# Sabotage Modal Logic: Some Model and Proof Theoretic Aspects

Guillaume Aucher, Johan van Benthem, Davide Grossi

## ▶ To cite this version:

# Sabotage Modal Logic:
# Some Model and Proof Theoretic Aspects

Guillaume Aucher[1], Johan van Benthem[2], and Davide Grossi[3]

[1] University of Rennes 1 – INRIA,
guillaume.aucher@irisa.fr
[2] University of Amsterdam, Stanford University, Tsinghua University,
J.vanBenthem@uva.nl
[3] University of Liverpool,
D.Grossi@liverpool.ac.uk

**Abstract.** We investigate some model and proof theoretic aspects of sabotage modal logic. The first contribution is to prove a characterization theorem for sabotage modal logic as the fragment of first-order logic which is invariant with respect to a suitably defined notion of bisimulation (called sabotage bisimulation). The second contribution is to provide a sound and complete tableau method for sabotage modal logic. We also chart a number of open research questions concerning sabotage modal logic, aiming at integrating it within the current landscape of logics of model update.

## 1 Introduction

Sabotage modal logic (SML) [13] expands the standard modal language with an edge-deletion modality $\blacklozenge\varphi$ whose intended reading is "after the deletion of at least one edge in the frame it holds that $\varphi$". As such it can be viewed as the modal logic of arbitrary edge deletion. Although it inspired several later formalisms in the dynamic epistemic logic tradition [14] (e.g., *graph modifiers logic* [3], *memory logic* [1], *swap logic* [2], *arrow update logic* [8]), and is directly related to recent work in theoretical computer science (e.g., [11,7]) and learning theory (e.g., [6]) it remains a rather under-investigated logic.

The only work focusing specifically on SML is, to the best of our knowledge, [10,9] where the undecidability of the satisfiability problem and the complexity of the model-checking problem of SML are established. Among the open questions concerning SML, that work points to the lack of a notion of bisimulation characteristic for SML. The present article addresses such question and can be regarded as an application of standard techniques and methods of modal correspondence theory [12] to sabotage modal logic. The article provides as well a sound and complete tableau method for SML. This contributes to the proof theory of SML, which has rather been neglected so far. In pursuing our investigations, the article establishes a few related model-theoretic results and aims at putting SML 'on the map' of current research in dynamic epistemic logic.

*Outline of the article.* Section 2 introduces SML and what is thus far known of its properties. That is the starting point of the article. Section 3 introduces a notion of bisimulation for SML—called sabotage bisimulation—and Section 4 characterizes SML as the fragment of first-order logic (FOL) which is invariant for sabotage bisimulation. Section 5 provides a sound and complete tableau method for sabotage modal logic. Section 6 concludes with some open research questions.

## 2   Preliminaries

In this section, we introduce the syntax and semantics of SML, recapitulate some key results from [9], and present a standard translation for SML (to FOL).

### 2.1   Syntax

Let $\mathbf{P}$ be a countable set of propositional atoms. The set of formulae of the sabotage modal language $\mathcal{L}^s$ is defined by the following grammar in BNF:

$$\mathcal{L}^s : \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Diamond\varphi \mid \blacklozenge\varphi$$

where $p \in \mathbf{P}$. The remaining set of Boolean connectives $\{\vee, \rightarrow\}$ and the modal operators $\Box$ and $\blacksquare$ can be defined in the standard way. The formula $\bot$ is an abbreviation for the formula $p \wedge \neg p$ (for a chosen $p \in \mathbf{P}$) and $\top$ is an abbreviation for $\neg\bot$. The iteration of $n$ sabotage operators or modalities will sometimes be denoted by $\blacklozenge_n$ and $\Diamond_n$, respectively, $\blacksquare_n$ and $\Box_n$. To save parenthesis, we use the following ranking of binding strength: $\blacksquare, \blacklozenge, \Box, \Diamond, \neg, \wedge, \vee, \rightarrow$.

