A Recursively Axiomatizable Subsystem of Levesque's Logic of Only Knowing

Ho Ngoc Duc

Automated Reasoning Project, RSISE
Australian National University
Canberra ACT 0200, Australia
duc@cisr.anu.edu.au

Abstract

A complete, recursive axiomatization of a subsystem of Levesque's only-knowing logic ([5]) is given. The sublogic is obtained by relaxing some of Levesque's semantic conditions while keeping the most important definitions unchanged. The axiom system is obtained by adding an axiom of Humberstone ([4]) to a subset of the set of Levesque's axioms. The completeness result is proved using the subordination method of Hughes and Cresswell ([3]).

1 Introduction

Levesque ([5]) introduced a first-order modal logic with a modal operator for "only knowing" in order to allow non-monotonic reasoning within the language. The truth-value for the sentence "the agent only knows α" is defined by the condition: "α is true in all and only accessible worlds". "Only knowing" can be defined in terms of two operators: "knowing at least" (truth in all accessible worlds) and "knowing at most" (truth in all inaccessible worlds.)[1] Levesque has given a non-recursive axiom system for his logic, proved its completeness for the unquantified case and conjectured its completeness for first-oder case. However, the conjecture has been shown to be incorrect by Halpern and Lakemeyer ([2]).

Levesque is the first to study non-monotonic reasoning using only-knowing, but not the first to propose the study of modal operators whose truth-conditions are defined in terms of inaccessible worlds. It seems, however, that he was unaware of

[1] For the intended application, knowledge and belief need not be distinguished. For further details and intuition see the original paper [5]. In some aspects Levesque's work resembles attempts to formalize the concept of provability within the language of a logic by introducing a modal operator of provability.

works in mainstream modal logic which have proposed similar operators, but only within a propositional framework and with different motivations. In [4], Humberstone proposed an elegant axiomatization for the basic propositional modal logic with two modalities very similar to Levesque's operators of "knowing at least" and "knowing at most". Basically, the axiom system consists of two copies of the modal system K, each for one operator, and an axiom schema governing the interaction of the two modal operators.

In the present paper we shall apply Humberstone's method to the first-order case and to a stronger modal logic (namely, $K45$ rather than K for one modal operator) to axiomatize a subsystem of Levesque's logic. Rather than trying to find a non-recursive axiom system for the full logic, which is not recursively axiomatizable, we shall try to axiomatize a part of it recursively and try to make this axiomatizable part as large as possible. We hope that such a complete axiomatization will shed some light on the first-order logic of only-knowing, which is still not yet well understood. Moreover, the axiom system to be obtained might serve as a basis for a complete (although non-recursive) axiomatization of the full logic.

The remainder of the paper is organized as follows. In the next section we review briefly Levesque's logic of only-knowing. Next, we define a subsystem of it by modifying the model definition. Then we give an axiom system for the sublogic and prove its soundness and completeness.

2 A review of Levesque's logic

The language $\mathcal{OL}$ is a modal first-order dialect with identity, but without individual constants or function symbols. It is built up from an infinite stock $Pred$ of predicate symbols of every arity, an infinite collection Var of individual variables, a special two-space equality symbol $=$, and a countably infinite set C of standard names, which are considered (like the equality symbol) to be logical symbols. Formulae are formed in the standard way using the logical connectives $\neg, \wedge$, the existential quantifier $\exists$, and two unary modal operators B and N. (The other connectives and the universal quantifier $\forall$ will be used freely as abbreviations.) $B\alpha$ is read as "the agent knows (at least) α" and $N\alpha$ as "the agent knows at most α". The modal operator O of only-knowing is defined as $O\alpha := B\alpha \wedge N\neg\alpha$. Formulae without free variables are called sentences.

Our emphasis will be on sentences, and the models we define later only deal with sentences. An atomic sentence is a predicate other than $=$ applied to names. The set of atomic sentences is denoted by $Atom$. If α is a formula, x a variable, and c a standard name, then $\alpha[x := c]$ denotes the formula we get from α when all free occurrences of x are replaced by c. A sentence is called objective if it does not contain any modal operator, basic if it does not contain the operator N, and subjective if each of its predicates falls within the scope of a modal operator.

