlds where an atom is assigned $*$. In the case of e_1 above, we will end up eliminating the worlds where p is assigned $*$, which means that $S(e_1)$ will be e_2. More generally, if the epistemic state is the set of all extended worlds that support $\{(p \vee q), (\neg p \vee r), (\neg q \vee r)\}$, then the clauses $(p \vee q)$ and $(s \vee \neg s)$ and their supersets will be believed at level 0, the clause $(q \vee r)$ and its supersets will believed at level 1 (after one application of S), and finally, the clause r and its supersets will be believed at level 2 (after two applications of S).

This idea of epistemic states as sets of extended worlds works fine for the quantifier-free part of the language, but it cannot be the whole story. To see why, note that logically valid sentences are supported by every extended world. This means that even in the least informed epistemic state, the one made up of all extended worlds, we would be required to believe all valid sentences of first-order logic at level 0 (making belief at level 0 undecidable in the first-order case).

To avoid this, the first-order part of the semantics of belief goes further. As we will see in the next section, the semantics of $\boldsymbol{O}$ uses Skolemization to eliminate existential variables, and the semantics of $\boldsymbol{B}_k$ uses the dual of Skolemization (also called Herbrandization) to eliminate universal variables, and then a bounded form of term substitution to produce a sentence with no quantifiers. We therefore use the following reductions:

1. $\boldsymbol{O}\phi$ will hold iff $\boldsymbol{O}\forall\vec{x}.\psi$ holds, where the formula ψ is a Skolemized version of ϕ with no quantifiers;
2. $\boldsymbol{B}_k\phi$ will hold iff there are terms $\vec{t_0}, \ldots, \vec{t_k}$ such that $\boldsymbol{B}_k(\psi^{\vec{x}}_{\vec{t_0}} \vee \cdots \vee \psi^{\vec{x}}_{\vec{t_k}})$ holds, where the formula ψ is a dual-Skolemized version of ϕ with no quantifiers;

The second item here is a bounded application of what is known as Herbrand's Theorem, a way of going from unsatisfiability in classical first-order logic (what we will henceforth call *fo-unsatisfiability*) to its propositional counterpart. The theorem is as follows:

Theorem 1 [Herbrand] *Let Φ be a set of formulas with no quantifiers. If the set Φ is fo-unsatisfiable (with the free variables interpreted universally) then so is some finite subset of* $\{\, \phi^{\vec{x}}_{\vec{t}} \mid \phi \in \Phi$ *and* $\vec{t}$ *is ground* $\}$.

As a special case (mirroring item 2 above), we have the following:

Corollary 1 *If ϕ is fo-valid, then for some n, there are ground terms $\vec{t_1}, \ldots, \vec{t_n}$ such that $(\psi^{\vec{x}}_{\vec{t_1}} \vee \cdots \vee \psi^{\vec{x}}_{\vec{t_n}})$ is fo-valid, where ψ is a dual-Skolemized version of ϕ.*

To see how these reductions avoid the problem of having to believe all classically valid sentences, consider for example $\exists x \forall y(P(y) \vee \neg P(x))$. This sentence is valid in first-order logic and so is supported by all extended worlds, as is $\exists x.\psi$, where ψ is its dual-Skolemized version, $(P(f(x)) \vee \neg P(x))$. However, there is no single term t such that ψ^x_t is supported by all extended worlds. Because of this, $\boldsymbol{B}_0 \exists x.\psi$ need not be true, according to the reduction above. (However, $\boldsymbol{B}_1 \exists x.\psi$ will be true in this case since there are two terms t and u such that $(\psi^x_t \vee \psi^x_u)$ is supported by all extended worlds, namely $t = a$ and $u = f(a)$. Other fo-valid sentences will require higher levels of belief.)

But having made this move to term substitution in beliefs, we need to do something related in the KB using Skolemization. Consider the sentence $\exists x.P(x)$. This will be believed in an epistemic state e at level 0 only if there is a t such that $P(t)$ is supported by all the extended worlds in e. This means that it is not sufficient that the extended worlds in e support an existential; they must all agree on some term t. So an epistemic state where say $\boldsymbol{O}\exists x[P(x) \wedge Q(x)]$ is true should be the set of extended worlds that support the Skolemized version of this KB, that is, something like $[P(a) \wedge Q(a)]$, for some Skolem constant a. In general, the Skolemization of the KB is needed to guarantee the existence of the terms now required for believing existentials.

These are the main ideas behind our logic.

2.1 The Logic $\mathcal{L}$

Here we briefly go over the logic $\mathcal{L}$, a variant of first-order logic considered in (Levesque and Lakemeyer 2001; Lakemeyer and Levesque 2019).[1] The language of $\mathcal{L}$ is a first-order dialect with = and an infinite supply of function and predicate symbols of every arity. In addition, the language also features a set $\mathcal{N}$ of standard names $^{\#}1, ^{\#}2, ^{\#}3, \ldots,$ which are syntactically treated like constants but which are intended to be isomorphic to the (fixed) domain of discourse. In other words, standard names can be thought of as constants that satisfy the unique name assumption and an infinitary version of domain closure. Among other things, standard names allow for a very simple, substitutional account of quantifiers. See (Levesque and Lakemeyer 2001) for more discussion on why standard names are useful. In the following we often simply write "name" instead of "standard name." Terms and atomic formulas, also called atoms, are defined in the usual way, and so are formulas using the connectives $\neg$ and $\wedge$ and the quantifier $\forall$. Other connectives like $\vee, \supset, \equiv$ and the quantifier $\exists$ are freely used as syntactic abbreviations. Any formula from $\mathcal{L}$ is also called an *objective* formula. A *sentence* is a formula without free variables.

Ground terms are terms without variables. Function symbols with names as arguments are called *primitive terms,* and predicate symbols with names as arguments are called *primitive atoms.* So $f(^{\#}1, ^{\#}2)$ is a primitive term while $f(g(^{\#}1), ^{\#}2)$ is ground but not primitive.

The semantics is defined in terms of *worlds*, which are mappings from the primitive terms into $\mathcal{N}$ and from the primitive atoms into $\{0, 1\}$.

The meaning of an arbitrary ground term is given in terms of its *coreferring* standard name. Formally, given a ground term t and a world w we define $|t|_w$ (read: the coreferring standard name for t given w) by:

1. If $t \in \mathcal{N}$, then $|t|_w = t$;
2. $|h(t_1, \ldots, t_k)|_w = w[h(n_1, \ldots, n_k)]$, where $n_i = |t_i|_w$.

The truth of a sentence wrt world w (written as $w \models \phi$) is defined inductively as follows:

1. $w \models p(t_1, \ldots, t_k)$ iff $w[p(n_1, \ldots, n_k)] = 1$ where $|t_i|_w = n_i$;
2. $w \models (t_1 = t_2)$ iff $|t_1|_w$ and $|t_2|_w$ are the same names;
3. $w \models \neg\phi$ iff $w \not\models \phi$;
4. $w \models (\phi \wedge \psi)$ iff $w \models \phi$ and $w \models \psi$;
5. $w \models \forall x.\, \phi$ iff $w \models \phi^x_n$ for all names n;

Here ϕ^x_n stands for ϕ with every free occurrence of x replaced by n. A sentence ϕ is valid ($\models \phi$) iff for all worlds $w, w \models \phi$.

Apart from standard names and equality, $\mathcal{L}$ behaves exactly like classical first-order logic: it is shown in (Levesque and Lakemeyer 2001) that a sentence without standard names and equality is valid iff it is valid in classical logic. Standard names are, of course, special in that sentences like $(^{\#}6 \neq {^{\#}7})$ are valid, for example. Also, since the domain of discourse is infinite, sentences like $\exists x \forall y(x = y)$ are unsatisfiable in $\mathcal{L}$.

3 A Logic of Limited Reasoning

The language includes all of $\mathcal{L}$ and extends it in the following way: for any objective formula ϕ of $\mathcal{L}$, and any non-negative integer k, $\boldsymbol{B}_k\phi$ and $\boldsymbol{O}\phi$ are also formulas. Note, in particular, that nested beliefs are ruled out.

We follow the convention of using ϕ and ψ for objective formulas, and α and β for arbitrary formulas. We use p to refer to ground atomic formulas (that is, atomic formulas including equalities without variables), and ρ and τ to refer to literals, with $\bar{\rho}$ as the complement of ρ. We let a, b, c, d refer to clauses, understood as finite sets of literals. Finally we use θ to refer to ground substitutions, mappings from variables to ground terms. For any formula α, $\alpha\theta$ is the sentence that results from replacing all free variables x in α by $\theta(x)$.

[1] As LL make use of $\mathcal{L}$ the same way as we do in this paper, this subsection is largely taken from (Lakemeyer and Levesque 2019).