A natural measure of syntactic complexity for sabotage formulae is given by their *sabotage depth* [9]. Let $\varphi \in \mathcal{L}^s$. The *sabotage depth* of $\varphi$, written $sd(\varphi)$, is inductively defined as follows: $sd(\top) = sd(p) := 0$, $sd(\neg\varphi) := sd(\varphi)$, $sd(\varphi_1 \wedge \varphi_2) := \mathtt{max}\{sd(\varphi_1), sd(\varphi_2)\}$, $sd(\Diamond\varphi) := sd(\varphi)$ and $sd(\blacklozenge\varphi) := sd(\varphi) + 1$.

### 2.2   Semantics

We will be working with standard Kripke models $\mathcal{M} = (W, R, V)$ where: $W$ is a non-empty set; $R \subseteq W \times W$; and $V : \mathbf{P} \longrightarrow 2^W$. The pair $(W, R)$ is called a frame, and is denoted by $\mathcal{F}$.

Such structures will be also interpreted as models for the binary fragment of FOL with equality[1] denoted $\mathcal{L}^1$. Sometimes we will use the following FOL terminology/notation. We say that a model $\mathcal{M}$ *satisfies* a formula $\varphi(x) \in \mathcal{L}^1$ (or a set $\Gamma(x) \subseteq \mathcal{L}^1$) with one free variable $x$ under the assignment of $w$ to $x$ if and only if $\varphi$ (respectively $\Gamma$) is true of $w$, in symbols, $\mathcal{M} \models \varphi(x)[w]$ (respectively, $\mathcal{M} \models \Gamma(x)[w]$). We say that a model $\mathcal{M}$ *realizes* a set $\Gamma(x) \subseteq \mathcal{L}^1$ with one free variable $x$ (i.e., a type) if and only if there exists an element $w \in W$ such that $\mathcal{M} \models \Gamma(x)[w]$.

---

[1] We refer the reader to [4, Ch. 2.4].

The satisfaction relation for $\mathcal{L}^s$ is defined as usual for the atomic and Boolean cases, and for the standard modalities. For the sabotage modality it is as follows:

$$(W, R, V), w \models \blacklozenge\varphi \Longleftrightarrow \exists(w', w'') \in R \text{ s.t. } (W, R \setminus \{(w', w'')\}, V), w \models \varphi \quad (1)$$

In other words, $\blacklozenge\varphi$ is satisfied by a pointed model if and only if there exist two $R$-related (possibly identical) states such that, once the edge between these two states is removed from $R$, $\varphi$ holds at the same evaluation state. The notions of validity and logical consequence are defined as usual.

We say that two pointed models $(\mathcal{M}, w)$ and $(\mathcal{M}', w')$ are *sabotage-related* (notation, $(\mathcal{M}, w) \xrightarrow{\blacklozenge} (\mathcal{M}', w')$) if and only if: $w' = w$; $W' = W$; $R' = R \setminus \{(w'', w''')\}$ for some $w'', w''' \in W$; $V' = V$. The set $\mathbf{r}(\mathcal{M}, w) = \{(\mathcal{M}', w') \mid (\mathcal{M}, w) \xrightarrow{\blacklozenge} (\mathcal{M}', w')\}$ denotes the set of all models which are sabotage-related to a given pointed model $(\mathcal{M}, w)$. Similarly, $\mathbf{r}^n(\mathcal{M}, w) = \{(\mathcal{M}', w') \mid (\mathcal{M}_1, w_1) \xrightarrow{\blacklozenge} (\mathcal{M}_2, w_2) \xrightarrow{\blacklozenge} \ldots \xrightarrow{\blacklozenge} (\mathcal{M}_{n+1}, w_{n+1}) \text{ \& } (\mathcal{M}_1, w_1) = (\mathcal{M}, w) \text{ \& } (\mathcal{M}_{n+1}, w_{n+1}) = (\mathcal{M}', w')\}$ denotes the set of all models which are related to $(\mathcal{M}, w)$ by a $\xrightarrow{\blacklozenge}$-path of length $n$. Finally, $\mathbf{r}^*(\mathcal{M}, w) = \{(\mathcal{M}', w') \mid (\mathcal{M}, w) \xrightarrow{\blacklozenge}^* (\mathcal{M}', w')\}$ denotes the set of all pointed models which are reachable from $(\mathcal{M}, w)$ by the reflexive and transitive closure of $\xrightarrow{\blacklozenge}$. We will often drop the reference to a given point in the model, which will be clear by the context, and simply write $\mathcal{M} \xrightarrow{\blacklozenge} \mathcal{M}'$ instead of $(\mathcal{M}, w) \xrightarrow{\blacklozenge} (\mathcal{M}', w')$.