The semantics for $\mathcal{OL}$ is a variant of the possible worlds semantics for first-order modal logic. A possible world is identified with a set $w \subseteq Atom$ of atomic sentences, which can be thought of as the set of atomic sentences which are true at that world[2]. We denote the set of all worlds with W_0: $W_0 = Pow(Atom)$, where $Pow(X)$ denotes the powerset of X, for any X. Note that we are concerned with sentences only, so no variable assignment is needed.

To assign truth-values to sentences containing the operators B and N, a set $W \subseteq W_0$ of accessible possible worlds is considered. "Knowing at least" means truth in all accessible worlds (i.e., worlds in W.) "Knowing at most" means truth in all inaccessible worlds, i.e., worlds in $W_0 \setminus W$. Given a set W of worlds and a world w, the relation $(W, w) \models \alpha$ (read: the pair (W, w) satisfies the sentence α,) is defined recursively as follows:

- For any atomic ϕ, $(W, w) \models \phi$ iff $\phi \in w$
- $(W, w) \models (n_i = n_j)$ iff n_i is the same name as n_j
- $(W, w) \models \neg\alpha$ iff $W, w \not\models \alpha$
- $(W, w) \models (\alpha \wedge \beta)$ iff $(W, w) \models \alpha$ and $(W, w) \models \beta$
- $(W, w) \models \exists x \alpha$ iff for some n, $(W, w) \models \alpha[x := n]$
- $(W, w) \models B\alpha$ iff for every $w' \in W$, $W, w' \models \alpha$
- $(W, w) \models N\alpha$ iff for every $w' \notin W$, $W, w' \models \alpha$

It is obvious that subjective sentences do not depend on the w in question, and objective sentences do not depend on the W chosen. So we can write $W \models \sigma$ and $w \models \phi$ in these cases. $(W, w) \models \Gamma$ iff $(W, w) \models \alpha$ for all $\alpha \in \Gamma$.

Some comments on the semantics definition are in order. Levesque's models differ in some ways from the usual semantics for first-order modal languages. First, a subset of the set of possible worlds, that are thought to be accessible from every world, rather than an explicit accessibility relation, is considered. Second, the standard names are rigid designators, denoting the same element of the domain, namely themselves, in every world. Standard names are taken to designate distinctly and exhaustively. Related to this domain restriction is the interpretation of quantification and equality: quantification is interpreted substitutionally, and equality means syntactically equal.

To define validity, Levesque does not consider arbitrary sets of worlds, but only those which are maximal in the following sense: W and W' are equivalent if for every basic sentence A we have $W \models B\alpha$ iff $W' \models B\alpha$. It can be shown that there is a way to extend each set of world to a maximal equivalent one ([5]). A sentence α is said to be valid iff $(W, w) \models \alpha$ for all pairs (W, w) where W is maximal. Halpern

[2] A consequence of the identification of possible worlds and sets of atomic sentences is that all worlds which are indistinguishable in the given language are considered identical. This restriction is not essential, however.

and Lakemeyer ([2]) also considers a stronger notion of validity: α strongly valid if $(W, w) \models \alpha$ for all pairs (W, w), including those where W is maximal.

We conclude our review of Levesque's logic with some results about it. It is easy to see that if α is a falsifiable objective sentence then the schema $N\alpha \rightarrow \neg B\alpha$ is both valid and strongly valid for all pairs (W, w). It follows that the set of all valid (strongly valid) sentences is not recursively axiomatizable: if there were a recursive axiomatization of that set, then we could enumerate recursively all falsifiable objective sentences, contrary to the fact that there is no such enumeration.

Levesque ([5]) has given a non-recursive axiom system and proved its completeness for the propositional case. However, for the first-order case the system is shown to be incomplete by Halpern and Lakemeyer ([2]).

3 The system $K45^*$

3.1 The modified semantics

Rather than trying to find a complete but non-recursive axiomatization of Levesque's valid sentences, we shall attempt to axiomatize a subset of it, namely those valid in a wider class of models. We consider the same language but a more general definition of model. In order to axiomatize the largest possible set of valid sentences we shall try to keep our definition of models as close as possible to the original definition by Levesque. The main difference is that we consider an accessibility relation instead of a subset of the set of possible worlds, that are accessible from every world. As in [5] and [2], we identify a possible world with a set $w \subseteq Atom$ of atomic sentences. However, we do not require the set of possible worlds to be the whole powerset of $Atom$.

