

A non-preferential semantics of non-monotonic modal logic

Michael Kaminski

Department of Computer Science
Technion – Israel Institute of Technology
Haifa 32000 Israel

*Dedicated to Victor Marek
on his 65th birthday*

Abstract

We present a new semantical description of non-monotonic modal logic whose underlying (monotonic) modal logic S is characterized by a class of *cluster-decomposable* Kripke interpretations: it is shown that S -expansions coincide with the theories of *isolated clusters* of the *canonical* model of S .

1 Introduction

Non-monotonic modal logic is intended to simulate the process of human reasoning by providing a formalism for deriving consistent conclusions from an incomplete description of the world. The language of non-monotonic modal logic contains a modal connective L (known) and its dual M (consistent), and the logic itself is defined *syntactically* by *expansions* for a set of axioms, which are fixpoints of a *monotonic* (provability) operator acting on sets of modal formulas. Loosely speaking, expansions are obtained by augmenting the underlying modal logic with the “inference rule” of the form

$$\text{“if } \neg\varphi \text{ is not derivable, then } \varphi \text{ is possible,”} \quad (1)$$

called *possibilitation* ((McDermott 1982)), or *negative introspection* ((Marek & Truszczyński 1993, p. 224)).

However, the fixpoint approach is not sufficiently clear because of its circularity/self-reference. This deficiency motivated the *minimal model semantics* of non-monotonic logic introduced by Schwarz in (Schwarz 1992). In this semantics the expansions for a set of axioms A are the theories of the models of A which are minimal with respect to a certain partial order, called the *preference relation*.

Schwarz’s semantical description of expansions applies to any underlying modal logic that can be characterized by a *cluster closed* class of Kripke interpretations, i.e., a class of Kripke interpretations that either is closed under *concatenation* with clusters, or consists

of *cluster-decomposable* Kripke interpretations and satisfies some natural closure condition.¹ In particular, modal logics K , T , and $S4$ are characterized by classes of Kripke interpretations possessing the former property, and logics $S4F$, $Sw5$, and a very important $KD45$ (that underlies *autoepistemic* logic) are characterized by classes of Kripke interpretations possessing the latter.

In this paper we present a *non-preferential* semantics of non-monotonic modal logic whose underlying modal logic S is characterized by a class of cluster-decomposable Kripke interpretations. We show that S -expansions for a set of axioms A coincide with the theories of *isolated clusters* of the *canonical* S, A -model. As a consequence of our approach we obtain that *autoepistemic* logic separates (models of) *knowledge* from (models of) *belief*.

The rest of the paper is organized as follows. In the next section we recall the definition of propositional non-monotonic modal logic and list some of its basic properties. Our main result – the canonical model semantics of non-monotonic modal logic is presented in Section 3. Then in Section 4 we discuss sufficient conditions for an expansion to be the theory of an isolated cluster of the canonical model. Finally, in Section 5 we apply our semantical description of expansions to autoepistemic logic.

2 Propositional non-monotonic logic

In this section we recall some basic properties of propositional monotonic and non-monotonic modal logics.

2.1 Propositional modal logic

We start with the language of classical propositional logic that contains propositional variables and only two classical propositional connectives \perp (a logical constant *falsity*) and \supset (implication). Connectives \top (*truth*), \neg (negation), \wedge (conjunction), and \vee (disjunction) are defined in a usual manner, e.g., $\neg\varphi$ is $\varphi \supset \perp$.

¹It should be also noted that there is a minimal model semantics of *ground* non-monotonic modal logic whose underlying (monotonic) modal logic is characterized by a class of cluster-decomposable Kripke interpretations, see (Donini, Nardi, & Rosati 1997).

The language of propositional modal logic is obtained from the language of classical propositional logic by extending it with a modal connective L (known). As usual, the dual connective M (consistent) is defined by $\neg L \neg$. Formulas not containing L are called *ground* and the set of all ground formulas is denoted by \mathbf{GFm} .

The modal logic K results from classical propositional logic² by adding the inference rule

NEC $\varphi \vdash L\varphi$

called *necessitation* and the axiom scheme

k $L(\varphi \supset \psi) \supset (L\varphi \supset L\psi)$.

The “classical” modal logics are obtained by adding to K all instances of some axiom schemes, e.g.,

t $L\varphi \supset \varphi$
d $M\top$
4 $L\varphi \supset LL\varphi$
5 $ML\varphi \supset L\varphi$
b $ML\varphi \supset \varphi$
w5 $ML\varphi \supset (\varphi \supset L\varphi)$
f $(\varphi \wedge ML\psi) \supset L(M\varphi \vee \psi)$

Adding **t** to K results in T, adding **4** to T results in S4, adding **5** to S4 results in S5, adding **w5** to S4 results in Sw5, and adding **f** to S4 results in S4F, etc., see (Marek & Truszczyński 1993, page 197).

For a modal logic S and a set of formulas A , called (proper) *axioms*, we define the (monotonic) *theory* of A , denoted $\mathbf{Th}_S(A)$, as $\mathbf{Th}_S(A) = \{\varphi : A \vdash_S \varphi\}$. As usual, we write $A \vdash_S \varphi$, if there exists a sequence of formulas $\varphi_1, \varphi_2, \dots, \varphi_n = \varphi$ such that each φ_i is an axiom from S, or belongs to A , or is obtained from some of the formulas $\varphi_1, \varphi_2, \dots, \varphi_{i-1}$ by *modus ponens* or *necessitation*.

The Kripke semantics of propositional modal logic is defined as follows. A *Kripke interpretation* is a triple $\mathfrak{M} = \langle U, R, I \rangle$, where U is a non-empty set of *possible worlds*, $R \subseteq U \times U$ is an *accessibility* relation on U , and I is an assignment to each world in U of a set of propositional variables. *In what follows we identify a (ground) propositional interpretation with the set of all propositional variables it satisfies.*

Definition 1 Let $\mathfrak{M} = \langle U, R, I \rangle$ be a Kripke interpretation and let $u \in U$. We say that the pair (\mathfrak{M}, u) *satisfies* a formula φ , denoted $(\mathfrak{M}, u) \models \varphi$, if the following holds.