Skolemization and dual-Skolemization will be key in defining the semantics of quantified beliefs. As already mentioned, an epistemic state will be said to only believe a KB just in case it only believes a Skolemization of KB, but, of course, it should not matter which variables and Skolem symbols are used. In particular, we want to make sure that two epistemic states which differ only in the choice of Skolem symbols are considered equivalent in the sense that they believe the same sentences that do not mention Skolem symbols. To be able to express this formally, we need a bit of machinery in order to rename Skolem functions or variables in a quantifier-free formula in a systematic way. For that we reserve two special infinite sets of function symbols G and H, to be used for Skolemization and dual-Skolemization, respectively. (For simplicity, we leave the arity of these function symbols unspecified). We call a formula *clean* if it does not use these Skolem symbols. (The primary aim in this paper is to specify a logic of belief over clean sentences, using the ones with Skolem symbols only as auxiliary support.) Let X be the set of variables in the language. Then a *Skolem renaming* is a bijection λ from $(X \cup G \cup H)$ to itself, such that $\lambda(x) \in X$, $\lambda(g) \in G$, and $\lambda(h) \in H$ for any $x \in X$, $g \in G$ and $h \in H$. Then $\phi\lambda$ is the result of replacing each symbol $s \in (X \cup G \cup H)$ in ϕ by $\lambda(s)$. Finally, we define three equivalence relations over quantifier-free formulas: $\phi \sim \phi'$ iff there is a Skolem renaming λ such that $\phi\lambda = \phi'$; $\phi \sim_G \phi'$ iff $\phi \sim \phi'$ where the λ satisfies $\lambda(h) = h$; and similarly $\phi \sim_H \phi'$ iff $\phi \sim \phi'$ where the λ satisfies $\lambda(g) = g$.

3.1 Extended Worlds and Epistemic States

As noted above, the semantics of the logic relies on a notion of extended world:

Definition 1 (World) *An extended world w is a function from ground atoms (including equality atoms) to $\{0, 1, *\}$. (When the context is clear, we will just call them "worlds.") An extended world w is called* standard *if there is a two-valued world w' from $\mathcal{L}$ such that for every ground atom p, $w[p] = 1$ iff $w' \models p$, and $w[p] = 0$ iff $w' \models \neg p$.*

Note that an extended world maps all ground atoms to values, not just the primitive ones as in $\mathcal{L}$. So, for example, we can have $w[P(n)] = 1$ for every standard name n, and still have $w[P(a)] = 0$. (However this can never happen with a standard world.)

Since worlds can support both the truth and falsity of sentences, we use two separate support relations, $\models_T$ and $\models_F$ defined as follows:

Definition 2 (World support) *For any world w and quantifier-free sentence ϕ, the relations $w \models_T \phi$ and $w \models_F \phi$ are defined recursively as follows:*

1. $w \models_T p$ *iff* $w[p] \neq 0$;
   $w \models_F p$ *iff* $w[p] \neq 1$.
2. $w \models_T \neg\phi$ *iff* $w \models_F \phi$;
   $w \models_F \neg\phi$ *iff* $w \models_T \phi$.
3. $w \models_T (\phi \vee \psi)$ *iff* $w \models_T \phi$ *or* $w \models_T \psi$;
   $w \models_F (\phi \vee \psi)$ *iff* $w \models_F \phi$ *and* $w \models_F \psi$.

For a set of quantifier-free sentences Φ we will sometimes write $w \models_T \Phi$ to mean that $w \models_T \phi$ for all $\phi \in \Phi$.

It is useful to define a notion of "strong entailment" based on the idea of extended worlds:

Definition 3 *Let ϕ and ψ be sentences without quantifiers. Then $\phi \Rightarrow \psi$ if for all extended worlds w, if $w \models_T \phi$ then $w \models_T \psi$.*

Note that strong entailment is a subset of logical entailment (that is, if $\phi \Rightarrow \psi$ then the sentence $(\phi \supset \psi)$ is valid in $\mathcal{L}$), but it is a proper subset: $(p \wedge (\neg p \vee q)) \not\Rightarrow q$, for example. There is, in fact, a close connection between the two notions:

Proposition 1 *Let ϕ and ψ be sentences without quantifiers that use only atomic sentences $p_0, \ldots, p_k$. Then $(\phi \supset \psi)$ is valid in $\mathcal{L}$ iff $\phi \Rightarrow (\psi \vee \bigvee(p_i \wedge \neg p_i))$.*

This property is closely related to a similar connection between tautological entailment, a fragment of relevance logic, and strong entailment. The difference is that the semantics of tautological entailment, originally due to Dunn (1976) and later adopted in (Levesque 1984; Patel-Schneider 1985; Frisch 1987) and others, offers a fourth truth value "neither true nor false support." While this additional truth value blocks arbitrary tautologies from being entailed, we still have the connection that ϕ strongly entails ψ iff

$$(\phi \wedge \bigwedge(p_i \vee \neg p_i) \text{ tautologically entails } (\psi \vee \bigvee(p_i \wedge \neg p_i)).$$

We opted to not include the fourth truth value as tautologies can easily be detected provided a (propositional) formula is in conjunctive normal form.

Turning now to epistemic states, here is their definition and how they work with quantifier-free sentences:

Definition 4 (Epistemic state) *An extended epistemic state is any set of extended worlds. (Again, when the context is clear, we drop the word "extended.")*

In the quantifier-free case, the beliefs of an epistemic state e at level 0 are exactly those sentences which are supported by all the worlds in e. The following definitions show how this set of worlds can be shrunk in an iterative fashion by successively eliminating certain worlds. Belief at higher levels will then be defined in terms of truth in all the remaining worlds.

Definition 5 (Unsupported literals)
$U(w) = \{p \mid w[p] = 0\} \cup \{\neg p \mid w[p] = 1\}$.

The unsupported literals of w are those that w says cannot be true.

Definition 6 (Eliminated world) *An epistemic state e eliminates world w if there is a ground atom p such that for every world $w' \in e$, if $U(w) \subseteq U(w')$, then $w'[p] = *$.*

Intuitively, e eliminates w if there is some p such that the claims made by w (in terms of what literals cannot be true) depend on p having value $*$. In other words, if we only kept worlds in e where p had value 0 or 1, no worlds would support the claims made by w.

Definition 7 (Successor epistemic state)
$S(e) = e - \{w \mid e \text{ eliminates } w\}$.

Note that if w is standard, it is never eliminated since for no p do we have $w[p] = *$.

Definition 8 (Epistemic state support) *For any epistemic state e, number k, and quantifier-free sentence ϕ, $e \rightarrow_k \phi$ if for all $w \in S^k(e)$, $w \models_{\mathrm{T}} \phi$. (Here $S^0(e) = e$.)*

Example 1 *Consider the simple case where we have only two atomic sentences p and q. Let $w_1[p] = w_1[q] = 1$, $w_2[p] = *$, $w_2[q] = 0$, and let $e = \{w_1, w_2\}$. It is easy to see that $e \rightarrow_0 p \wedge (p \supset q)$ yet $e \not\rightarrow_0 q$ because of w_2. Note, however, that e eliminates w_2, that is, $S(e) = \{w_1\}$ and we obtain $e \rightarrow_1 q$.*

As we will see in the subsections to follow, $\boldsymbol{B}_k$ will be defined in terms of $\rightarrow_k$, and we will go over a slightly more complex example in more detail once we have the full definition of $\boldsymbol{B}_k$. While $\rightarrow_k$ is not closed under logical entailment, it is not hard to see that it is closed under strong entailment: if $\phi \Rightarrow \psi$ and $e \rightarrow_k \phi$, then $e \rightarrow_k \psi$. Furthermore, on the path towards eventual completeness, we have properties like this:

Lemma 1 *Suppose $e \rightarrow_k p$ and $e \rightarrow_k (\neg p \vee q)$. Then $e \rightarrow_{k+1} q$.*

Proof: Suppose $w \in S^k(e)$ and $w \not\models_{\mathrm{T}} q$. For any $w' \in S^k(e)$, $w' \models_{\mathrm{T}} p$, $w' \models_{\mathrm{T}} (\neg p \vee q)$, and if $U(w) \subseteq U(w')$, then $w' \not\models_{\mathrm{T}} q$ and so $w'[p] = *$. So if $w \in S^k(e)$ and $w \not\models_{\mathrm{T}} q$, then $S^k(e)$ eliminates w. It follows that for all $w \in S^{k+1}(e)$, $w \models_{\mathrm{T}} q$. Hence $e \rightarrow_{k+1} q$. ∎

Regarding Skolem renaming, we have the following:

Definition 9 *Let λ be a Skolem renaming. For any world w, $w\lambda$ is the world defined by $w\lambda[p] = w[p\lambda]$. For any epistemic state e, $e\lambda = \{w\lambda \mid w \in e\}$. Finally, for any two epistemic states e and e', we say that $e \sim e'$ iff there is a Skolem renaming λ such that $e\lambda = e'$. (The relations $e \sim_G e'$ and $e \sim_H e'$ are analogous.)*

Lemma 2 *For any quantifier-free sentence ψ and any Skolem renaming λ, if $e\lambda \rightarrow_k \psi$ then $e \rightarrow_k \psi\lambda$.*

Proof: The proof is by induction on k, but here we consider just the case $k = 0$. So assume that $e\lambda \rightarrow_0 \psi$ and that $w \in e$. Then $w\lambda \in e\lambda$ and so $w\lambda \models_{\mathrm{T}} \psi$. Then (by an easy induction on the length of ψ), $w \models_{\mathrm{T}} \psi\lambda$. Therefore, $e \rightarrow_0 \psi\lambda$. ∎

3.2 Skolemization

The definitions above deal with the quantifier-free part of the language; quantifiers within beliefs are handled by Skolemization and dual-Skolemization.