The reason why an accessibility relation is used is motivated as follows. Levesque assumes $K45$ as the underlying logic of (implicit) belief. This logic is known to be determined by the class of all transitive and Euclidean Kripke models. He notices that $K45$ is already determined by the class of all Kripke models (W, R, V) where $R = W \times W'$ for some subset W' of W. Thus, for simplifying matters, a set W of worlds is considered instead of a relation R. The relation R is implicitly understood as $W_0 \times W$, where W_0 is the set of all possible worlds. However, the reduction only works when no other than basic sentences are considered: it depends on the fact that a sentence α has a $K45$-model if and only if it has a model where the accessibility relation can be reduced to a (subset of the) set of worlds. The proof of this depends on the preservation of the truth value of any sentence under generated submodels ([1, Chapter 3.4]). This preservation is not guaranteed when arbitrary sentences are allowed. Roughly speaking, the reason for the failure is this: the truth of a basic sentence in a world depends only on the truth of its subsentences in those worlds that can be reached from the current world in finitely many (including zero) steps through the accessibility relation; but this is no longer true if non-basic sentences are allowed.

Thus, we modify Levesque's definition of model as follows. Let $S \subseteq W_0$ be a nonempty set of "possible worlds", whose elements are sets of atomic sentences. Let $R \subseteq S \times S$ be a transitive and Euclidean binary relation on S, and $w \in S$ a possible world (which can be viewed as a truth-value assignment to atomic sentences). Such a triple (S, R, w) is called a $K45^*$-model. The sentences of $\mathcal{OL}$ are valuated according to the following rules:

- For any atomic ϕ, $(S, R, w) \models \phi$ iff $\phi \in w$
- $(S, R, w) \models (n_i = n_j)$ iff n_i is the same name as n_j
- $(S, R, w) \models \neg\alpha$ iff $(S, R, w) \not\models \alpha$
- $(S, R, w) \models (\alpha \wedge \beta)$ iff $(S, R, w) \models \alpha$ and $(S, R, w) \models \beta$
- $(S, R, w) \models \exists x\alpha$ iff for some standard name $c \in C$, $(S, R, w) \models \alpha[x := c]$
- $(S, R, w) \models B\alpha$ iff for every $w' \in S$, if $(w, w') \in R$ then $(S, R, w') \models \alpha$
- $(S, R, w) \models N\alpha$ iff for every $w' \in S$, if $(w, w') \notin R$, $(S, R, w') \models \alpha$

A sentence α is said to be $K45^*$-valid ($\models_{K45^*} \alpha$) if $(S, R, w) \models \alpha$ for all $K45^*$-models (S, R, w). If $S = W_0 (= Pow(Atom))$ and $R = W_0 \times W$ for some $W \subseteq W_0$, then we have the original model of Levesque.

3.2 The proof theory

The axiom system $K45^*$ consists of the following parts: an adequate first-oder basis (which handles identity and standard names properly), the $K45$-axioms and rules for the operator B, the K-axioms and rules for the operator N, the Barcan sentence, and the Humberstone axiom schema governing the interaction between the two operators N and B ([4]).

Definition 1 Let L stand for B or N. Let $\mathbf{S}, \mathbf{S}', \ldots$ be strings of any length (including zero) of the operators B and N. The proof system $K45^*$ consists of the following axiom schemas and rules of inference:

FO. All instances of theorems of FOL (in the language $\mathcal{OL}$).

ID. $(n_i = n_i) \wedge (n_i \neq n_j)$ where n_i and n_j are distinct names.