- If φ is a propositional variable p , then $(\mathfrak{M}, u) \models \varphi$ if and only if $p \in I(u)$;
- $(\mathfrak{M}, u) \not\models \perp$;
- $(\mathfrak{M}, u) \models \varphi \supset \psi$ if and only if $(\mathfrak{M}, u) \not\models \varphi$ or $(\mathfrak{M}, u) \models \psi$;
- $(\mathfrak{M}, u) \models L\varphi$ if and only if for every v such that $(u, v) \in R$, $(\mathfrak{M}, v) \models \varphi$.

²We assume a standard set of propositional axioms and the inference rule *modus ponens*.

The set of all formulas satisfied by (\mathfrak{M}, u) is called the *theory* of (\mathfrak{M}, u) and is denoted by $\mathbf{Th}(\mathfrak{M}, u)$.

We say that a Kripke interpretation \mathfrak{M} *satisfies* a formula φ , denoted $\mathfrak{M} \models \varphi$, if and only if for every $u \in U$, $(\mathfrak{M}, u) \models \varphi$, and we say that \mathfrak{M} *satisfies* a set of formulas A (or is a *model* of A), denoted $\mathfrak{M} \models A$, if and only if $\mathfrak{M} \models \varphi$ for every $\varphi \in A$. The set of all formulas satisfied by \mathfrak{M} is called the *theory* of \mathfrak{M} and is denoted by $\mathbf{Th}(\mathfrak{M})$.

For Proposition 2 below we shall need the following notation.

Let R be a binary relation on a set U and let i be a non-negative integer. The binary relation R^i on U is defined by the following induction.

- $R^0 = \{(u, u) : u \in U\}$, i.e., R^0 is the *diagonal* relation on U , and
- $R^{i+1} = \{(u, v) : \text{for some } w \in U, (u, w) \in R^i \text{ and } (w, v) \in R\}$.

That is, $(u, v) \in R^i$ if and only if there exists a path (of edges) of length i from u to v .

Let $\mathfrak{M} = \langle U, R, I \rangle$ be a Kripke interpretation and let $u \in U$. We denote by \mathfrak{M}^u the Kripke interpretation $\langle U^u, R^u, I^u \rangle$, where $U^u = \{v : (u, v) \in \bigcup_{i=0}^{\infty} R^i\}$,³ and $R^u = R|_{U^u}$ and $I^u = I|_{U^u}$ are the restrictions of R and I on U^u , respectively.

Proposition 2 Let $\mathfrak{M} = \langle U, R, I \rangle$ be a Kripke interpretation and let $u \in U$. Then for a formula φ , $(\mathfrak{M}, u) \models \varphi$ if and only if $(\mathfrak{M}^u, u) \models \varphi$.

The proof is by a straightforward induction on the formula complexity and is omitted.

Definition 3 Let \mathcal{C} be a class of Kripke interpretations and let S be a modal logic. We say that S is *characterized* by \mathcal{C} , if the following holds. For every set of formulas A and every formula φ , $A \vdash_S \varphi$ if and only if for every Kripke interpretation $\mathfrak{M} \in \mathcal{C}$, $\mathfrak{M} \models A$ implies $\mathfrak{M} \models \varphi$.⁴

The logic K is characterized by the class of all Kripke interpretations. In particular, a set of formulas is consistent if and only if it has a model. The logic K + **d** is characterized by the class of all Kripke interpretations whose accessibility relation has no *dead ends*⁵, the logic K + **b** is characterized by the class of all Kripke interpretations with a symmetric accessibility relation, and the logic K + **5** is characterized by the class of all Kripke interpretations with a *euclidean* accessibility relation⁶, etc.

Our semantical description of non-monotonic logic is based on the notion of the *canonical* model which

³That is, U^u consists of all worlds of U which are reachable from u by means of R .

⁴That is, \mathcal{C} is sound and complete for S.

⁵A Kripke interpretation $\mathfrak{M} = \langle U, R, I \rangle$ has no dead ends, if for every $u \in U$ there exists a $v \in U$ such that $(u, v) \in R$.

⁶A relation $R \subseteq U \times U$ is euclidean, if for every $u, v, w \in U$, the containment $(u, v), (u, w) \in R$ implies $(v, w) \in R$.

we recall below (see (Marek & Truszczyński 1993, Sections 7.2 and 7.3)). In what follows S and A are a modal logic and a set of formulas, respectively.

Definition 4 A set of formulas Γ is said to be S, A -consistent, if for no finite subset Γ' of Γ , $A \vdash_S \neg \bigwedge_{\varphi \in \Gamma'} \varphi$.

Definition 5 Maximal (with respect to inclusion) S, A -consistent sets of formulas are called S, A -maximal.

Example 6 Let $\mathfrak{M} = \langle U, R, I \rangle$ be a model of S that satisfies A and let $u \in U$. It follows from the definition of \models that $\mathbf{Th}(\mathfrak{M}, u)$ is S, A -maximal.

Proposition 7 (See (Marek & Truszczyński 1993, Lemma 7.30, p. 204).) *Let Γ be an S, A -maximal set of formulas. Then for each formula φ , exactly one member of $\{\varphi, \neg\varphi\}$ is in Γ .*

To define the S, A -canonical model we need one more bit of notation. For a set of formulas Γ we define the set of formulas Γ^- by

$$\Gamma^- = \{\varphi : L\varphi \in \Gamma\}.$$

Definition 8 The S, A -canonical model $\mathfrak{M}_{S, A} = \langle U_{S, A}, R_{S, A}, I_{S, A} \rangle$ is defined as follows.

- $U_{S, A}$ is the set of all S, A -maximal sets of formulas.
- $R_{S, A} = \{(u, v) \in U_{S, A} \times U_{S, A} : u^- \subseteq v\}$.
- $I_{S, A}(u)$ is the set of all propositional variables which belong to u .