Assume that the sentence to be Skolemized only uses the $\exists$, $\vee$ and $\neg$ connectives. We must first rewrite the sentence with new variables as necessary so that each quantifier is over a distinct variable. We call a variable an *E-variable* if the $\exists$ that quantifies it appears within an even number of $\neg$ symbols, and an *A-variable* otherwise. We define a *Skolemization* of a sentence ϕ to be any formula that results from eliminating the quantifiers in ϕ, and replacing each E-variable y of ϕ by a distinct term $g(x_1, \ldots, x_k)$, where $g \in G$, and the x_i are all the A-variables that y appears within the scope of in ϕ. We define a *dual-Skolemization* of a sentence to be similar, except that each A-variable x is replaced by a distinct term $h(y_1, \ldots, y_k)$, where $h \in H$, and the y_i are all the E-variables that x appears within the scope of in ϕ. So Skolemization eliminates existential variables and leaves the universal ones free, while dual-Skolemization does the reverse. Note that, in contrast to the usual definition of (dual-)Skolemization, we also remove all remaining quantifiers. This is done just for convenience as we will need to replace the now free variables by other terms.

The main property of Skolemization is that the fo-satisfiability of a set of sentences reduces to the fo-satisfiability of their Skolemizations:

Proposition 2 *Let $\Phi = \{\phi_1, \phi_2, \ldots\}$ be a set of clean sentences. Let $\Psi = \{\forall\psi_1, \forall\psi_2, \ldots\}$ be a set where each ψ_i is a Skolemization of ϕ_i using its own distinct Skolem symbols. Then Φ is fo-unsatisfiable iff Ψ is fo-unsatisfiable.*

Note that the definition of Skolemization above does not specify exactly which Skolem symbols to use in eliminating existentially-quantified variables or which new variables to use in replacing variables that are quantified more than once. We do have that if ϕ_1 and ϕ_2 are Skolemizations of a clean ϕ, then $\phi_1 \sim_G \phi_2$, and if ψ_1 and ψ_2 are dual-Skolemizations of a clean ψ, then $\psi_1 \sim_H \psi_2$. It will be convenient, however, to name one of these Skolemizations as a representative:

Definition 10 *Assume that all the formulas of the language are ordered in some way. For any sentence ϕ, SKO(ϕ) denotes the Skolemization of ϕ that appears first in this ordering, and DSKO(ϕ) does the same for dual-Skolemization.*

Note that while *DSKO* eliminates universally quantified variables, $\neg$*DSKO*$(\neg\phi)$ is like *SKO* in eliminating existentially quantified variables. So a clean sentence ϕ is fo-unsatisfiable iff $\forall$*SKO*(ϕ) is fo-unsatisfiable iff $\forall\neg$*DSKO*$(\neg\phi)$ is fo-unsatisfiable.

3.3 Equality and Standard Names

Extended worlds have no special provisions for equality sentences or for the denotations of terms. These are handled in the logic using two special sets of formulas:

Definition 11
UNA $= \{(n \neq n') \mid n$ *and* n' *are distinct standard names*$\}$.

Definition 12 *The set EQ is UNA together with the following infinite set of formulas:*

1. $(x = x)$,
2. $\neg(x = y) \vee (y = x)$,
3. $\neg(x = y) \vee \neg(y = z) \vee (x = z)$,
4. $\neg(x_1 = y_1) \vee \cdots \vee \neg(x_k = y_k) \vee (f(x_1, \ldots, x_k) = f(y_1, \ldots, y_k))$, *for every k-ary function symbol f,*
5. $\neg(x_1 = y_1) \vee \cdots \vee \neg(x_k = y_k) \vee \neg P(x_1, \ldots, x_k) \vee P(y_1, \ldots, y_k)$, *for every k-ary predicate symbol P.*

The main property we need is this:

Proposition 3 *A sentence ϕ is valid in $\mathcal{L}$ iff $\{\neg\phi\} \cup$ EQ is fo-unsatisfiable.*

Corollary 2 *Let ϕ and ψ be clean sentences. Then the sentence $(\phi \supset \psi)$ is valid in $\mathcal{L}$ iff $\{SKO(\phi), \neg DSKO(\psi)\} \cup EQ$ is fo-unsatisfiable.*

In the following it will often be convenient to refer to the set of all ground instances of a set of quantifier-free formulas:

Definition 13 *For any set of quantifier-free formulas Φ, $GND(\Phi) = \{\phi\theta \mid \phi \in \Phi$ and θ a ground substitution$\}$.*

Herbrand's Theorem can now be restated as follows: if a set of quantifier-free formulas Φ is fo-unsatisfiable, then some finite subset of $GND(\Phi)$ is fo-unsatisfiable.

Definition 14 *For any set of quantifier-free formulas Φ, let $REP[\Phi]$ be defined by*

$$REP[\Phi] = \{w \mid w \models_{\text{T}} \psi \text{ for every } \psi \in GND(\Phi) \cup GND(EQ)\}.$$

In other words, $REP[\Phi]$ is the set of all worlds which support the truth of all ground instances of Φ and EQ.

3.4 Truth and Validity

We are now ready to define truth and validity in the logic of limited belief:

For any extended world w, extended epistemic state e and sentence α, the relation $e, w \models \alpha$ is defined recursively as follows:

1. $e, w \models p$ iff $w[p] \neq 0$, where p is a ground atom;
2. $e, w \models \neg\alpha$ iff $e, w \not\models \alpha$;
3. $e, w \models (\alpha \vee \beta)$ iff $e, w \models \alpha$ or $e, w \models \beta$;
4. $e, w \models \exists x\alpha$ iff for some n, $e, w \models \alpha^x_n$;
5. $e, w \models \boldsymbol{B}_k\phi$ iff there are substitutions $\theta_0, \ldots, \theta_k$ such that $e \rightarrow_k (DSKO(\phi)\theta_0 \vee \cdots \vee DSKO(\phi)\theta_k)$;
6. $e, w \models \boldsymbol{O}\phi$ iff $e \sim_G REP[SKO(\phi)]$.

When α is objective, we sometimes omit the e and write $w \models \alpha$; when α is subjective, we sometimes omit the w and write $e \models \alpha$. We say that e is *representable* iff $e \models \boldsymbol{O}\phi$ for some clean sentence ϕ. Finally, for any sentence α, we write $\models \alpha$ and say that α is *valid* iff $e, w \models \alpha$ for every representable e and every standard w.

Rules (1)-(4) are the usual ones (as in $\mathcal{L}$). Note that validity is defined wrt standard worlds only, so that for objective sentences, validity in this logic agrees with validity in $\mathcal{L}$.

Rules (5) and (6) are admittedly quite a hand full, as they need to deal with the intricacies of quantifiers in the context of limited belief. It turns out that, if we only consider the special case of quantifier-free sentences without equality, the rules can be drastically simplified with an intuitive reading. Therefore, it is instructive to look at this simple case first together with an example. So let ϕ be a quantifier-free objective sentence without equality. Then Rules (5) and (6) can be rewritten as:

5' $e, w \models \boldsymbol{B}_k\phi$ iff for all w', if $w' \in S^k(e)$ then $w' \models_{\text{T}} \phi$;

6' $e, w \models \boldsymbol{O}\phi$ iff for all w', $w' \in e$ iff $w' \models_{\text{T}} \phi$.

In other words, ϕ is believed at level k if all those worlds support ϕ that remain after removing eliminated worlds from e in k rounds, and ϕ is all that is believed if e is the largest set of worlds supporting ϕ.

To see how $\boldsymbol{B}_k$ works for different values of k, let us look at the following example.