BK. $B(\alpha \rightarrow \beta) \rightarrow (B\alpha \rightarrow B\beta)$

B4. $B\alpha \rightarrow BB\alpha$

B5. $\neg B\alpha \rightarrow B\neg B\alpha$

NK. $N(\alpha \to \beta) \to (N\alpha \to N\beta)$

BF. $\forall x L\alpha \to L\forall x\alpha$

HU. $\neg\mathbf{S}\neg(K\alpha \wedge N\beta) \to \mathbf{S}'(\alpha \vee \beta)$

MP. From α and $(\alpha \to \beta)$ infer β

UG. From $\alpha[x := n_1], \ldots, \alpha[x := n_k]$ infer $\forall x\alpha$, where the n_i range over all standard names in α and one not in α

BN. From α infer $B\alpha$

NN. From α infer $N\alpha$

A sentence α is called a theorem of $K45^*$ (in symbol: $\vdash_{K45^*} \alpha$) if it is an axiom or can be derived from the axioms using the specified inference rules. A sentence α is provable from a set Γ of sentences ($\Gamma \vdash_{K45^*} \alpha$) if there are some $\beta_1, \ldots, \beta_n$ from Γ such that $\beta_1 \wedge \ldots \wedge \beta_n \to \alpha$ is a $K45^*$-theorem.

It should be noted that (HU.) actually stands for infinitely many axioms, each for a pair $(\mathbf{S}, \mathbf{S}')$ of strings consisting of the operators B and N. As we only deal with the logic $K45^*$, no confusion can occur, and we can write simply $\vdash$ instead of $\vdash_{K45^*}$, or speak simply of "theorem" instead of $K45^*$-theorem etc.

4 Soundness and completeness of $K45^*$

Theorem 2 (Soundness) The logic $K45^*$ is sound wrt the class of $K45^*$-models, i.e., every $K45^*$-theorem is $K45^*$-valid.

Proof We need to check that all axioms are valid in all $K45^*$-models and the rules lead from valid sentences to valid ones. The only case that is not straightforward is (HU). Assume that there is a model (S, R, w) such that $(S, R, w) \models \neg\mathbf{S}\neg(B\alpha \wedge N\beta)$ and $(S, R, w) \not\models \mathbf{S}'(\alpha \vee \beta)$. Then (i) there is a world $w_1 \in S$ such that $(S, R, w_1) \models B\alpha \wedge N\beta$ and (ii) there exists a world w_2 such that $(S, R, w_2) \not\models \alpha \vee \beta$. If $(w_1, w_2) \in R$ then from $(S, R, w_1) \models B\alpha$ one can infer $(S, R, w_2) \models \alpha$; if $(w_1, w_2) \notin R$ then from $(S, R, w_1) \models N\beta$ one can infer $(S, R, w_2) \models \beta$, both contradicting (ii). $\Box$

Now we are going to state and prove the completeness theorem for $K45^*$. The completeness of $K45^*$ can be proved by extending Humberstone's method ([4]) to the first-order case, taking into account the restrictions on models, in particular, the special interpretations of standard names, the identification of possible worlds with sets of atomic sentences, and the requirements imposed on the accessibility relation. Our proof is based on the subordination method ([3]).

First, we need some auxiliary notions and results. We sometimes need to restrict our attention to a sublanguage of $\mathcal{OL}$ which is generated by a subset $Pred' \subseteq Pred$ of its predicates, i.e., whose formulae are built up using only the predicates from $Pred'$, the equality symbol, the standard names and the variables according to the standard formation rules. In particular, if α is a sentence of $\mathcal{OL}$ then $\mathcal{L}_\alpha$ denotes the sublanguage of $\mathcal{OL}$ generated by the set of all predicates occuring in α. The set of atoms of $\mathcal{L}_\alpha$ is denoted by $Atom_\alpha$.

As to derivability, it is clear that syntactic proofs are relative to a language and to a logic formulated in it. Since we consider only the logic $K45^*$, we have to care about the language only. Let $\mathcal{L}$ be $\mathcal{OL}$ or any of its sublanguages. A set $\Gamma \subseteq \mathcal{L}$ is called $\mathcal{L}$-consistent if there is no finite subset $\{\alpha_1, \ldots, \alpha_n\}$ of Γ such that $\vdash_{K45^*} \neg(\alpha_1 \wedge \ldots \wedge \alpha_n)$. Otherwise it is said to be $\mathcal{L}$-inconsistent. The set Γ is $\mathcal{L}$-maximal consistent if it is $\mathcal{L}$-consistent and every proper extension of it (within $\mathcal{L}$) is $\mathcal{L}$-inconsistent. Using the standard (Lindenbaum) method it can be shown that every $\mathcal{L}$-consistent set can be extended to an $\mathcal{L}$-maximal consistent one. The following lemma states some basic facts about maximal consistent sets.