Theorem 9 (See (Marek & Truszczyński 1993, Theorem 7.32, p. 206).) *For any formula φ and any $u \in U_{S, A}$, $(\mathfrak{M}_{S, A}, u) \models \varphi$ if and only if $\varphi \in u$.*

Remark 10 Let $\mathfrak{M} = \langle U, R, I \rangle$ be a model of S that satisfies A and let $u, v \in U$ be such that $(u, v) \in R$. It follows from the definition of \models that $(\mathbf{Th}(\mathfrak{M}, u))^- \subseteq \mathbf{Th}(\mathfrak{M}, v)$. This together with Example 6 imply that \mathfrak{M} naturally embeds into $\mathfrak{M}_{S, A}$ via

$$u \rightarrow \mathbf{Th}(\mathfrak{M}, u), \quad u \in U.^7$$

Therefore, $\mathfrak{M}_{S, A}$ contains *all* information about A , because it “includes” all models of S which satisfy A .

For Schwarz’s minimal model semantics of non-monotonic modal logic in Section 2.3 we shall also need the following definition.

Definition 11 Kripke interpretations of the form $\langle U, U \times U, I \rangle$ are called *clusters* and are denoted by $\langle U, I \rangle$.

It can be readily seen that $S5$ is characterized by the class of all clusters, (e.g., see (Marek & Truszczyński 1993, Theorem 7.52, p. 216)).

The following simple lemma motivates Definition 13 below.

⁷This embedding is, actually, the *filtration* of \mathfrak{M} through the set of *all* modal formulas, see (Hughes & Cresswell 1968, pp. 136).

Lemma 12 Let $\mathfrak{M}' = \langle U', I' \rangle$ and $\mathfrak{M}'' = \langle U'', I'' \rangle$ be such that

$$\mathbf{Th}(\mathfrak{M}') \cap \mathbf{GFm} = \mathbf{Th}(\mathfrak{M}'') \cap \mathbf{GFm}.$$

Then $\mathbf{Th}(\mathfrak{M}') = \mathbf{Th}(\mathfrak{M}'')$.

The proof of the lemma is by a straightforward induction on the formula complexity. Alternatively, the lemma immediately follows from (Marek & Truszczyński 1993, Theorem 8.10, p. 228; Theorem 8.12, p. 229; and Corollary 8.19, p. 233). We leave the details to the reader.

Definition 13 A cluster $\mathfrak{M} = \langle U, I \rangle$ is called *maximal*, if for each (ground) propositional interpretation ι satisfying $\mathbf{Th}(\mathfrak{M}) \cap \mathbf{GFm}$ there is a world $u \in U$ such that $\iota = I(u)$.⁸

For a cluster \mathfrak{M} we denote by $[\mathfrak{M}]$ the cluster $\langle U, I \rangle$, where

- U consists of all propositional interpretations satisfying $\mathbf{Th}(\mathfrak{M}) \cap \mathbf{GFm}$ and
- I is the identity function on U .

By definition, $[\mathfrak{M}]$ is maximal, and, by Lemma 12, $\mathbf{Th}([\mathfrak{M}]) = \mathbf{Th}(\mathfrak{M})$.

2.2 Propositional non-monotonic modal logic

Here we recall the definition of propositional non-monotonic modal logic based on the McDermott and Doyle fixpoint equation ((McDermott & Doyle 1980)). Definition 14 below is an extension to a modal logic S of McDermott’s original definition ((McDermott 1982)) that only dealt with the classical modal logics T , $S4$, or $S5$. A general form of McDermott’s definition is as follows.

Definition 14 ((Marek & Truszczyński 1993, Definition 9.2, p. 252)) Let A be a set of modal formulas (axioms) and let S be a modal logic. An S -consistent set of formulas E is called an S -expansion for A if

$$E = \mathbf{Th}_S(A \cup \{M\varphi : E \not\models_S \neg\varphi\}). \quad (2)$$

That is, S -expansions for A can be thought of as the “deductive closures” of A in S extended with the *possibilitation* rule (1).

The following property of expansions immediately follows from (Marek & Truszczyński 1993, Theorem 8.12, p. 229, and Theorem 9.4, p. 253).

Proposition 15 *Let A be a set of formulas, S be a modal logic, and let E be an S -expansion for A . Then E is the theory of a cluster.*

Proposition 16 *Let \mathfrak{M} be a cluster satisfying A . Then $\mathbf{Th}(\mathfrak{M})$ is an S -expansion for A if and only if*

$$\mathbf{Th}(\mathfrak{M}) = \mathbf{Th}_S(A \cup \{M\varphi : \mathfrak{M} \models M\varphi\}).$$

⁸Recall that we identify a propositional interpretation with the set of all propositional variables it satisfies.

Proof By (2), it suffices to show that for each formula φ ,

$$\mathfrak{M} \models M\varphi \quad (3)$$

if and only if

$$\mathbf{Th}(\mathfrak{M}) \not\models_S \neg\varphi. \quad (4)$$

By soundness and completeness of the Kripke semantics, (4) is equivalent to $\neg\varphi \notin \mathbf{Th}(\mathfrak{M})$, which, in turn, is equivalent to (3), because \mathfrak{M} is a cluster. ■

2.3 The minimal model semantics

This section contains Schwarz's semantics of propositional non-monotonic modal logic ((Schwarz 1992)). It is based on Definitions 17–20 below.

Definition 17 Let $\mathfrak{M}' = \langle U', R', I' \rangle$ and $\mathfrak{M}'' = \langle U'', R'', I'' \rangle$ be Kripke interpretations such that $U' \cap U'' = \emptyset$.⁹ The *concatenation* of \mathfrak{M}'' to \mathfrak{M}' , denoted $\mathfrak{M}' \odot \mathfrak{M}''$, is the Kripke interpretation $\langle U, R, I \rangle$, where

- $U = U' \cup U''$,
- $R = R' \cup (U' \times U'') \cup R''$, and
- I is defined by $I(u) = \begin{cases} I'(u) & \text{if } u \in U' \\ I''(u) & \text{if } u \in U'' \end{cases}$.

Definition 18 ((Schwarz 1992)) Let $\mathfrak{M} = \langle U, I \rangle$ and $\mathfrak{M}' = \langle U', R, I' \rangle$ be a cluster and a Kripke interpretation, respectively. We say that $\mathfrak{M}' \odot \mathfrak{M}$ is *preferred* over \mathfrak{M} , denoted $\mathfrak{M}' \odot \mathfrak{M} \sqsubset \mathfrak{M}$, if there is a world $u' \in U'$ and a ground formula θ such that $I'(u') \models \theta$, but $\mathfrak{M} \models \neg\theta$.