The set of sabotage modal formulae which are satisfied by a pointed model $(\mathcal{M}, w)$, i.e., the *sabotage modal logic theory* of $w$ in $\mathcal{M}$, is denoted $\mathbb{T}^s(\mathcal{M}, w)$. We say that two pointed models $(\mathcal{M}, w)$ and $(\mathcal{M}', w')$ are *sabotage modally equivalent* —notation: $(\mathcal{M}, w) \leftrightsquigarrow_s (\mathcal{M}', w')$—if and only if they satisfy the same sabotage modal formulae, that is, they have the same sabotage modal logic theory.

SML can express properties that are beyond the reach of standard modal logic. An example is the property "there are at most $n$ successors" with $1 \leq n$ (see also Example 1 below):

$$\exists^{\leq n} y \, (xRy). \quad (2)$$

This property can be expressed in SML with:

$$\Box\bot \vee \bigvee_{1 \leq i \leq n} \blacklozenge_i \Box\bot. \quad (3)$$

A dual formula expresses the property "there are at least $n$ successors". In fact, SML can even define frames up to isomorphisms. Indeed, one can easily show that the formula $\Diamond\top \wedge \Box\Diamond\top \wedge \blacksquare\Box\bot$ is true in a model if and only if its underlying frame consists of one reflexive point.

Finally, let us recapitulate the findings of [10,9]. They are proved with a multi-modal version of SML, but all our results and methods are easily generalizable to this multi-modal setting.

**Theorem 1 ([10,9]).** *The model-checking problem of* SML *is PSPACE-complete.* SML *lacks the finite model property, the tree-model property and its satisfiability problem is undecidable.*

### 2.3   A Standard Translation for SML

A standard translation for SML was first sketched in the technical report [10]. In this section we describe such translation and its correctness in detail. This is essential to prepare the later sections of the article.

**Setting up the translation** In order to define a standard translation from the language of SML to the free variable fragment of FOL with equality one needs to keep track of the changes that the sabotage operators introduce in the model.

   This can be achieved by indexing the standard translation with a set $E$ consisting of pairs of variables. The idea is that when the standard translation is applied to the outermost operator of a given formula, this set is empty. As the analysis proceeds towards inner operators, each sabotage operator ♦ in the formula will introduce a new pair of variables in $E$, which will be bound by an existential quantifier. Here is the formal definition:

**Definition 1 (Standard translation for SML).** *Let $E$ be a set of pairs $(y, z)$ of variables—edges—and $x$ be a designated variable. The translation $ST_x^E$ : $\mathcal{L}^s \longrightarrow \mathcal{L}^1$ is recursively defined as follows:*

$$ST_x^E(p) = P(x)$$
$$ST_x^E(\bot) = x \neq x$$
$$ST_x^E(\neg\varphi) = \neg ST_x^E(\varphi)$$
$$ST_x^E(\varphi_1 \wedge \varphi_2) = ST_x^E(\varphi_1) \wedge ST_x^E(\varphi_2)$$
$$ST_x^E(\Diamond\varphi) = \exists y \left( xRy \wedge \bigwedge_{(v,w)\in E} \neg(x = v \wedge y = w) \wedge ST_y^E(\varphi) \right)$$
$$ST_x^E(\blacklozenge\varphi) = \exists y, z \left( yRz \wedge \bigwedge_{(v,w)\in E} \neg(y = v \wedge z = w) \wedge ST_x^{E\cup\{(y,z)\}}(\varphi) \right)$$