Lemma 3 Suppose that Γ is $\mathcal{L}$-maximal consistent. Then

i. If $\Gamma \vdash_{K45^*} \alpha$ then $\alpha \in \Gamma$.

ii. $(\alpha \wedge \beta) \in \Gamma$ iff $\alpha \in \Gamma$ and $\beta \in \Gamma$

iii. $\neg\alpha \in \Gamma$ iff $\alpha \notin \Gamma$.

iv. $\exists x\alpha \in \Gamma$ iff $\alpha[x := c] \in \Gamma$ for some name $c \in C$.

Proof The proofs of (i.–iii.) are standard and is omitted here. The statement (iv.) holds for $K45^*$ because of the special interpretation of the standard names and the quantifiers, but it does not hold for arbitrary first-order modal systems. The 'if' direction of (iv.) is trivial: if $\alpha[x := c] \in \Gamma$ for some name $c \in C$ then $\exists x\alpha \in \Gamma$. To show the 'only if' direction, assume that $\alpha[x := c] \notin \Gamma$ for all names $c \in C$. Then $\neg\alpha[x := c] \in \Gamma$ for all $c \in C$, hence $\Gamma \vdash \forall x\neg\alpha$, by the rule (UG). Therefore $\forall x\neg\alpha \in \Gamma$, so $\exists x\alpha \notin \Gamma$. $\Box$

For any set Γ of sentences the following abbreviations are used: $B^-(\Gamma) =_{def} \{\alpha : B\alpha \in \Gamma\}$ and $N^-(\Gamma) =_{def} \{\alpha : N\alpha \in \Gamma\}$. We observe that for any $\mathcal{L}$-maximal consistent set Γ, if $\neg B\beta \in \Gamma$ then $\{\neg\beta\} \cup B^-(\Gamma)$ is consistent, and if $\neg N\beta \in \Gamma$ then $\{\neg\beta\} \cup N^-(\Gamma)$ is consistent. To prove the first observation, we assume that $\{\neg\beta\} \cup B^-(\Gamma)$ were inconsistent. Then there are some $\gamma_1, \ldots, \gamma_n$ such that $\vdash \neg(\neg\beta \wedge \gamma_1 \wedge \ldots \wedge \gamma_n)$. It follows that $\vdash \gamma_1 \wedge \ldots \wedge \gamma_n \rightarrow \beta$, hence $\vdash B\gamma_1 \wedge \ldots \wedge B\gamma_n \rightarrow B\beta$. As $B\gamma_i \in \Gamma$ for $i = 1, \ldots, n$, $B\beta$ must be in Γ too. But this is impossible because Γ is consistent. The second observation can be proved in a similar way.

Theorem 4 The logic $K45^*$ is complete wrt the class of $K45^*$-models, that is, every $K45^*$-valid sentence is a $K45^*$-theorem.

Proof We need to show that if $\neg\alpha$ is not provable in $K45^*$ then α has a model. The model of α is constructed in several steps. First, we define a tree structure that serves as the skeleton for our model. Next, we associate to every node of the tree a maximal consistent set of sentences of a suitable language. Then we associate to every node of the tree a different possible world (i.e., set of atomic sentences). Finally, an accessibility relation is defined on that set and a suitable world is chosen to complete our model.

Consider the structure $(\mathbb{Z}^*, R^+, R^-)$, where $\mathbb{Z}^*$ is the set of finite sequences of integers ordered by the following relations R^+ and R^-:

xR^+y iff $y = x * \langle n\rangle$ for some $n \geq 0$ $(n \in \mathbb{Z})$

xR^-y iff $y = x * \langle n\rangle$ for some $n < 0$ $(n \in \mathbb{Z})$

($*$ denotes the operation of concatenation.) This structure can be viewed as an infinite tree with $\mathbb{Z}^*$ as the set of nodes, the empty sequence $\langle\rangle$ as the root and $(R^+ \cup R^-)$ as the set of arcs.