Definition 19 ((Schwarz 1992)) Let \mathcal{C} be a class of Kripke interpretations. A cluster \mathfrak{M} is called *\mathcal{C} -minimal* for a set of modal formulas A if $\mathfrak{M} \models A$ and for every Kripke interpretation \mathfrak{M}' such that $\mathfrak{M}' \odot \mathfrak{M} \in \mathcal{C}$ and $\mathfrak{M}' \odot \mathfrak{M} \models A$, $\mathfrak{M}' \odot \mathfrak{M} \not\models A$.

Definition 20 ((Schwarz 1992)) A class \mathcal{C} of Kripke interpretations is called *cluster closed* if it contains all clusters and at least one of the two following conditions is satisfied.

1. For every cluster \mathfrak{M} and every Kripke interpretation $\mathfrak{M}' \in \mathcal{C}$, $\mathfrak{M}' \odot \mathfrak{M} \in \mathcal{C}$.
2. Every Kripke interpretation in \mathcal{C} is of the form $\mathfrak{M}' \odot \mathfrak{M}$, where \mathfrak{M} is a cluster. Moreover, for every $\mathfrak{M}' \odot \mathfrak{M} \in \mathcal{C}$, where \mathfrak{M} is a cluster, and every cluster \mathfrak{N} , $\mathfrak{M}' \odot \mathfrak{N} \in \mathcal{C}$.

At last, we have arrived at Schwarz's description of S-expansions.

Theorem 21 ((Schwarz 1992, Theorem 3.1), see also (Marek & Truszczyński 1993, Corollary 9.22, p. 266).) *Let S be a modal logic characterized by a class C of cluster closed Kripke interpretations and let A be a set of formulas. A set of formulas E is an S-expansion for A if and only if there exists a cluster M such that M is C-minimal for A and $E = \mathbf{Th}(\mathfrak{M})$.*

⁹In what follows, renaming the interpretations' worlds, if necessary, we always assume that their sets of worlds are disjoint.

3 The canonical model semantics

This section contains the main result of our paper – the canonical model semantics of non-monotonic modal logic, see Theorem 24 below. To state the theorem we need the following definitions.

Definition 22 A connected component $\mathfrak{M} = \langle U, R, I \rangle$ of $\mathfrak{M}_{S,A}$ is called an *isolated cluster*, if $R = U \times U$.

Definition 23 (Cf. case 2 of Definition 20 and (Donini, Nardi, & Rosati 1997, Definition 3.11).) Kripke interpretations of the form $\mathfrak{M}' \odot \mathfrak{M}$, where \mathfrak{M} is a cluster, are called *cluster-decomposable*, and \mathfrak{M} is called the *final cluster* of $\mathfrak{M}' \odot \mathfrak{M}$.¹⁰

Theorem 24 *Let S be a modal logic characterized by a class of cluster-decomposable Kripke interpretations and let A be a set of formulas. A set of formulas E is an S-expansion for A if and only if E is the theory of an isolated cluster of $\mathfrak{M}_{S,A}$.*

The proof of the “if” part of the theorem is presented in Section 3.1 and the proof of the “only if” part is presented in Section 3.2.

3.1 Proof of the “if” part of Theorem 24

Let $\mathfrak{M} = \langle U, I \rangle$ be an isolated cluster of $\mathfrak{M}_{S,A}$ and assume to the contrary that $\mathbf{Th}(\mathfrak{M})$ is not an S-expansion for A. Then, by Proposition 16, there exists a formula $\theta \in \mathbf{Th}(\mathfrak{M})$ such that

$$A \cup \{M\varphi : \mathfrak{M} \models M\varphi\} \not\models_S \theta.$$

By completeness, there exists a model $\mathfrak{M}' = \langle U', R', U' \rangle$ of S such that

$$\mathfrak{M}' \models A \cup \{M\varphi : \mathfrak{M} \models M\varphi\}, \quad (5)$$

but for some $u' \in U'$, $(\mathfrak{M}', u') \not\models \theta$.

Since \mathfrak{M}' satisfies both S and A, $\mathbf{Th}(\mathfrak{M}', u') \in U_{S,A}$, see Remark 10. However, $\mathbf{Th}(\mathfrak{M}', u') \notin U$, because $\theta \in \mathbf{Th}(\mathfrak{M})$. Thus, we shall arrive at a contradiction with our assumption that \mathfrak{M} is an isolated cluster of $\mathfrak{M}_{S,A}$, if we show that $\mathbf{Th}(\mathfrak{M}', u')$ is connected to some world in U via accessibility relation $R_{S,A}$. In fact, we shall prove that $\mathbf{Th}(\mathfrak{M}', u')$ is connected to each world $u \in U$.

For the proof assume to the contrary that $(u', u) \notin R_{S,A}$. That is, $u'^- \not\subseteq u$. Hence, for some formula φ , $(\mathfrak{M}', u') \models L\varphi$, but $\varphi \notin u$. Then, by Proposition 7, $\neg\varphi \in u$. Since \mathfrak{M} is a cluster, by Theorem 9, $\mathfrak{M} \models M\neg\varphi$. Thus, by (5), $(\mathfrak{M}', u') \models M\neg\varphi$, which contradicts our assumption $(\mathfrak{M}', u') \models L\varphi$. This completes the proof of the “if” part of Theorem 24.

¹⁰An axiomatization of the modal logic characterized by the class of all cluster-decomposable Kripke interpretations can be found in (Tiomkin & Kaminski 2007).

3.2 Proof of the “only if” part of Theorem 24

We precede the proof of the “only if” part of Theorem 24 with the following definition and auxiliary lemmas.

Definition 25 (Cf. (Marek & Truszczyński 1993, Definition 11.37, p. 344).) Let $\mathfrak{M} = \langle U, R, I \rangle$ be a Kripke interpretation and let $U' \subseteq U$. The cluster $\langle U', I|_{U'} \rangle$ is called a *terminal cluster* of \mathfrak{M} , if it satisfies conditions 1 and 2 below.

1. The restriction of R on U' is a total relation, i.e., $U' \times U' \subseteq R$.
2. For every $v \in U$ the following holds. If for some $u \in U'$, $(u, v) \in R$, then $v \in U'$.