Example 2 *Suppose there are only three atomic sentences, p, q and r, and so $3 \times 3 \times 3$ possible (extended) worlds. Now let $e = \{w \mid w \models_{\text{T}} \{p, (\neg p \vee q), (\neg q \vee r)\}\}$. In other words, $e \models \boldsymbol{O}(p \wedge (\neg p \vee q) \wedge (\neg q \vee r))$. Then e does not include the nine worlds where $w[p] = 0$, the three worlds where $w[p] = 1$ and $w[q] = 0$, and the single world w_{110}. (The notation w_{xyz} here means the world where $w[p] = x$, $w[q] = y$, and $w[r] = z$.) This means that e is the following set of fourteen worlds:*

$$\{w_{***}, w_{**0}, w_{**1}, w_{*0*}, w_{*00}, w_{*01}, w_{*1*}, w_{*10}, w_{*11}, w_{1**}, w_{1*0}, w_{1*1}, w_{11*}, w_{111}\}.$$

Some observations:

- $e \models \boldsymbol{B}_0 p$, $e \not\models \boldsymbol{B}_0 q$, $e \not\models \boldsymbol{B}_0 r$.
  We can see that q is not believed by virtue of w_{*0*}, w_{*00}, and w_{*01}, and that r is not believed by virtue of w_{**0}, w_{*00}, w_{*10}, and w_{1*0}.
- $e \models \boldsymbol{B}_1 p$, $e \models \boldsymbol{B}_1 q$, $e \not\models \boldsymbol{B}_1 r$.
  To calculate $S(e)$, we need the value of $U(w)$ for each $w \in e$. These are as follows:

  $$\begin{array}{ll} U(w_{***}) = \{\}, & U(w_{**0}) = \{r\}, \\ U(w_{*0*}) = \{q\}, & U(w_{*00}) = \{q, r\}, \\ U(w_{*1*}) = \{\neg q\}, & U(w_{*10}) = \{\neg q, r\}, \\ U(w_{1**}) = \{\neg p\}, & U(w_{1*0}) = \{\neg p, r\}, \\ U(w_{11*}) = \{\neg p \neg q\}, & U(w_{111}) = \{\neg p, \neg q, \neg r\}. \end{array}$$

  $$\begin{array}{l} U(w_{**1}) = \{\neg r\}, \\ U(w_{*01}) = \{q, \neg r\}, \\ U(w_{*11}) = \{\neg q, \neg r\}, \\ U(w_{1*1}) = \{\neg p, \neg r\}, \}. \end{array}$$

  It is easy to see that if $q \in U(w)$ then $w[p] = *$. This means that e eliminates each such world: w_{*0*}, w_{*00} and w_{*01}. (This is why q will be believed.) Also, e eliminates w_{*10} and w_{1*0}. Because of w_{111}, e does not eliminate w_{***}, w_{**1}, w_{*1*}, w_{*11}, w_{1**}, w_{1*1} w_{11*}, and w_{111}. Finally, e does not eliminate w_{**0} because w_{*10} and w_{1*0} assign different atoms to $*$. (This is why r is still not believed.) In the end, we have that $S(e) = \{w_{***}, w_{**0}, w_{**1}, w_{*1*}, w_{*11}, w_{1**}, w_{1*1} w_{11*}, w_{111}\}$. Hence, clearly $e \models B_1 q$ since for all $w \in S(e)$ we have that $w \models_{\text{T}} q$.
- $e \models \boldsymbol{B}_2 p$, $e \models \boldsymbol{B}_2 q$, $e \models \boldsymbol{B}_2 r$.
  For every $w \in S(e)$ other than w_{**0}, $S(e)$ does not eliminate w because of w_{111}. However, $S(e)$ will eliminate w_{**0} (because the w_{*10} and w_{1*0} of the previous step are now gone). So $S(S(e)) = \{w_{***}, w_{**1}, w_{*1*}, w_{*11}, w_{1**}, w_{1*1} w_{11*}, w_{111}\}$, and r is believed at level 2.

We will see below that these properties regarding $\boldsymbol{B}_0$, $\boldsymbol{B}_1$, and $\boldsymbol{B}_2$ follow from the general way belief relates to Resolution (see especially Lemma 5).

Let us now go back to Rules (5) and (6) in their full generality. In essence they reduce belief and only-believing to the quantifier-free case after Skolemization or dual-Skolemization, also taking into account equality. This is illustrated by the following theorem:

Theorem 2 *For clean sentences ϕ and ψ, $\models (\boldsymbol{O}\phi \equiv \boldsymbol{O}\forall\phi')$ and $\models (\boldsymbol{B}_k\psi \equiv \boldsymbol{B}_k\exists\psi')$, where $\phi' = SKO(\phi)$ and $\psi' = DSKO(\psi)$.*

Proof: Follows immediately from the fact that $SKO(\forall\phi') = \phi'$ and $DSKO(\exists\psi') = \psi'$. ∎

Note that even though in this logic there is not a unique epistemic state that satisfies $\boldsymbol{O}\phi$, the states that satisfy it are all very similar, in that they satisfy the same subjective sentences:

Lemma 3 *Suppose $e_1 \sim_G e_2$. For any clean sentence ϕ, $e_1 \models \boldsymbol{O}\phi$ iff $e_2 \models \boldsymbol{O}\phi$.*

Proof: Follows from the fact that $\sim_G$ is an equivalence relation. ∎

Lemma 4 *Suppose $e_1 \sim_G e_2$. For any clean sentence ψ, $e_1 \models \boldsymbol{B}_k\psi$ iff $e_2 \models \boldsymbol{B}_k\psi$.*

Proof: Suppose that $e_2\lambda = e_1$ for some bijection λ from G to G, and $e_1 \models \boldsymbol{B}_k\psi$, so that there are substitutions $\theta_0, \ldots, \theta_k$ such that $e_1 \rightarrow_k [DSKO(\psi)\theta_0 \vee \cdots \vee DSKO(\psi)\theta_k]$. By Lemma 2, $e_2 \rightarrow_k [\bigvee DSKO(\psi)\theta_i]\lambda$. Since ψ is clean and $DSKO$ does not introduce symbols from G, $[\bigvee DSKO(\psi)\theta_i]\lambda = \bigvee DSKO(\psi)(\theta_i\lambda)$. So there are substitutions $\theta'_0, \ldots, \theta'_k$ such that $e_2 \rightarrow_k \bigvee DSKO(\psi)\theta'_i$. Therefore, $e_2 \models \boldsymbol{B}_k\psi$. ∎

Corollary 3 *Suppose $e_1 \sim_G e_2$. For any clean subjective sentence σ, $e_1 \models \sigma$ iff $e_2 \models \sigma$.*

Proof: The proof is by induction on σ with Lemmas 3 and 4 as base cases. ∎

Theorem 3 *For any clean objective sentence ϕ and clean subjective σ, $\models (\boldsymbol{O}\phi \supset \sigma)$ iff $e \models \sigma$, where $e = REP[SKO(\phi)]$.*

Proof: The ($\Rightarrow$) direction is immediate, and the ($\Leftarrow$) direction follows directly from Corollary 3. ∎

Note that there is no special rule in the logic for equality. When it comes to truth involving equality, standard worlds deliver all the expected properties from $\mathcal{L}$. When it comes to belief involving equality, however, the axioms of equality *EQ* (including *UNA*) are conceptually added to the knowledge base (via *REP*) to be reasoned with, like anything else. So although there are extended worlds w where $w \not\models_{\mathrm{T}} (n = n)$ and $w \not\models_{\mathrm{T}} (n' \neq n)$ for distinct names n' and n, once the axioms of equality are taken into account, we get these sentences as beliefs (that is, $\boldsymbol{B}_0(n = n)$ and $\boldsymbol{B}_0(n' \neq n)$ are both valid).

Before looking at general properties of this logic, let us consider a simple example involving equality and standard names:

Example 3 $\models \boldsymbol{O}[\forall x. x \neq {}^{\#}5 \supset P(x)] \supset \boldsymbol{B}_1 P({}^{\#}7)$.

Let $e = REP[x \neq {}^{\#}5 \supset P(x)]$. By Theorem 3, it is sufficient to show that $e \models \boldsymbol{B}_1 P({}^{\#}7)$. We have that $e \rightarrow_0 ({}^{\#}7 \neq {}^{\#}5 \supset P({}^{\#}7))$ and $e \rightarrow_0 ({}^{\#}7 \neq {}^{\#}5)$ from UNA. So by Lemma 1, $e \rightarrow_1 P({}^{\#}7)$, and therefore $e \models \boldsymbol{B}_1 P({}^{\#}7)$.

4 Properties of the Logic

In this section, we confirm that the logic of belief defined above satisfies the desiderata listed at the beginning of the paper, and then consider other logical properties of the belief.

4.1 Satisfying the Desiderata

Cumulativity

Theorem 4 $\models (\boldsymbol{B}_k\psi \supset \boldsymbol{B}_{k+1}\psi)$.