We now build successively a model based on this tree. Let α be a sentence whose negation is not provable in $K45^*$. (Then $\{\alpha\}$ is $\mathcal{L}_\alpha$-consistent and can be extended to a maximal consistent set.) First, we define recursively a function T on $\mathbb{Z}^*$ that associates to every $x \in \mathbb{Z}^*$ a maximal consistent set T_x of $\mathcal{L}_\alpha$-sentences. The function T associates to the empty sequence $\langle\rangle \in \mathbb{Z}^*$ an $\mathcal{L}_\alpha$-maximal consistent set containing α. Now let $x \in \mathbb{Z}^*$ be any point in the structure $(\mathbb{Z}^*, R^+, R^-)$ (i.e., x is any finite sequence of integers.) Assume that T_x has been defined, we define $T_{x*\langle n\rangle}$ for each successor $x * \langle n\rangle$ of x $(n \in \mathbb{Z})$ as follows. Let s_e be a surjective function from the set $\{x * \langle n\rangle : n \geq 0\}$ onto the set $\{\neg\beta : \neg B\beta \in T_x\}$. (Such a function always exists because T_x has at most the cardinality of $\mathbb{Z}$.) Then for any non-negative integer $n \geq 0$, $T_{x*\langle n\rangle}$ is taken to be an $\mathcal{L}_\alpha$-maximal consistent extension of $\{s_e(x*\langle n\rangle)\} \cup B^-(T_x)$. Likewise, let s_o be a surjective function from $\{x*\langle m\rangle : m < 0\}$ onto $\{\neg\beta : \neg N\beta \in T_x\}$. For any negative integer $m < 0$, $T_{x*\langle m\rangle}$ is taken to be an $\mathcal{L}_\alpha$-maximal consistent extension of $\{s_o(x * \langle m\rangle)\} \cup B^-(T_x)$. The definition of T guarantees that the following conditions are satisfied:

Lemma 5 1. if xR^+y then $B^-(T_x) \subseteq T_y$

2. if $\neg B\alpha \in T_x$ then there is an y such that xR^+y and $\neg\alpha \in T_y$

3. if xR^-y then $N^-(T_x) \subseteq T_y$

4. if $\neg N\alpha \in T_x$ then there is an y such that xR^-y and $\neg\alpha \in T_y$

Let $x = \langle z_1, \ldots, z_m\rangle$ be any point in the tree structure. Let L_i be B if $z_i \geq 0$ and N if $z_i < 0$, for $i = 1, \ldots, m$. Using the maximal consistency of the sets associated with the nodes we can easily check that if $\alpha \in T_x$ then $\neg L_1 \ldots L_m \neg\alpha \in T_{\langle\rangle}$, and if $L_1 \ldots L_m\beta \in T_{\langle\rangle}$ then $\beta \in T_x$.

We now define $K45^*$-model based on this labeled tree structure. To define the set of worlds we join the canonical valuation of the atomic sentences at each node with a different set of atoms not occuring in $\mathcal{L}_\alpha$. Formally, let $f : \mathbb{Z}^* \mapsto Pow(Atom \setminus Atom_\alpha)$ be an one-to-one function from $\mathbb{Z}^*$ (the set of nodes) into the powerset the set of all atoms not occuring in $\mathcal{L}_\alpha$. Such a function exists because the cardinality of the former is smaller than that of the latter. For any $x \in \mathbb{Z}^*$, the world $w(x)$ is defined by: $w(x) =_{def} (T_x \cap Atom_\alpha) \cup f(x)$, that is, $w(x)$ consists of all atomic sentences occuring in T_x and all atomic sentences not occuring in $\mathcal{L}_\alpha$ which are associated with x. If $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$, hence $w(x_1) \neq w(x_2)$, i.e., the worlds correspond exactly to the elements of $\mathbb{Z}^*$[3].

Having the set $S = \{w(x) : x \in \mathbb{Z}^*\}$ of possible worlds, we now move on to define the accessibility relation $R \subseteq S \times S$. First, we include in R all pairs $(w(x), w(y))$ such that xR^+y. Next, for any pair of (not necessarily distinct) worlds $(w(x), w(y))$ such that neither xR^+y nor xR^-y, if $B^-(T_x) \subseteq T_y$ then add $(w(x), w(y))$ to R.

The main results about the relation R we have just constructed are the following:

Lemma 6 Let R be the relation defined above.

(1) $w(x)Rw(y)$ iff $B^-(T_x) \subseteq T_y$

(2) if $(w(x), w(y)) \notin R$ then $N^-(T_x) \subseteq T_y$

(3) R is transitive

(4) R is Euclidean

Proof (Lemma 6) (1) is trivial according to our construction.