Obviously, final clusters (see Definition 23) are terminal. However, terminal clusters are not necessarily final. This can happen for one of the following reasons.

1. A Kripke interpretation may have more than one terminal cluster.
2. Even if a Kripke interpretation has only one terminal cluster, the former may contain a world not connected to *each* world of the latter.

Lemma 26 *Let S be a modal logic, A be a set of formulas, $\mathfrak{M}' = \langle U', I' \rangle$ be a maximal cluster satisfying both S and A , and let $U = \{\mathbf{Th}(\mathfrak{M}', u') : u' \in U'\}$. Then $\mathfrak{M} = \langle U, I_{S,A}|_U \rangle$ is a terminal cluster of $\mathfrak{M}_{S,A}$.*

Proof Condition 1 of Definition 25 is satisfied by Remark 10.

For the proof of condition 2 of Definition 25, let $u = \mathbf{Th}(\mathfrak{M}', u') \in U$ and $v \in U_{S,A}$ be such that $(u, v) \in R_{S,A}$, i.e., $u^- \subseteq v$. By the definition of U , we have to show that for some $v' \in U'$, $v = \mathbf{Th}(\mathfrak{M}', v')$.

First, we observe that

$$\mathbf{Th}(\mathfrak{M}') \subseteq v. \quad (6)$$

Indeed, let $\varphi \in \mathbf{Th}(\mathfrak{M}')$. Then $\mathfrak{M}' \models L\varphi$. In particular, $(\mathfrak{M}', u') \models L\varphi$, implying $L\varphi \in u$, and $\varphi \in v$ follows from $u^- \subseteq v$.

Therefore,

$$\mathbf{Th}(\mathfrak{M}') \cap \mathbf{GF}m \subseteq v \cap \mathbf{GF}m = \mathbf{Th}(I_{S,A}(v)).^{11} \quad (7)$$

Since \mathfrak{M}' is maximal, it follows from (7) that there is a world $v' \in U'$ such that

$$I'(v') = I_{S,A}(v), \quad (8)$$

and the proof will be complete if we show that $\mathbf{Th}(\mathfrak{M}', v') = v$. That is, for each formula φ , $(\mathfrak{M}', v') \models \varphi$ if and only if $\varphi \in v$. The proof is by induction on the complexity of φ .

Basis: The case of \perp is trivial and the case of a propositional variable immediately follows from (8) and the definition of $I_{S,A}$ (Definition 8).

¹¹Of course, by $\mathbf{Th}(I_{S,A}(v))$ we mean the *ground* theory of $I_{S,A}(v)$. Recall that we identify a propositional interpretation with the set of all propositional variables it satisfies.

Induction step: The case of implication is immediate and is omitted. Let φ be of the form $L\psi$ and assume $(\mathfrak{M}', v') \models L\psi$. Then $\mathfrak{M}' \models L\psi$, because \mathfrak{M}' is a cluster, and $L\psi \in v$ follows from (6).

Conversely, assume $L\psi \in v$. Then, by Theorem 9, $(\mathfrak{M}_{S,A}, v) \models L\psi$, implying $(\mathfrak{M}_{S,A}, u) \models ML\psi$, because $(u, v) \in R_{S,A}$. Applying Theorem 9 one more time, we see that $ML\psi \in u = \mathbf{Th}(\mathfrak{M}', u')$. In other words, $(\mathfrak{M}', u') \models ML\psi$. Since \mathfrak{M}' is a cluster, $(\mathfrak{M}', v') \models L\psi$, which completes the proof of the lemma. \blacksquare

The next lemma deals with the axiom schemes **bm** and **5m** below.

bm $MLM\varphi \supset M\varphi$

5m $MLM\varphi \supset LM\varphi$

Remark 27 Note that schemes **5m** and **bm** result in substitution of $M\varphi$ for φ in **b** and **5**, respectively, which motivates our notation.

Lemma 28 *Each cluster-decomposable Kripke interpretation satisfies both **bm** and **5m**.*

Proof Let $\mathfrak{M} = \langle U, R, I \rangle$ be a cluster-decomposable Kripke interpretation, $u \in U$, and let φ be a formula such that $(\mathfrak{M}, u) \models MLM\varphi$. We have to show that $(\mathfrak{M}, u) \models M\varphi$ and $(\mathfrak{M}, u) \models LM\varphi$.

It follows from $(\mathfrak{M}, u) \models MLM\varphi$ that there is a world $v \in U$ such that $(\mathfrak{M}, v) \models LM\varphi$. Let $\langle U', I|_{U'} \rangle$ be the final cluster of \mathfrak{M} . Since $\{v\} \times U' \subseteq R$, there is a world $v' \in U'$ such that $(\mathfrak{M}, v') \models \varphi$. Thus, for each world $w \in U$, $(\mathfrak{M}, w) \models M\varphi$, because $(w, v') \in R$. This implies both $(\mathfrak{M}, u) \models M\varphi$ and $(\mathfrak{M}, u) \models LM\varphi$. \blacksquare

Our last auxiliary lemma is as follows.

Lemma 29 *Let S be a modal logic characterized by a class of cluster-decomposable Kripke interpretations and let A be a set of formulas. Let $\mathfrak{M} = \langle U, I_{S,A}|_U \rangle$ be a terminal cluster of $\mathfrak{M}_{S,A}$ and let $u \in U_{S,A}$ and $v \in U$ be such that $(u, v) \in R_{S,A}$. Then $\mathfrak{M}_{S,A}^u \models \{M\varphi : \mathfrak{M} \models M\varphi\}$.¹²*

Proof Let $\mathfrak{M} \models M\varphi$. Since \mathfrak{M} is a terminal cluster of $\mathfrak{M}_{S,A}$, it is also a terminal cluster of $\mathfrak{M}_{S,A}^u$ and, in addition,

$$(\mathfrak{M}_{S,A}^u)^v = \mathfrak{M}_{S,A}^v = \mathfrak{M},$$

because $v \in U$. Therefore, by Proposition 2, $(\mathfrak{M}_{S,A}^u, v) \models M\varphi$, which, by (Hughes & Cresswell 1996, Modalities in S5, pp. 59–60) implies

$$(\mathfrak{M}_{S,A}^u, v) \models LM(LM)^n\varphi, \quad n = 0, 1, \dots, \quad (9)$$

where $(LM)^0\varphi$ is φ and $(LM)^{n+1}\varphi$ is $LM(LM)^n\varphi$.