Proof: This follows from the fact that $S(e) \subseteq e$, and so if $e \rightarrow_k \gamma$, then $e \rightarrow_{k+1} \gamma$. ∎

Expressiveness We use the following property of Skolemization:

Proposition 4 *For any objective sentence ϕ, there is a substitution θ^* such that $SKO(\phi)\theta^* = DSKO(\phi)\theta^*$.*

Theorem 5 *For any clean sentence ϕ, there is an epistemic state e such that $e \models \boldsymbol{O}\phi$, and moreover $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_0\phi)$.*

Proof: Let $e = REP[SKO(\phi)]$. Then clearly $e \models \boldsymbol{O}\phi$. To show that $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_0\phi)$, by Theorem 3, it suffices to show that $e \models \boldsymbol{B}_0\phi$. Now let θ^* be as in Proposition 4. Let w be any element of $e = REP[SKO(\phi)]$. It follows that $w \models_{\mathrm{T}} SKO(\phi)\theta^*$, and hence $w \models_{\mathrm{T}} DSKO(\phi)\theta^*$. Therefore, $e \rightarrow_0 DSKO(\phi)\theta^*$ and so $e \models \boldsymbol{B}_0\phi$. ∎

Soundness and Eventual Completeness We will be proving soundness and eventual completeness using propositional Resolution. First, we turn to *clauses*, which are finite sets of literals, interpreted disjunctively. In other words, $w \models_{\mathrm{T}} c$ iff for some $\rho \in c$, $w \models_{\mathrm{T}} \rho$; $w \models_{\mathrm{F}} c$ iff for every $\rho \in c$, $w \models_{\mathrm{F}} \rho$. So the empty clause, written [], is understood to satisfy $w \not\models_{\mathrm{T}} []$ and $w \models_{\mathrm{F}} []$ for every w. We define the conversion of a quantifier-free formula into clausal form as follows:

Definition 15 *Assume that ϕ has no quantifiers and has been rewritten so that it does not use $\vee$, $\supset$, or $\equiv$. Then $CNF(\phi)$ is a finite set of clauses defined inductively by:*

1. *$CNF(\phi) = \{\{\phi\}\}$, when ϕ is a literal;*
2. *$CNF(\phi \wedge \psi) = CNF(\phi) \cup CNF(\psi)$;*
3. *$CNF(\neg\neg\phi) = CNF(\phi)$;*
4. *$CNF(\neg(\phi \wedge \psi)) = \{a \cup b \mid a \in CNF(\neg\phi), b \in CNF(\neg\psi)\}$.*

Note that for any w and any quantifier-free sentence ϕ, $w \models_{\mathrm{T}} \phi$ iff for every $c \in CNF(\phi)$, $w \models_{\mathrm{T}} c$. Finally, we define the direct resolvents of a set of ground clauses:

Definition 16 *For any set of ground clauses C, let $RP(C)$ be the set of clauses defined by*

$$C \cup \{(a \cup b) \mid \text{for some } p,\ (\{p\} \cup a) \in C,\ (\{\neg p\} \cup b) \in C\}.$$

Notice that *RP* applies one step of propositional Resolution to C, and in general, RP^k applies k steps. We will be using the following property of Resolution:

Proposition 5 *Let C be a a set of ground clauses and d a non-tautologous ground clause. Then $C \cup \{\neg d\}$ is unsatisfiable iff for some k and d', $d' \subseteq d$ and $d' \in RP^k(C)$.*

As a special case of this proposition, we have the usual "refutation completeness" of Resolution: C is unsatisfiable iff for some k, $[] \in RP^k(C)$.

The following lemma shows a tight connection between the notion of eliminating worlds from $\{e \mid w \models_{\mathrm{T}} C\}$, where C is a set of ground clauses, and Resolution:

Lemma 5 *Let C be any set of ground clauses and let $e = \{w \mid w \models_{\mathrm{T}} C\}$. Then $S^k(e) = \{w \mid w \models_{\mathrm{T}} RP^k(C)\}$.*

Proof: The lemma holds by induction on k. Here we only show the base case for $k = 1$: if $e = \{w \mid w \models_{\mathrm{T}} C\}$ then $S(e) = \{w \mid w \models_{\mathrm{T}} RP(C)\}$.

($\Rightarrow$) We show that if $w \not\models_{\mathrm{T}} RP(C)$ then e eliminates w, and so $w \notin S(e)$. Since $w \not\models_{\mathrm{T}} RP(C)$, there is $(\{p\} \cup b) \in C$, $(\{\neg p\} \cup d) \in C$, such that $w \not\models_{\mathrm{T}} (b \cup d)$. So $(b \cup d) \subseteq U(w)$. Therefore, for any $w' \in e$ such that $U(w) \subseteq U(w')$, it follows that $w' \not\models_{\mathrm{T}} (b \cup c)$ and therefore $w'[p] = *$. Hence e eliminates w.

($\Leftarrow$) We show that if $w \models_{\mathrm{T}} RP(C)$ then e does not eliminate w, and so $w \in S(e)$. To do so, we show that for every p, there is a $w' \in e$ such that $U(w) \subseteq U(w')$ and where $w'[p] \neq *$. First, suppose that $w[p] \neq *$; then let $w' = w$ and the claim is satisfied. Otherwise, if $w[p] = *$, define w' to be like w except on p, where $w'[p] = 1$ if for some $(\{p\} \cup b) \in C$, $w \not\models_{\mathrm{T}} b$, and 0 otherwise. So $U(w) \subseteq U(w')$ and $w'[p] \neq *$. To show that $w' \in e$, we show that for any $d \in C$, $w' \models_{\mathrm{T}} d$. There are three cases.

1. If d does not include p or $\neg p$, then $w' \models_{\mathrm{T}} d$ since $w \models_{\mathrm{T}} d$.
2. If $d = (\{p\} \cup d')$ then there are two subcases: if $w \not\models_{\mathrm{T}} d'$, then $w'[p] = 1$ and so $w' \models_{\mathrm{T}} d$; if $w \models_{\mathrm{T}} d'$, then $w' \models_{\mathrm{T}} d$ and so $w' \models_{\mathrm{T}} d$.
3. If $d = (\{\neg p\} \cup d')$ then there are two subcases: if $w'[p] = 0$, clearly, $w' \models_{\mathrm{T}} d$; if $w'[p] = 1$, then there is an $(\{p\} \cup b) \in C$ where $w \not\models_{\mathrm{T}} b$. Since $w \models_{\mathrm{T}} RP(C)$, $w \models_{\mathrm{T}} (b \cup d')$ and so $w \models_{\mathrm{T}} d'$. It follows that $w' \models_{\mathrm{T}} d$ and so $w' \models_{\mathrm{T}} d$. ∎

In a way, the lemma can be seen as providing a semantic justification of a resolution step in terms of eliminating certain non-standard worlds.

Corollary 4 *For any set C of ground clauses which contains $GND(EQ)$, $S^k(REP[C]) = REP[RP^k(C)]$.*

Proof: Since $GND(EQ) \subseteq C$, $REP[C] = \{w \mid w \models_{\mathrm{T}} C\}$ and $REP[RP^k(C)] = \{w \mid w \models_{\mathrm{T}} RP^k(C)\}$. The corollary then follows immediately from the previous lemma. ∎

Lemma 6 *Let C be a set of ground clauses containing $GND(EQ)$ and ψ a sentence without quantifiers. Then $C \cup \{\neg\psi\}$ is unsatisfiable iff there is a k such that $REP[C] \rightarrow_k \psi$.*

Proof: ($\Rightarrow$) Suppose $C \cup \{\neg\psi\}$ is unsatisfiable and let d be any non-tautologous clause in $CNF(\psi)$. Then $C \cup \{\neg d\}$ is unsatisfiable and by Proposition 5, there is an i and a $c \in RP^i(C)$ such that $c \subseteq d$. So for any $w \in REP[RP^i(C)]$, $w \models_{\mathrm{T}} d$. By Corollary 4, $REP[RP^i(C)] = S^i(REP[C])$. So $REP[C] \rightarrow_i d$. Now let k be the maximum of these i values over all clauses of $CNF(\psi)$. Then $REP[C] \rightarrow_k \psi$.

($\Leftarrow$) Suppose $C \cup \{\neg\psi\}$ is satisfiable. Then there is a standard world w in $REP[C]$ such that $w \models_{\mathrm{T}} \neg\psi$ and so $w \not\models_{\mathrm{T}} \psi$. Since w is standard, for every k, $w \in S^k(REP[C])$. ∎

Theorem 6 *Let ϕ and ψ be clean objective sentences. Then $\models (\phi \supset \psi)$ iff for some k, $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$.*

Proof: $\models (\phi \supset \psi)$ iff (by Corollary 2) $C \cup \{\neg DSKO(\psi)\}$ is fo-unsatisfiable, where $C = CNF(SKO(\phi)) \cup EQ$ iff (by Theorem 1) some finite subset of $GND(C \cup \{\neg DSKO(\psi)\})$ is unsatisfiable iff there exist $\theta_0, \ldots, \theta_s$ such that $GND(C) \cup \{\neg\gamma\}$ is unsatisfiable, where $\gamma = [DSKO(\psi)\theta_0 \vee \cdots \vee DSKO(\psi)\theta_s]$ iff (by Lemma 6) there is an r such that $REP[GND(C)] \rightarrow_r \gamma$ iff (letting k be the maximum of r and s and noting that $REP[GND(C)] = REP[SKO(\phi)]$) there is a k such that $REP[SKO(\phi)] \models \boldsymbol{B}_k\psi$ iff (by Theorem 3) there is a k such that $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$. ∎

Tractability We will address tractability in two steps, first by considering the simpler case of quantifier-free formulas without equality and then the full first-order case with equality.