To prove (2) we have to consider two cases. If xR^-y then the claim of (2) is true because of the construction of the function T. If not xR^-y then there must be some $\beta \in \mathcal{L}_\alpha$ such that $B\beta \in T_x$ and $\beta \notin T_y$ (otherwise $B^-(T_x) \subseteq T_y$ holds and $(w(x), w(y))$ would have been added to R.) Assume that $N\gamma \in T_x$ but $\gamma \notin T_y$. We show that this assumption leads to a contradiction. x and y are finite sequences of integers, thus $x = \langle z_1, \ldots, z_m \rangle$ and $y = \langle t_1, \ldots, t_n \rangle$ for some integers $z_1, \ldots, z_m, t_1, \ldots, t_n$. For $i = 1, \ldots, m$ let L_i be B if $z_i \geq 0$ and N if $z_i < 0$, and for $j = 1, \ldots, n$ let L'_i be B if $t_j \geq 0$ and N if $t_j < 0$. From $B\beta \wedge N\gamma \in T_x$ we can infer $\neg L_1 \ldots L_m \neg(B\beta \wedge N\gamma) \in T_{\langle\rangle}$. Consider the instance of axiom (HU.) $\neg L_1 \ldots L_m \neg(B\beta \wedge N\gamma) \rightarrow L'_1 \ldots L'_n(\beta \vee \gamma)$. $T_{\langle\rangle}$ contains this sentence because of its maximal consistency, and $\neg L_1 \ldots L_m \neg(B\beta \wedge N\gamma) \in T_{\langle\rangle}$ as argued, so $L'_1 \ldots L'_n(\beta \vee \gamma)$ must belong to $T_{\langle\rangle}$. It follows that $(\beta \vee \gamma) \in T_y$, which is impossible because neither β nor γ belongs to T_y.

To show (3), let us assume that $w(x)Rw(y)$ and $w(y)Rw(z)$. Assume that $B\beta \in T_x$. Because of (1) it suffices to show that $\beta \in T_z$. By axiom B4, $BB\beta \in T_x$, thus $B\beta \in T_y$ and so $\beta \in T_z$.

Finally, let $w(x)Rw(y)$ and $w(x)Rw(z)$. We show that $w(y)Rw(z)$. Assume that $B\beta \in T_y$, we show that $\beta \in T_z$. Suppose that this is not the case, then $\neg\beta \in T_z$, so

[3]Without the diversion through the sublanguage $\mathcal{L}_\alpha$, this correspondence is not guaranteed, because different nodes in $\mathbb{Z}^*$ can be labeled with the same maximal consistent set of sentences.

$\neg B\beta \in T_x$. From axiom B5 we can infer $B\neg B\beta \in T_x$, so $\neg B\beta \in T_y$, contradicting our assumption that $B\beta \in T_y$. Thus, we have proved that R is Euclidean. $\square$

The last step to prove our completeness theorem is the following lemma:

Lemma 7 Let $S = \{w(x) : x \in \mathbb{Z}^*\}$ and $R \subseteq S \times S$ be the relation defined above. For all $w(x) \in S$ and all $\beta \in \mathcal{L}_\alpha$, $(S, R, w(x)) \models \beta$ if and only if $\beta \in T_x$.

Proof (Lemma 7) We prove the lemma by induction on the complexity of sentences. If $\beta \in Atom_\alpha$ then $\beta \in T_x$ iff $\beta \in w(x)$ iff $(S, R, w(x)) \models \beta$.

If β is $\neg\gamma$ then $(S, R, w(x)) \models \beta$ iff $(S, R, w(x)) \not\models \gamma$ iff $\gamma \notin T_x$ (by induction hypothesis) iff $\neg\gamma \in T_x$ (by $\mathcal{L}_\alpha$-maximal consistency of T_x and lemma 3) iff $\beta \in T_x$.