The key clauses concern $\Diamond$-formulae and ♦-formulae. Let us start with the latter. Formula ♦$\varphi$ is translated as the first order formula stating the following: that there exists some $R$-edge denoted by $(y, z)$; that such edge is different from any edge possibly denoted by the pairs in $E$; that the translation of $\varphi$ should now be carried out with respect to the set $E \cup \{(y, z)\}$; and that this translation is realized at $x$.

   As to the former clause, it says that formula $\Diamond\varphi$ is translated as the first order formula with $x$ free, which states the existence of a state $y$ accessible from $x$ via an edge which is different from all the edges in the set $E$, and that the translation of $\varphi$ is realized at $y$.

   Setting up the translation like this allows one to book-keep the removal of edges via $E$. The removal of edges is handled by imposing the existence of states which are different from the ones reachable via the 'removed' edges. In other words edge removal is simulated by imposing the existence of edges which are then not used to interpret inner modal operators.

It is important to notice the following feature of the translation. Depending on the chosen $E$, $ST^E$ can possibly yield formulae with several free variables, e.g.: $ST_x^{(v,w)}\Diamond p = \exists y\,(xRy \land \neg(x=v \land y=w) \land p)$. However, if $ST^E$ is applied to a formula $\varphi$ by setting $E = \emptyset$, that is to say, if the translation is initiated with an empty $E$, then, at each successive application of $ST^E$ to subformulae of $\varphi$, the variables occurring in $W$ will be bound by some quantifiers introduced at previous steps. For any $\varphi$, $ST_x^\emptyset(\varphi)$ yields a FOL formula with only $x$ free.

**Correctness of the translation** We prove now the correctness of the translation proposed in Definition 1.

**Theorem 2.** *Let $\mathcal{M}, w$ be a pointed model and $\varphi \in \mathcal{L}^s$:*

$$\mathcal{M}, w \models \varphi \Longleftrightarrow \mathcal{M} \models ST_x^\emptyset(\varphi)[w]$$

*Proof (Sketch).* By induction on the structure of $\varphi$. We omit the Boolean and modal cases. The case for the sabotage operator $\blacklozenge$ is proven by the following series of equivalences:

$$\mathcal{M}, w \models \blacklozenge\varphi \Longleftrightarrow \mathcal{M}, w \xrightarrow{\blacklozenge} \mathcal{M}', w \models \varphi \qquad\qquad \text{semantics of } \blacklozenge \; (1)$$

$$\Longleftrightarrow \mathcal{M}, w \xrightarrow{\blacklozenge} \mathcal{M}' \models ST_x^\emptyset(\varphi)[w] \qquad\qquad\qquad \text{IH}$$

$$\Longleftrightarrow \mathcal{M} \models \exists y, z \left( yRz \land ST_x^{\{(y,z)\}}(\varphi)[w] \right) \quad \text{sem. of } \blacklozenge \; (1) \text{ and Def. 1}$$

$$\Longleftrightarrow \mathcal{M} \models ST_x^\emptyset(\blacklozenge\varphi)[w] \qquad\qquad\qquad\qquad\qquad \text{Def. 1} \quad \square$$

We conclude the section with the following observation:

**Proposition 1.** SML *is not contained in any fixed variable fragment of* FOL.

*Proof.* We show SML contains formulae that are not definable in any fixed variable fragment of FOL. Consider the above FOL formulae with counting quantifier of Expression (2) with $1 \leq n$. Clearly, for each integer $n$, Expression (2) is definable in FOL (without counting quantifiers) using a fixed number of variables. But no fixed variable fragment can define Expression (2) for all integers $n$. Since Expression (2) is equivalent to (3) it follows that although SML is FOL-definable (Corollary 2) it is not definable in any fixed variable fragment of FOL. $\square$

## 3 Bisimulation for SML

In this section, we introduce a notion of bisimulation for SML.