Let $w \in U^u$ and let i be the minimal integer for which $(u, w) \in R_{S,A}^i$. We shall prove by induction on i that for all $n = 0, 1, \dots$, $(\mathfrak{M}_{S,A}^u, w) \models M(LM)^n\varphi$. Then the lemma will follow with $n = 0$.

¹²In fact, $\mathfrak{M}_{S,A}^u$ is cluster-decomposable and \mathfrak{M} is the final cluster of $\mathfrak{M}_{S,A}^u$, but we do not need this for the proof of Theorem 24.

Basis: $i = 0$, i.e., $w = u$. Then the containment $(u, v) \in R_{S,A}$ and (9) imply

$$(\mathfrak{M}_{S,A}^u, u) \models MLM(LM)^n \varphi. \quad n = 0, 1, \dots$$

Then, by **bm**, $(\mathfrak{M}_{S,A}^u, u) \models M(LM)^n \varphi$.

Induction step: $(u, w) \in R_{S,A}^{i+1}$. Then there is a world w_i such that $(u, w_i) \in R_{S,A}^i$ and $(w_i, w) \in R_{S,A}$. By the induction hypothesis,

$$(\mathfrak{M}_{S,A}^u, w_i) \models MLM(ML)^n \varphi, \quad n = 0, 1, \dots$$

Then, by **5m**, $(\mathfrak{M}_{S,A}^u, w_i) \models LM(LM)^n \varphi$, which together with $(w_i, w) \in R_{S,A}$, implies $(\mathfrak{M}_{S,A}^u, w) \models M(LM)^n \varphi$. This completes the induction step and the proof of the lemma. ■

Proof of the “only if” part of Theorem 24 Let E be an S-expansion for A . By Proposition 15, there exists a cluster $\mathfrak{M}' = \langle U', I' \rangle$ such that $E = \mathbf{Th}(\mathfrak{M}')$. We may assume that \mathfrak{M}' is maximal, see the construction in the end of Section 2.1.

Let $U = \{\mathbf{Th}(\mathfrak{M}', u') : u' \in U'\}$. Then, by Example 6, $U \subseteq U_{S,A}$ and, by Remark 10, $U \times U \subseteq R_{S,A}$. That is, $\mathfrak{M} = \langle U, I_{S,A}|_U \rangle$ is a cluster. By Theorem 9 and the definition of \models , for each $u' \in U'$,

$$I_{S,A}(\mathbf{Th}(\mathfrak{M}', u')) = I'(u').$$

Hence, by Lemma 12,

$$\mathbf{Th}(\mathfrak{M}) = \mathbf{Th}(\mathfrak{M}') = E.$$

Therefore, the proof will be complete if we show that \mathfrak{M} is an isolated cluster of $\mathfrak{M}_{S,A}$. Since, by Lemma 26, \mathfrak{M} is a terminal cluster of $\mathfrak{M}_{S,A}$, it suffices to prove that for each $u \in U_{S,A}$ and each $v \in U$, $(u, v) \in R_{S,A}$ implies $u \in U$. By the definition of U , the containment $u \in U$ will follow, if we show that for some $u' \in U'$, $u = \mathbf{Th}(\mathfrak{M}', u')$. The proof is similar to that of Lemma 26.

First, we observe that

$$\mathbf{Th}(\mathfrak{M}') \subseteq u. \quad (10)$$

Indeed, let $\theta \in \mathbf{Th}(\mathfrak{M}')$. Since $\mathbf{Th}(\mathfrak{M}')$ is an S-expansion for A and $\mathbf{Th}(\mathfrak{M}) = \mathbf{Th}(\mathfrak{M}')$, by Proposition 16, $A \cup \{M\varphi : \mathfrak{M} \models \varphi\} \models \theta$. Hence, by Lemma 29, $\mathfrak{M}_{S,A}^u \models \theta$, and, by Proposition 2, $\mathfrak{M}_{S,A} \models \theta$ as well. Therefore, $\theta \in u$ follows from Theorem 9.

Therefore,

$$\mathbf{Th}(\mathfrak{M}') \cap \mathbf{GFm} \subseteq u \cap \mathbf{GFm} = \mathbf{Th}(I_{S,A}(u)).^{13} \quad (11)$$

Since \mathfrak{M}' is maximal, it follows from (11) that there is a world $u' \in U'$ such that

$$I'(u') = I_{S,A}(u), \quad (12)$$

and the proof will be complete if we show that $\mathbf{Th}(\mathfrak{M}', u') = u$. That is, for each formula φ , $(\mathfrak{M}', u') \models$

¹³Recall that by $\mathbf{Th}(I_{S,A}(v))$ we mean the *ground* theory of $I_{S,A}(v)$.

φ if and only if $\varphi \in u$. The proof is by induction on the complexity of φ .

Basis: The case of \perp is trivial and the case of a propositional variable immediately follows from (12) and the definition of $I_{S,A}$ (Definition 8).

Induction step: The case of implication is immediate and is omitted. Let φ be of the form $L\psi$ and assume $(\mathfrak{M}', u') \models L\psi$. Then $\mathfrak{M}' \models L\psi$, because \mathfrak{M}' is a cluster, and $L\psi \in u$ follows from (10).

Conversely, let $L\psi \in u$. Then, by Theorem 9, $(\mathfrak{M}_{S,A}, u) \models L\psi$. Assume to the contrary that $L\psi \notin \mathbf{Th}(\mathfrak{M}', u')$. Then, by Proposition 7, $M\neg\psi \in \mathbf{Th}(\mathfrak{M}', u')$. Since \mathfrak{M}' is a cluster, $\mathfrak{M}' \models M\neg\psi$. Then, by (10) and Theorem 9, $(\mathfrak{M}_{S,A}, u) \models M\neg\psi$, in contradiction with $(\mathfrak{M}_{S,A}, u) \models L\psi$. This completes the induction step and the proof of the “only if” part of the theorem. ■

4 Remarks on the proof of Theorem 24

Note that the assumption that S is characterized by a cluster-decomposable class of Kripke interpretations was not used in the proof of the “if” part of Theorem 24 at all, and in the proof of the “only if” part of the theorem we used only the schemes **bm** and **5m** (which, by Lemma 28, belong to S). This naturally leads to the following questions.