Looking over the details of the proof of soundness and eventual completeness above, it is not hard to see that, if ϕ and ψ are quantifier-free sentences without equality, then $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$ iff for all non-tautologous clauses $d \in CNF(\psi)$, there is a $d' \in RP^k(CNF(\phi))$ such that $d' \subseteq d$. In other words, in order to figure out whether $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$ holds it suffices to check, for each clause in the CNF of ψ, whether it contains complementary literals or is subsumed by some clause obtained by at most k resolution steps from the clauses in the CNF of ϕ.

Using this, we can show that it is possible to efficiently decide if $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$ under these assumptions: the k is small, the query ψ is small, and if the KB ϕ itself is large, it is only because it is a large conjunction of sentences that are themselves small.

Theorem 7 *There is a procedure that takes as input a number k, and quantifier-free sentences $\phi_1, \ldots, \phi_N$ and ψ, and decides whether or not $\models (\boldsymbol{O}[\phi_1 \wedge \cdots \wedge \phi_N] \supset \boldsymbol{B}_k\psi)$. Moreover, if the inputs are restricted such that for some number c, $k \leq c$, $|\phi_i| \leq c$, and $|\psi| \leq c$, then the procedure will run in time that is polynomial in N.*

Proof: Here is the procedure. First, calculate $D = CNF(\psi)$ and $C_i = CNF(\phi_i)$. (These will be polynomial in N since ψ and each ϕ_i are bounded.) Then calculate $C = RP^k(C_1 \cup \cdots \cup C_N)$, which will also be polynomial in N since k is bounded. Finally, check that for each non-taut. $d \in D$, there is a $d' \in C$ such that $d' \subseteq d$. ∎

Note that even in this quantifier-free case, the tractability can fail for sentences that are not in the required form above. As a simple example, note that for any quantifier-free ϕ, we have that $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_0\psi)$ whenever $\psi = \bigwedge CNF(\phi)$, but there is no guarantee that this conversion to CNF can be determined efficiently.

One may wonder whether existing technology like SAT-solvers would be of any help in computing the beliefs at level k in the quantifier-free case. We believe that this is unlikely mainly because the question whether or not a set of clauses is satisfiable seems to tell us very little if anything about the beliefs at a fixed level k. For example, consider an unsatisfiable sentence ϕ and let q be an atomic sentence not occurring in ϕ. Suppose $\models \boldsymbol{O}\phi \supset \boldsymbol{B}_{17}q$ yet $\not\models \boldsymbol{O}\phi \supset \boldsymbol{B}_k q$ for all $k < 17$. In other words, the inconsistency is detected only at level 17 and beyond. A SAT-solver may well be able to figure out that ϕ is unsatisfiable, but it will not be able to give us finer-grained distinctions.

Let us now turn to the full first-order case with equality. Again, from the discussion of soundness and completeness we have that

$\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$ iff there are substitutions $\theta_0, \ldots, \theta_k$ such that for all non-tautologous $d \in CNF(DSKO(\psi)\theta_0 \vee \cdots \vee DSKO(\psi)\theta_k)$ there is a $c \in RP^k(GND(CNF(SKO(\phi)) \cup EQ))$ such that $c \subseteq d$.

We want to show that under certain reasonable assumptions, it will be possible to efficiently decide if $\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$. The problem with this is is that we cannot simply calculate $RP^k(GND(CNF(SKO(\phi)) \cup EQ))$ since this is an infinite set of clauses. To get around this, we restrict *EQ* to be finite and we replace *RP* by *RQ*, the first-order version of Resolution that handles clauses with variables (via most general unifiers and so on). We will end up using something like this:

$\models (\boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi)$ iff there are substitutions[2] $\theta_0, \ldots, \theta_k$ such that for all non-tautologous $d \in CNF(DSKO(\psi)\theta_0 \vee \cdots \vee DSKO(\psi)\theta_k)$there is a $c \in RQ^k(CNF(SKO(\phi)) \cup EQ')$ and a θ such that $c\theta \subseteq d$.

where EQ' is *EQ* restricted to the function and predicate symbols appearing in $SKO(\phi)$ or $DSKO(\psi)$, with *UNA* restricted to a finite set of standard names. Here we will be able to calculate $RQ^k(CNF(SKO(\phi)) \cup EQ')$, and the rest will involve guessing the appropriate substitutions. The precise definitions are as follows:

Definition 17 *For any two literals ρ and τ, $MGU[\rho, \tau]$ is the set of most general unifiers of ρ and τ (empty if the two literals do not unify).*

Definition 18 *For any set C of clauses, $F(C)$ is the union of $F(c)$ for all $c \in C$, where*

$$F(c) = \{c\} \cup F(\{c\theta \mid \{\rho, \tau\} \subseteq c, \rho \neq \tau, \theta \in MGU[\rho, \tau]\}).$$

Definition 19 *For any set C of clauses,*

$$RQ(C) = C \cup \{(a \cup b)\theta \mid \{\rho\} \cup a \in F(C), \{\tau\} \cup b \in F(C), \theta \in MGU[\bar{\rho}, \tau]\}.$$

[2]In contrast to the rest of the paper, here we need substitutions which may be non-ground.

Note that Definition 18 realizes what is known as *Factoring*, that is, the unification of literals within the same clause, which is needed for Resolution to be complete. In the definitions of F and *RQ*, we assume the clauses in C use distinct variables, and that just one θ is chosen (if one exists) so that the new clauses also have distinct variables. The main property of this first-order Resolution is the following generalization of Proposition 5:

Proposition 6 *Let C be a a set of clauses and d a non-tautologous ground clause. Then $C \cup \{\neg d\}$ is unsatisfiable iff for some k, d', and θ, $d'\theta \subseteq d$ and $d' \in RQ^k(C)$.*

We now prove that calculating what is believed at level k can be efficiently computed under the assumptions that the query is not too large, the k is not too large, and while the KB may be large, it is because it is a large conjunction of sentences that are not too large:

Theorem 8 *There is a procedure for deciding if $\models (\boldsymbol{O}[\phi_1 \wedge \cdots \wedge \phi_N] \supset \boldsymbol{B}_k\psi)$ that runs in time that is polynomial in N under the assumption that for some constant c, $k \leq c$, $|\phi_i| \leq c$, and $|\psi| \leq c$.*

Proof: Here is a sketch of the procedure. First, calculate $C = CNF(SKO(\bigwedge \phi_i))$. (This will be polynomial in N, since each ϕ_i is bounded.) Let $r = |\bigwedge \phi_i|$.

Then calculate EQ' from the given ϕ_i and ψ. Here the main complication is to limit the number of elements from *UNA* to a finite subset UNA'. It can be shown that UNA' can be restricted to those elements from *UNA* which mention the names in ϕ and ψ plus $max\{2^k, (k+1) * |\psi|\}$ new names. (The size of UNA' is polynomial since k and ψ are bounded.)

Having a finite EQ' in hand, we calculate $RQ^k(C \cup EQ')$, which will also be polynomial. Next, guess at the $(k+1)$ substitutions θ_j and calculate $D = CNF(\bigvee DSKO(\psi)\theta_j)$. (Again, the k and ψ are bounded. The "guessing" of appropriate substitutions can be made determinate by trying all potential MGUs between terms in $DSKO(\psi)$ and terms in $RQ^k(C \cup EQ')$, of which there are only polynomially many.) Finally, check that for each non-tautologous $d \in D$, there is a $c \in C$ such that $c\theta \subseteq d$ for some θ. (This is a special case of theta-subsumption.) ▮

4.2 Logical Transformations within Beliefs

In this subsection, we consider belief equivalence with respect to logical operators within a belief. The main theorem is this:

Theorem 9 (Equivalent beliefs) *For any k and any clean sentences ϕ, ψ, and χ, the following sentences are valid:*

$\boldsymbol{B}_k\phi \equiv \boldsymbol{B}_k(\phi \wedge \phi)$,
$\boldsymbol{B}_k\phi \equiv \boldsymbol{B}_k(\phi \vee \phi)$,
$\boldsymbol{B}_k\phi \equiv \boldsymbol{B}_k\neg\neg\phi$,
$\boldsymbol{B}_k(\phi \wedge \psi) \equiv \boldsymbol{B}_k(\psi \wedge \phi)$,
$\boldsymbol{B}_k(\phi \vee \psi) \equiv \boldsymbol{B}_k(\psi \vee \phi)$,
$\boldsymbol{B}_k(\phi \wedge (\psi \wedge \chi)) \equiv \boldsymbol{B}_k((\phi \wedge \psi) \wedge \chi))$,
$\boldsymbol{B}_k(\phi \vee (\psi \vee \chi)) \equiv \boldsymbol{B}_k((\phi \vee \psi) \vee \chi))$,
$\boldsymbol{B}_k(\phi \wedge (\psi \vee \chi)) \equiv \boldsymbol{B}_k((\phi \wedge \psi) \vee (\phi \wedge \chi))$,
$\boldsymbol{B}_k(\phi \vee (\psi \wedge \chi)) \equiv \boldsymbol{B}_k((\phi \vee \psi) \wedge (\phi \vee \chi))$,

$\boldsymbol{B}_k(\neg(\phi \wedge \psi)) \equiv \boldsymbol{B}_k(\neg\phi \vee \neg\psi)$,
$\boldsymbol{B}_k(\neg(\phi \vee \psi)) \equiv \boldsymbol{B}_k(\neg\phi \wedge \neg\psi)$.