### 3.1 Sabotage Bisimulation

**Definition 2 (s-bisimulation).** *Let $\mathcal{M}_1 = (W_1, R_1, V_1)$ and $\mathcal{M}_2 = (W_2, R_2, V_2)$ be two Kripke models. A non-empty relation $Z \subseteq \mathbf{r}^*(\mathcal{M}_1, w) \times \mathbf{r}^*(\mathcal{M}_2, v)$ is an* s-bisimulation *between the two pointed models $(\mathcal{M}_1, w)$ and $(\mathcal{M}_2, v)$—notation, $Z : (\mathcal{M}_1, w) \leftrightarroweq_s (\mathcal{M}_2, v)$—if the following conditions are satisfied:*

**Atom:** *If $(\mathcal{M}_1, w)Z(\mathcal{M}_2, v)$ then $\mathcal{M}_1, w \models p$ iff $\mathcal{M}_2, v \models p$, for any atom $p$.*

**Zig$_\Diamond$:** *If $(\mathcal{M}_1, w)Z(\mathcal{M}_2, v)$ and there exists $w' \in W_1$ s.t. $wR_1w'$ then there exists $v' \in W_2$ s.t. $vR_2v'$ and $(\mathcal{M}_1, w')Z(\mathcal{M}_2, v')$;*

**Zag$_\Diamond$:** *If $(\mathcal{M}_1, w)Z(\mathcal{M}_2, v)$ and there exists $v' \in S_s$ s.t. $vR_1v'$ then there exists $w' \in W_1$ s.t. $wR_1w'$ and $(\mathcal{M}_1, w')Z(\mathcal{M}_2, v')$;*

**Zig$_\blacklozenge$:** *If $(\mathcal{M}_1, w)Z(\mathcal{M}_2, v)$ and there exists $\mathcal{M}_1'$ such that $(\mathcal{M}_1, w) \xrightarrow{\blacklozenge} (\mathcal{M}_1', w)$, then there exists $\mathcal{M}_2'$ such that $(\mathcal{M}_2, v) \xrightarrow{\blacklozenge} (\mathcal{M}_2', v)$ and $(\mathcal{M}_1', w)Z(\mathcal{M}_2', v)$;*

**Zag$_\blacklozenge$:** *If $(\mathcal{M}_1, w)Z(\mathcal{M}_2, v)$ and there exists $\mathcal{M}_2'$ such that $(\mathcal{M}_2, v) \xrightarrow{\blacklozenge} (\mathcal{M}_2', v)$, then there exists $\mathcal{M}_1'$ such that $(\mathcal{M}_1, w) \xrightarrow{\blacklozenge} (\mathcal{M}_1', w)$ and $(\mathcal{M}_1', w)Z(\mathcal{M}_2', v)$.*

*We write $(\mathcal{M}_1, w) \leftrightarroweq_s (\mathcal{M}_2, v)$ if there exists an s-bisimulation $Z$ s.t. $(\mathcal{M}_1, w)Z(\mathcal{M}_2, v)$.*

It is worth spending a few words about Definition 2. The notion of s-bisimulation strengthens the standard modal bisimulation with the 'zig' and 'zag' conditions for the sabotage modality. Just like the sabotage modality is an 'external' modality so is s-bisimulation an 'external' notion of bisimulation. Standard bisimulation keeps the model fixed and changes the evaluation point along the accessibility relation of the Kripke model, s-bisimulation keeps the evaluation point fixed and changes the model by picking one among the sabotage-accessible ones.

### 3.2   Bisimulation and Modal Equivalence in SML

We first show that s-bisimulation implies SML equivalence.