- Is $K + \mathbf{bm} + \mathbf{5m}$ characterized by a class of cluster-decomposable Kripke interpretations (cf. (Tiomkin & Kaminski 2007))?
- What is the weakest modal logic S such that for a set of axioms A , the S-expansions for A are exactly the theories of the isolated clusters of $\mathfrak{M}_{S,A}$?

A negative answer to the first question would strengthen Theorem 24 by replacing its prerequisite that S is characterized by a class of cluster-decomposable Kripke interpretations with **bm**, **5m** \in S, and the second question (for which we have no answer) is of interest for an obvious reason.

A trivial example below shows that indeed, $K + \mathbf{bm} + \mathbf{5m}$ is not characterized by a class of cluster-decomposable Kripke interpretations.

Example 30 Obviously, $K + \mathbf{bm} + \mathbf{5m} + L\perp$ is consistent, because it is satisfied by any Kripke interpretation with the empty accessibility relation.¹⁴ However, no cluster-decomposable Kripke interpretation satisfies $K + \mathbf{bm} + \mathbf{5m} + L\perp$, because all cluster-decomposable Kripke interpretations satisfy \mathbf{d} which is the negation of $L\perp$.

Example 30 does not provide any hint on a possible strengthening of Theorem 24, because all expansions contain \mathbf{d} . Also, by Theorem 31 below, *non-monotonic* $K + \mathbf{bm} + \mathbf{5m}$ and $K + \mathbf{bm} + \mathbf{5m} + \mathbf{d}$ are equivalent.

¹⁴Theories containing $L\perp$ are trivial: in such theories all formulas of the form $L\varphi$ are true, and all formulas of the form $M\varphi$ are false. Theories containing $L\perp$ are called *Ver*, see (Hughes & Cresswell 1996, pp. 66 and 108).

Theorem 31 Let S be a modal logic and let A be a set of formulas. A set of formulas E is an S -expansion for A if and only if E is an $(S + \mathbf{d})$ -expansion for A .

Proof The “only if” part of the theorem is a particular case of (Marek & Truszczyński 1993, Theorem 9.6, p. 256), and for the proof of the “if” part we proceed as follows.

Let E be an $(S + \mathbf{d})$ -expansion for A , i.e.,

$$E = \mathbf{Th}_{S+\mathbf{d}}(A \cup \{M\varphi : E \not\vdash_{S+\mathbf{d}} \neg\varphi\}). \quad (13)$$

We have to show that

$$E = \mathbf{Th}_S(A \cup \{M\varphi : E \not\vdash_S \neg\varphi\}).$$

It follows from (13) that E contains \mathbf{d} , which implies

$$\{M\varphi : E \not\vdash_{S+\mathbf{d}} \neg\varphi\} = \{M\varphi : E \not\vdash_S \neg\varphi\}. \quad (14)$$

In addition, since E is $(S + \mathbf{d})$ -consistent, $E \not\vdash_S \neg\top$, implying

$$\mathbf{d} \in \{M\varphi : E \not\vdash_S \neg\varphi\}. \quad (15)$$

Therefore,

$$\begin{aligned} E &= \mathbf{Th}_{S+\mathbf{d}}(A \cup \{M\varphi : E \not\vdash_{S+\mathbf{d}} \neg\varphi\}) \\ &= \mathbf{Th}_{S+\mathbf{d}}(A \cup \{M\varphi : E \not\vdash_S \neg\varphi\}) \\ &= \mathbf{Th}_S(A \cup \{M\varphi : E \not\vdash_S \neg\varphi\}), \end{aligned}$$

where the second equality follows from (14) and the last equality follows from (15). \blacksquare

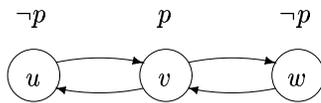
Consequently, to eliminate the degenerate case of Example 30, the first question should be restated as follows.

- Is $\mathbf{K} + \mathbf{bm} + \mathbf{5m} + \mathbf{d}$ characterized by a class of cluster-decomposable Kripke interpretations (cf. (Tiomkin & Kaminski 2007))?

This seems to be a difficult question to which we have no clue.

We conclude this section with two examples which show that \mathbf{bm} and $\mathbf{5m}$ are independent in $\mathbf{K} + \mathbf{d}$.¹⁵

Example 32 Consider a Kripke interpretation \mathfrak{M} depicted below.



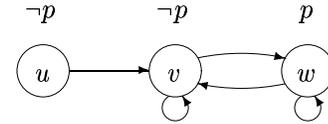
That is, $\mathfrak{M} = \langle U, R, I \rangle$, where

- $U = \{u, v, w\}$,
- $R = \{(u, v), (v, u), (v, w), (w, v)\}$,
- $I(u) = I(w) = \emptyset$, and $I(v) = \{p\}$.

Since R has no dead ends, $\mathfrak{M} \models \mathbf{d}$, and, since R is symmetric, $\mathfrak{M} \models \mathbf{b}$. Therefore, by Remark 27, $\mathfrak{M} \models \mathbf{bm}$. In addition, $(\mathfrak{M}, u) \models \mathbf{MLMp}$, because $(\mathfrak{M}, v) \models \mathbf{LMp}$. However, $(\mathfrak{M}, u) \not\models \mathbf{LMp}$. Hence, $\mathfrak{M} \not\models \mathbf{5m}$.

¹⁵Semantical characterizations of \mathbf{bm} and $\mathbf{5m}$ can be found in ((in collaboration with Dana Scott) 1977, p. 67).

Example 33 Consider a Kripke interpretation \mathfrak{M} depicted below.



That is, $\mathfrak{M} = \langle U, R, I \rangle$, where

- $U = \{u, v, w\}$,
- $R = \{(u, v)\} \cup \{v, w\}^2$,
- $I(u) = I(v) = \emptyset$, and $I(w) = \{p\}$.

Note that $\langle \{v, w\}, I|_{\{v, w\}} \rangle$ is the only terminal cluster of \mathfrak{M} . However, $\langle \{v, w\}, I|_{\{v, w\}} \rangle$ is not a final cluster, because $(u, w) \notin R$ (cf. the note following Definition 25).