Suppose that β is $(\gamma_1 \wedge \gamma_2)$. Then $(S, R, w(x)) \models \beta$ iff $(S, R, w(x)) \models \gamma_1$ and $(S, R, w(x)) \models \gamma_2$ iff $\gamma_1 \in T_x$ and $\gamma_2 \in T_x$ (by induction hypothesis) iff $(\gamma_1 \wedge \gamma_2) \in T_x$ (by $\mathcal{L}_\alpha$-maximal consistency of T_x and lemma 3) iff $\beta \in T_x$.

If β is $\exists x\gamma$ then $(S, R, w(x)) \models \beta$ iff $(S, R, w(x)) \models \gamma[x := c]$ for some standard name $c \in C$ iff $\gamma[x := c] \in T_x$ for some $c \in C$ (by induction hypothesis) iff $\exists x\gamma \in T_x$ (by $\mathcal{L}_\alpha$-maximal consistency of T_x and lemma 3) iff $\beta \in T_x$.

Now let β be $B\gamma$. For all $w(y) \in S$, if $B\gamma \in T_x$ and $w(x)Rw(y)$ then $\gamma \in T_y$, by lemma 6, (1). By induction hypothesis $(S, R, w(y)) \models \gamma$, so $(S, R, w(x)) \models B\gamma$, i.e., $(S, R, w(x)) \models \beta$. Conversely, if $B\gamma \notin T_x$ then $\neg B\gamma \in T_x$, hence $\neg\gamma \in T_y$ for some y such that xR^+y, according to lemma 5, so $\gamma \notin T_y$. By induction hypothesis, $(S, R, w(y)) \not\models \gamma$. But from xR^+y it follows that $w(x)Rw(y)$, so $(S, R, w(y)) \not\models B\gamma$.

Finally, let β be $N\gamma$. For all $w(y) \in S$, if $N\gamma \in T_x$ and $(w(x), w(y)) \notin R$ then $\gamma \in T_y$, by lemma 6, (2). By induction hypothesis $(S, R, w(y)) \models \gamma$, so $(S, R, w(x)) \models N\gamma$, i.e., $(S, R, w(x)) \models \beta$. Conversely, if $N\gamma \notin T_x$ then $\neg N\gamma \in T_x$, hence $\neg\gamma \in T_y$ for some y such that xR^-y, because of lemma 5, so $\gamma \notin T_y$. By induction hypothesis, $(S, R, w(y)) \not\models \gamma$. But from xR^-y it follows that $(w(x), w(y)) \notin R$, so $(S, R, w(y)) \not\models N\gamma$. $\square$

From lemma 7 we can infer $(S, R, w(\langle\rangle)) \models \alpha$, because $\alpha \in T_{\langle\rangle}$. Thus, every valid $K45^*$-sentence is derivable in $K45^*$. $\square$

5 Conclusion

We have defined a subsystem of Levesque's logic of only-knowing and prove its completeness. Our logic is obtained by dropping some of Levesque's restrictions on models, thus allowing a larger class of models. Its axiomatization is obtained by adding an axiom of Humberstone to a subset of the set of Levesque's axioms. The main completeness is obtained using the subordination technique, combined with some tricks to take care of the restrictions imposed on models. The tricks are necessary because we try to modify Levesque's definition of models as little as possible.

As the works of Levesque ([5]) and Halpern and Lakemeyer ([2]) show, the interaction of the operators B and N is very subtle, which makes a complete axiomatization of the original logic of only-knowing very hard to find. Instead of trying to capture this interaction "in one go" with a single (non-recursive) axiom schema (in the style of Levesque's N versus B axiom), we try to axiomatize only parts of this interaction using a recursive axiom schema. The obvious question to be asked now is: which (non-recursive) axioms must be added to $K45^*$ in order to yield a complete axiom system for Levesque's logic? We leave this question for future investigation.

References

[1] B. Chellas. *Modal Logic. An Introduction.* Cambridge UP, 1980.

[2] J. Y. Halpern and G. Lakemeyer. Levesque's axiomatization of only knowing is incomplete. *Artificial Intelligence*, 74:381–387, 1995.

[3] G. E. Hughes and M. J. Cresswell. *A Companion to Modal Logic.* Methuen, London, 1984.

[4] I. L. Humberstone. Inaccessible worlds. *Notre Dame Journal of Formal Logic*, 24(3):346–352, 1983.

[5] H. Levesque. All I know: A study in autoepistemic logic. *Artificial Intelligence*, 42:263–309, 1990.