**Proposition 2 ($\leftrightarroweq_s \subseteq \leftrightsquigarrow_s$).** *For any two pointed models $(\mathcal{M}_1, w)$ and $(\mathcal{M}_2, v)$ it holds that: $(\mathcal{M}_1, w) \leftrightarroweq_s (\mathcal{M}_2, v) \Longrightarrow (\mathcal{M}_1, w) \leftrightsquigarrow_s (\mathcal{M}_2, v)$.*

*Proof.* The proof is by induction on the syntax of $\varphi$. Assume $(\mathcal{M}_1, w_1)Z(\mathcal{M}_2, w_2)$. **Base:** The Atom clause of Definition 2 covers the case of atoms and nullary operators. **Step:** The Boolean cases are as usual. The Zig$_\Diamond$ and Zag$_\Diamond$ clauses of Definition 2 take care of $\Diamond$-formulae in the standard way. As to $\blacklozenge$-formulae, assume $\mathcal{M}_1, w_1 \models \blacklozenge\varphi$. By the semantics of $\blacklozenge$ we have that $\mathcal{M}_1 \xrightarrow{\blacklozenge} \mathcal{M}_1', w \models \varphi$ and, by clause Zig$_\blacklozenge$ of Definition 2, it follows that $\mathcal{M}_2 \xrightarrow{\blacklozenge} \mathcal{M}_2'$ and $(\mathcal{M}_1', w)Z(\mathcal{M}_2', v)$. By IH we conclude that $\mathcal{M}_2', v \models \varphi$ and, consequently, $\mathcal{M}_2, v \models \blacklozenge\varphi$. Similarly, from $\mathcal{M}_2, v \models \blacklozenge\varphi$ we conclude $\mathcal{M}_1, w \models \blacklozenge\varphi$ by clause Zag$_\blacklozenge$ of Definition 2.  □

Just like for the standard modal language, the converse of Proposition 2 can be proven under the assumption that the models at issue are $\omega$-saturated. Before introducing such notion let us fix some notation. Given a finite set $Y$, the expansion of $\mathcal{L}^1$ with a finite set of constants $Y$ is denoted $\mathcal{L}_Y^1$, and the expansion of a Kripke model $\mathcal{M}$ to $\mathcal{L}_Y^1$ is denoted $\mathcal{M}_Y$.[2]

**Definition 3 ($\omega$-saturation).** *A model $\mathcal{M} = (W, R, V)$ is $\omega$-saturated if, and only if, for every $Y \subseteq W$ such that $|Y| < \omega$, the expansion $\mathcal{M}_Y$ realizes every set $\Gamma(x)$ of $\mathcal{L}_Y^1$-formulae whose finite subsets $\Gamma'(x) \subseteq \Gamma(x)$ are all realized in $\mathcal{M}_Y$.*

---

[2] For more on $\omega$-saturation we refer the reader to [4, Ch. 2] and [5, Ch. 2].

Intuitively, a model $\mathcal{M}$ is $\omega$-saturated if for any set of formulae $\Gamma(x, y_1, \ldots, y_n)$ over a finite set of variables, once some interpretation of $y_1, \ldots, y_n$ is fixed to, e.g., $w_1, \ldots, w_n$, and all finite subsets of $\Gamma(x)[w_1, \ldots, w_n]$ are realizable in $\mathcal{M}$, then the whole of $\Gamma(x)[w_1, \ldots, w_n]$ is realizable in $\mathcal{M}$. From a modal point of view, Definition 3 requires that if for any subset of $\Gamma$ there are accessible states satisfying it at the evaluation point, then there are accessible states satisfying the whole of $\Gamma$ at the evaluation point. This is precisely the property used in the proof of the following proposition.