Since R has no dead ends, $\mathfrak{M} \models \mathbf{d}$, and, since, obviously, R is euclidean, $\mathfrak{M} \models \mathbf{5}$. Therefore, by Remark 27, $\mathfrak{M} \models \mathbf{5m}$. In addition, $(\mathfrak{M}, v) \models \mathbf{LMp}$, implying $(\mathfrak{M}, u) \models \mathbf{MLMp}$. However, $(\mathfrak{M}, u) \not\models \mathbf{Mp}$. Hence, $\mathfrak{M} \not\models \mathbf{bm}$.

5 Autoepistemic logic: knowledge vs. belief

In this section we apply Theorem 24 to *autoepistemic* logic ((Moore 1987)) to show that autoepistemic expansions separate knowledge from belief in $\mathbf{KD45}$ -canonical models. First we recall some basic facts about the *logic of belief* $\mathbf{KD45}$ and its relationship to autoepistemic logic.

The modal logic $\mathbf{KD45}$ results in adding to \mathbf{K} the schemes \mathbf{d} , $\mathbf{4}$, and $\mathbf{5}$. It is well-known that, for a set of proper axioms A , the connected components of the canonical $\mathbf{KD45}, A$ -model $\mathfrak{M}_{\mathbf{KD45}, A}$ are of the form

$$\mathfrak{M} = \langle U, U \times U_c, I \rangle, \quad (16)$$

where U_c is a non-empty subset of U . That is, $\langle U_c, I|_{U_c} \rangle$ is the final cluster of \mathfrak{M} and each world in $U \setminus U_c$ can see the whole cluster $\langle U_c, I|_{U_c} \rangle$, but nothing more. A Kripke interpretation \mathfrak{M} of the form (16) will be called a *KD45-model*.

The distinction between *knowledge* and *belief* in a Kripke interpretation \mathfrak{M} is as follows. A formula φ is *known* in \mathfrak{M} , if $\mathfrak{M} \models \varphi$, whereas φ is *believed* in \mathfrak{M} , if $\mathfrak{M} \models L\varphi$.

Since in clusters satisfiability of $L\varphi$ implies satisfiability of φ , in the theory of a cluster the modal connective L should be *naturally* interpreted as *known*, see (Voorbraak 1990; Schwarz 1992).¹⁶ This is in contrast with the theory of a $\mathbf{KD45}$ -model, because in $\mathbf{KD45}$ -models belief does not necessarily imply knowledge. In particular, an agent may believe something that does not hold in its own world. Consequently, in the theory of a proper (non-cluster) $\mathbf{KD45}$ -model L should be (also

¹⁶Cf. (Stalaker 1993; Moore 1985; 1987; Konolige 1988), where L , for no reason, is interpreted as *believed*.

naturally) interpreted as *believed*, cf. (Voorbraak 1990). Namely, an agent in each world of an KD45-model believes in what is satisfied by the model's final cluster. In particular, the belief set of an agent is the same in all worlds.

We shall use an equivalent definition of autoepistemic expansions given by (Shvarts 1990, Proposition 2.1) (see also (Marek & Truszczyński 1993, Corollary 10.44, p. 313)) according to which autoepistemic expansions coincide with KD45-expansions. Note that even though, autoepistemic expansions are based on beliefs of an agent, they are theories of clusters which correspond to knowledge.

Now, applying the above discussion to canonical KD45-models, we see that its isolated clusters reflect knowledge, whereas the other connected components, which are *proper* KD45 models, reflect belief. Thus, by Theorem 24, autoepistemic expansions for a set of formulas A coincide with the theories of the "knowledge" components of $\mathfrak{M}_{\text{KD45},A}$.

Acknowledgments

This research was supported by the Technion president fund and by the fund for the promotion of research at the Technion.

References

- Donini, F.; Nardi, D.; and Rosati, R. 1997. Ground nonmonotonic modal logics. *Journal of Logic and Computation* 7:523–548.
- Hughes, G., and Cresswell, M. 1984. *A companion to modal logic*. London: Methuen and Co.
- Hughes, G., and Cresswell, M. 1996. *A new introduction to modal logic*. New York, New York: Routledge.
- (in collaboration with Dana Scott), E. L. 1977. *The "Lemmon notes": an introduction to modal logic*. Oxford: Basil Blackwell. Edited by Krister Segerberg.
- Konolige, K. 1988. On the relation between default and autoepistemic logic. *Artificial Intelligence* 35:343–382.
- Marek, W., and Truszczyński, M. 1993. *Nonmonotonic Logic*. Berlin: Springer-Verlag.
- McDermott, D., and Doyle, J. 1980. Non-monotonic logic I. *Artificial Intelligence* 13:41–72.
- McDermott, D. 1982. Non-monotonic logic II: Non-monotonic modal theories. *Journal of the ACM* 29:33–57.
- Moore, R. 1985. Semantical considerations on non-monotonic logics. *Artificial Intelligence* 25:75–94.
- Moore, R. 1987. Possible-world semantics for autoepistemic logic. In Ginsberg, M., ed., *Readings in Nonmonotonic Reasoning*, 137–142. Los Altos, California: Morgan Kaufmann.
- Schwarz, G. 1992. Minimal model semantics for non-monotonic modal logics. In *Proceedings of the Seventh*

Annual IEEE Symposium on Logic in Computer Science, 34–43. Los Alamitos, California: IEEE Computer Society Press.

Shvarts, G. 1990. Autoepistemic modal logics. In Parikh, R., ed., *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge*, 97–110. Los Altos, California: Morgan Kaufmann.

Stalnaker, R. 1993. A note on nonmonotonic modal logic. *Artificial Intelligence* 64:183–196.

Tiomkin, M., and Kaminski, M. 2007. The modal logic of cluster-decomposable kripke interpretations. *Notre Dame Journal of Formal Logic* 48:511–520.

Voorbraak, F. 1990. The logic of objective knowledge and rational belief. In van Eijck, J., ed., *Proceedings of the European Workshop on Logics in Artificial Intelligence – JELIA '90*, volume 478 of *Lecture Notes in Artificial Intelligence*, 499–515. Berlin: Springer.