These are all shown to be valid in roughly the same way.

First note that the items in the theorem are easy to prove when the formulas do not contain quantifiers. This is because propositional beliefs at any level are closed under strong entailment (see Definition 3). For example, to prove that $\boldsymbol{B}_k(\phi \vee \psi)$ logically implies $\boldsymbol{B}_k(\psi \vee \phi)$, we would simply observe that $(\phi \vee \psi) \Rightarrow (\psi \vee \phi)$. For space reasons we are not able to include a proof of the general case except to note that it follows from this lemma:

Lemma 7 *Let ϕ and ψ be clean formulas for which there are dual-Skolemizations ϕ' and ψ' such that $\phi'\theta \Rightarrow \psi'\theta$ for every θ. Then $\models (\boldsymbol{B}_k\phi \supset \boldsymbol{B}_k\psi)$.*

Although the above theorem sanctions many logical transformations within beliefs (for example, converting a belief into *CNF*), it does not sanction Modus Ponens:

Theorem 10 (Non-equivalent beliefs) *There are clean objective sentences without quantifiers ϕ and ψ such that $\models (\phi \equiv \psi)$, but $\not\models (\boldsymbol{B}_k\phi \equiv \boldsymbol{B}_k\psi)$.*

Proof: Let p and q be distinct atomic sentences. Let $\phi = (p \wedge (\neg p \vee q))$ and $\psi = (\phi \wedge q)$. Then $\models (\phi \equiv \psi)$, but $\not\models (\boldsymbol{B}_0\phi \equiv \boldsymbol{B}_0\psi)$. ∎

4.3 Logical Combinations of Beliefs

The final thing we consider in this section is the relationship between beliefs and logical combinations of other beliefs. For the most part, the arguments will be similar to those seen above. We need something new, however, to deal with standard names. Recall that in the logic, quantifiers outside the context of a belief are interpreted in the usual way using standard names. Some notation: if a is a constant and n is a standard name, we let ϕ^a_n denote the formula that results from replacing all occurrences of a in ϕ by n. We get the following:

Lemma 8 *Let e be representable, and let ψ be any sentence without quantifiers that uses $a \in H$ as a constant. If $e \rightarrow_k \psi$ then $e \rightarrow_k \psi^a_n$ for any name n.*

Lemma 9 *If γ is a quantifier-free sentence that uses $a \in H$ as a constant, then $\models (\boldsymbol{B}_k\exists\gamma \supset \boldsymbol{B}_k\exists\gamma^a_n)$ for any standard name n.*

Lemma 10 *Let ϕ be a formula without quantifiers. Then $\models (\boldsymbol{B}_k\exists\phi^x_n \supset \boldsymbol{B}_k\exists\phi)$.*

Here is the main theorem of this subsection involving the disjunction, conjunction, and quantification of belief:

Theorem 11 (Combinations of beliefs) *For any clean sentences ϕ and ψ and any clean formula γ with a single free variable x:*

1. $\models (\boldsymbol{B}_k\phi \vee \boldsymbol{B}_k\psi) \supset \boldsymbol{B}_k(\phi \vee \psi)$.
2. $\models \boldsymbol{B}_k(\phi \wedge \psi) \supset (\boldsymbol{B}_k\phi \wedge \boldsymbol{B}_k\psi)$, *but the converse fails.*
3. $\models \exists x\boldsymbol{B}_k\gamma \supset \boldsymbol{B}_k\exists x\gamma$.
4. $\models \boldsymbol{B}_k\forall x\gamma \supset \forall x\boldsymbol{B}_k\gamma$, *but the converse fails.*

Proof:

1. We show that $\models (\boldsymbol{B}_k\phi \supset \boldsymbol{B}_k(\phi \vee \psi))$; the other disjunct is analogous. Suppose *DSKO*$((\phi \vee \psi))$ is $(\phi' \vee \psi')$ where ϕ' is a dual Skolemization of ϕ. We have that $\phi'\theta \Rightarrow (\phi' \vee \psi')\theta$ for every θ, and so the result follows from Lemma 7.
2. We show that $\models (\boldsymbol{B}_k(\phi \wedge \psi) \supset \boldsymbol{B}_k\phi)$; the other conjunct is analogous. Suppose *DSKO*$((\phi \wedge \psi))$ is $(\phi' \wedge \psi')$ where ϕ' is a dual Skolemization of ϕ. We have that $(\phi' \wedge \psi')\theta \Rightarrow \phi'\theta$ for every θ, and so the result follows from Lemma 7. Regarding the converse, let ϕ be $\exists x.P(x)$, ψ be $\exists y.Q(y)$, and let
$$e = REP[(P(\#1) \vee P(\#2)) \wedge (Q(\#1) \vee Q(\#2))].$$
Then $e \models \boldsymbol{B}_1\phi$ and $e \models \boldsymbol{B}_1\psi$, but $e \not\models \boldsymbol{B}_1\exists x\exists y.P(x) \wedge Q(y)$ and so $e \not\models \boldsymbol{B}_1(\phi \wedge \psi)$.
3. If $e \models \exists x\boldsymbol{B}_k\gamma$, then for some standard name n, $e \models \boldsymbol{B}_k\gamma^x_n$. Suppose *DSKO*$(\exists x\gamma)$ is γ_1 and *DSKO*(γ^x_n) is ${\gamma_2}^x_n$, so that $\gamma_1 \sim_H \gamma_2$. Then by Theorem 2, $e \models \boldsymbol{B}_k\exists{\gamma_2}^x_n$. By Lemma 7, $e \models \boldsymbol{B}_k\exists{\gamma_1}^x_n$, and so by Lemma 10, $e \models \boldsymbol{B}_k\exists\gamma_1$. So by Theorem 2, $e \models \boldsymbol{B}_k\exists x\gamma$.
4. Let n be any standard name. Suppose *DSKO*$(\forall x\gamma)$ is ${\gamma_1}^x_a$ where $a \in H$, and let ${\gamma_2}^x_n$ be a dual-Skolemization of γ^x_n that does not use a anywhere, so that $\gamma_1 \sim_H \gamma_2$. Now suppose $e \models \boldsymbol{B}_k\forall x\gamma$. Then by Theorem 2, $e \models \boldsymbol{B}_k\exists{\gamma_1}^x_a$. By Lemma 7, $e \models \boldsymbol{B}_k\exists{\gamma_2}^x_a$ and by Lemma 9, $e \models \boldsymbol{B}_k\exists{\gamma_2}^x_n$ since a does not appear in γ_2. So by Theorem 2, $e \models \boldsymbol{B}_k\gamma^x_n$. Since this applies to any standard name n, it follows that if $e \models \boldsymbol{B}_k\forall x\gamma$ then $e \models \forall x\boldsymbol{B}_k\gamma$.
Regarding the converse, let $\gamma = (x \neq \#1 \vee x \neq \#2)$ and $e = REP[EQ]$. Then for every standard name n, $e \models \boldsymbol{B}_0\gamma^x_n$, and so $e \models \forall x\boldsymbol{B}_0\gamma$. But for any $a \in H$, $e \not\models \boldsymbol{B}_0\gamma^x_a$, and so $e \not\models \boldsymbol{B}_0\forall x\gamma$. ∎

Note that unlike belief in classical possible-world semantics we do not get the usual closure of belief under conjunction or under universal quantification. (These are the converses in the theorem above that are shown not to hold.) We do however get "eventual" versions of these closures.

Theorem 12 *For ϕ, ψ, γ as in Theorem 11, any k, and any representable e,*

1. *If $e \models (\boldsymbol{B}_k\phi \wedge \boldsymbol{B}_k\psi)$, then for some k', $e \models \boldsymbol{B}_{k'}(\phi \wedge \psi)$.*
2. *If $e \models \forall x\boldsymbol{B}_k\gamma$, then for some k', $e \models \boldsymbol{B}_{k'}\forall x\gamma$.*

The theorem follows easily from soundness and eventual completeness (Theorem 6) and this lemma:

Lemma 11 *Let e be representable with $e \models \boldsymbol{O}\phi$ for some clean ϕ. Then for any clean ψ and any k, $e \models \boldsymbol{B}_k\psi$ iff $\models \boldsymbol{O}\phi \supset \boldsymbol{B}_k\psi$.*

We remark that, besides satisfying the same principles of soundness, eventual completeness, cumulativity, and tractability as in LL, Theorems 9, 11 and 12 hold in both logics, but the proofs are quite different, of course. So do the two logic have the same valid sentences? The answer is: almost. There are subtle differences in the way elements of *UNA* are handled. For example, if KB $= (\#1 = \#2 \vee p)$, then p is believed at level 0 in the case of LL, while we obtain p only at level 1. This is because we require one resolution step using ($^\#1 \neq {}^\#2$) from UNA. Essentially, LL get p for free because of unsatisfiable atoms like ($^\#1 = {}^\#2$) are filtered out automatically when testing for subsumption. While this means that sometimes one may need to go to a higher belief level than in LL to obtain a belief, we regard the overall impact as minor. In fact, we conjecture that the two logics are identical when restricted to sentences without equality. Alternatively, we believe that it would be easy to modify the way subsumption is handled in LL so that the two logics become identical for the whole language.

5 Related Work

As already mentioned, modeling belief in terms of possible worlds (Kripke 1959) goes back to the seminal work of Hintikka (Hintikka 1962) (see also (Halpern and Moses 1992) for a survey and in-depth study). Since the use of classical (two-valued) worlds invariably leads to logical omniscience, making reasoning intractable in the propositional and undecidable in the first-order case, Hintikka already suggested the use of impossible worlds to address this issue (Hintikka 1975). In KR this idea was developed further by considering four-valued worlds as a basis of weaker models of belief or inference, see (Levesque 1984; Cadoli and Schaerf 1996; Frisch 1987; Fagin *et al.* 1990; Delgrande 1995; Cadoli and Schaerf 1996) for propositional and (Patel-Schneider 1985; Lakemeyer 1996; Lakemeyer and Levesque 2016) for first-order approaches. The underlying inference mechanisms are also closely related to *tautological entailment* (Dunn 1976), a fragment of relevance logic (Anderson and Belnap 1975).

The idea to consider levels of belief within a modal epistemic logic, where epistemic states are represented as sets of clauses, originated in (Liu et al. 2004), was further refined in (Liu and Levesque 2005; Lakemeyer and Levesque 2013; 2014; 2016; Schwering 2017), and culminated in LL's most recent proposal (Lakemeyer and Levesque 2019). Except for LL, belief levels in this line of work are defined in terms of splitting, either on a clause in the epistemic state (Liu et al. 2004; Liu and Levesque 2005; Lakemeyer and Levesque 2013), an arbitrary ground literal (Lakemeyer and Levesque 2014), or on the possible denotations of a term (Lakemeyer and Levesque 2016; Schwering 2017). For example, splitting on a literal l at belief level k means adding l to the epistemic state and then checking whether the belief in question obtains at level $k - 1$, and doing the same for the complement of l. All of these belief models, again with the exception of LL, have the property that each belief level is closed under *unit propagation*, that is, resolution with a clause containing a single literal. In all cases reasoning is tractable at every belief level k. A notable exception to the use of clauses as semantic primitive is (Klassen et al. 2015), where epistemic states are defined in terms of a three-valued variant of neighborhood semantics (Montague 1968; Scott 1970). Belief levels are again defined in terms of splitting on literals, and tractability obtains at every level. However, the work is limited to the propositional case. Among the above proposals, LL is the only one which is expressive, eventually complete, and tractable for the full language. We remark that (Lakemeyer and Levesque 2014; Klassen et al. 2015; Lakemeyer and Levesque 2016) are eventually complete for the propositional fragment.

Beginning with (Dalal 1996), there has also been work on tractable entailment relations of increasing complexity, again limited to a propositional language. Perhaps the most advanced such proposal is (D'Agostino 2015), which is based on a three-valued nondeterministic semantics first considered in (Crawford and Etherington 1998). The author defines a k-consequence relation, which features splitting on arbitrary formulas and closure under unit propagation. The k-consequence relation is eventually complete and a proof-theoretic account is also provided.

6 Conclusions

In this paper we proposed a model of limited belief, which shares desirable properties such as expressiveness, tractability and eventual completeness with the recently proposed logic by Lakemeyer and Levesque. In contrast to LL's semantics, which is based on clauses, ours uses three-valued possible worlds. Besides being more appealing semantically, defining belief in terms of a set of possible worlds lends itself naturally to extensions involving nested beliefs. While the details remain to be worked out, it does not seem to be hard to obtain properties like full introspection ($\boldsymbol{B}_k\alpha \supset \boldsymbol{B}_k\boldsymbol{B}_k\alpha$ and $\neg\boldsymbol{B}_k\alpha \supset \boldsymbol{B}_k\neg\boldsymbol{B}_k\alpha$), at least for the propositional fragment of the language. Another avenue that deserves further investigation would be to include more "easy" inferences at all levels. For example, as mentioned above, (Liu et al. 2004) and its descendants close beliefs under *unit propagation*. If we want to add such easy inferences without having to go to higher belief levels, at least two problems need to be addressed: (1) given the expressiveness of our language, where function symbols can be nested arbitrarily, unit propagation needs to be restricted to avoid undecidability; (2) it remains to be seen how to integrate a form of unit propagation into our possible-world framework. For the latter ideas from (Crawford and Etherington 1998; D'Agostino 2015), where unit propagation is given a semantic justification, may be useful.

Acknowledgments

We thank the reviewers for their thoughtful comments, which helped improve the presentation of the paper. Gerhard Lakemeyer also acknowledges his affiliation with the University of Toronto while this work was conducted.

References

A.R. Anderson and N.D. Belnap. 1975. *Entailment: The Logic of Relevance and Necessity*. Princeton University Press.

M. Cadoli and M. Schaerf. 1996. On the complexity of entailment in propositional multivalued logics. *Annals of Mathematics and Artificial Intelligence*, 18(1):29–50.

J.M. Crawford and D. Etherington. 1998. A non-deterministic semantics for tractable inference. In *Proc. of AAAI-98*, 286–291.

M. Dalal. 1996. Semantics of an anytime family of reasoners. In *Proc. of the 12th European Conference on Artificial Intelligence (ECAI-96)*, 360–364.

M. D'Agostino. 2015. An informational view of classical logic. Theor. Comput. Sci. 606, 79–97.

J.P. Delgrande. 1995. A framework for logics of explicit belief. *Computational Intelligence*, 11(1):47–88.

J.M. Dunn. 1976. Intuitive semantics for first-degree entailments and coupled trees. *Philosophical Studies*, 29:149–168.

R. Fagin, J.Y. Halpern, and M.Y. Vardi. 1990. A nonstandard approach to the logical omniscience problem. In *Proc. of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge (TARK-90)*, 41–55.

A. Frisch. 1987. Inference without chaining. In *Proc. of IJCAI-87*, 515–519.

J.Y. Halpern and Y. Moses. 1992. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54(3):319–379.

J. Hintikka. 1962. *Knowledge and Belief.* Cornell University Press.

J. Hintikka. 1975. Impossible Possible Worlds Vindicated. *Journal of Philosophical Logic,* 4, 475–484.

T.Q. Klassen, S.A. McIlraith, and H.J. Levesque. 2015. Towards Tractable Inference for Resource-Bounded Agents. AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning, 89–95, AAAI Press.

S. Kripke. 1959. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24:1–14.

G. Lakemeyer. 1996. *Limited Reasoning in First-Order Knowledge Bases with Full Introspection Artif. Intell.* 84(1-2): 209–255.

G. Lakemeyer and H.J. Levesque. 2013. Decidable Reasoning in a Logic of Limited Belief with Introspection and Unknown Individuals. *23rd International Joint Conference on Artificial Intelligence (IJCAI).*

G. Lakemeyer and H.J. Levesque. 2014. Decidable Reasoning in a Fragment of the Epistemic Situation Calculus. *Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR-14),* Vienna.

G. Lakemeyer and H.J. Levesque. 2016. Decidable Reasoning in a Logic of Limited Belief with Function Symbols. In *Proc of KR-16*, 288–297.

G. Lakemeyer and H.J. Levesque. 2019. A Tractable, Expressive, and Eventually Complete First-Order Logic of Limited Belief. In *Proc. IJCAI-19*, 1764–1771.

H.J. Levesque. 1984. A logic of implicit and explicit belief. In *Proc. of AAAI-84*, 198–202.

H.J. Levesque and G. Lakemeyer. 2001. The Logic of Knowledge Bases. MIT Press.

Y. Liu, G. Lakemeyer, and H.J. Levesque. 2004. A Logic of Limited Belief for Reasoning with Disjunctive Information. *Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR),* 587–597.

Y. Liu and H.J. Levesque. 2005. Tractable reasoning in first-order knowledge bases with disjunctive information. *In Proc. of the Twentieth National Conference on Artificial Intelligence (AAAI-05).*

J. McCarthy. 1963. Programs with common sense. M. Minsky (ed.), *Semantic Information Processing*, MIT Press, Cambridge, 403–418.

R. Montague. 1968. Pragmatics. In Klibansky, R. (Ed.), *Contemporary Philosophy*, Firenze: La Nuova Italia Editrice, 102—122.

P. Patel-Schneider. 1985. A decidable first-order logic for knowledge representation. In *Proc. of IJCAI-85*, 455–458.

D. Scott. 1970. Advice on Modal Logic. In Lambert, K. (Ed.), *Philosophical Problems in Logic*, Dordrecht, Holland: D.Reidel.

C. Schwering. 2017. A Reasoning System for a First-Order Logic of Limited Belief. In *Proc. of IJCAI-17*, 1247–1253.