**Proposition 3 ($\leftrightsquigarrow_s \subseteq \leftrightarrows_s$).** *For any two $\omega$-saturated pointed models $(\mathcal{M}_1, w_1)$ and $(\mathcal{M}_2, w_2)$ it holds that: $(\mathcal{M}_1, w_1) \leftrightsquigarrow_s (\mathcal{M}_2, w_2) \Longrightarrow (\mathcal{M}_1, w_1) \leftrightarrows_s (\mathcal{M}_2, w_2)$.*

*Proof.* It suffices to show that $\leftrightsquigarrow_s$ is an s-bisimulation (Definition 2). **Base:** The condition Atom is straightforwardly satisfied. **Step:** The proof for conditions $\mathrm{Zig}_\Diamond$ and $\mathrm{Zag}_\Diamond$ proceeds as usual for basic modal languages. We prove that the condition $\mathrm{Zig}_\blacklozenge$ is satisfied. Assume $(\mathcal{M}_1, w_1) \leftrightsquigarrow_s (\mathcal{M}_2, w_2)$ and $(\mathcal{M}_1, w_1) \xrightarrow{\blacklozenge} (\mathcal{M}_1', w_1)$. We show that there exists $(\mathcal{M}_2', w_2)$ such that $(\mathcal{M}_2, w_2) \xrightarrow{\blacklozenge} (\mathcal{M}_2', w_2)$ and $(\mathcal{M}_1', w_1) \leftrightsquigarrow_s (\mathcal{M}_2', w_2)$. We have that for any finite $\Gamma \subseteq \mathbb{T}^s(\mathcal{M}_1', w_1)$ the following sequence of equivalences holds:

$$\mathcal{M}_1, w_1 \models \blacklozenge \bigwedge \Gamma \Longleftrightarrow \mathcal{M}_2, w_2 \models \blacklozenge \bigwedge \Gamma$$
$$\Longleftrightarrow \mathcal{M}_2 \models ST_x^\emptyset \left( \blacklozenge \bigwedge \Gamma \right)[w_2]$$
$$\Longleftrightarrow \mathcal{M}_2 \models \exists y, z \left( yRz \wedge ST_x^{\{(y,z)\}} \left( \bigwedge \Gamma \right) \right)[w_2]$$

The first equivalence holds by the assumption of sabotage equivalence between $(\mathcal{M}_1, w_1)$ and $(\mathcal{M}_2, w_2)$. The second one follows by Theorem 2 and the third one by Definition 1. From this, by $\omega$-saturation of $\mathcal{M}_2$ we can conclude that:

there are $y, z \in \mathcal{M}_2$ such that $yRz$ and $\mathcal{M}_2 \models ST_x^{\{(y,z)\}}(\mathbb{T}^s(\mathcal{M}_1', w_1))[w_2]$.

By Theorem 2 there exists then a model $\mathcal{M}_2'$ such that $\mathcal{M}_2 \xrightarrow{\blacklozenge} \mathcal{M}_2'$ and $\mathcal{M}_2' \models ST_x^\emptyset(\mathbb{T}^s(\mathcal{M}_1', w_1))[w_2]$. By Theorem 2 we conclude that $(\mathcal{M}_1', w_1) \leftrightsquigarrow_s (\mathcal{M}_2', w_2)$, which completes the proof of the $\mathrm{Zig}_\blacklozenge$ clause. In the same way it can be proven that also the condition $\mathrm{Zag}_\blacklozenge$ is satisfied. $\square$

We have thus established a match between sabotage modal equivalence and sabotage bisimulation for the class of $\omega$-saturated models.

## 4   Characterization of SML by Invariance

In this section, we characterize SML as the one free variable fragment of FOL which is invariant under s-bisimulation.[3]

---

[3] Recall that the standard translation $ST^\emptyset$ of a sabotage modal logic formula always produces a FOL formula with only one free variable.

**Theorem 3 (Characterization of SML by s-bisimulation invariance).** *An $\mathcal{L}_1$-formula is equivalent to the translation of an $\mathcal{L}^s$ formula if, and only if, it is invariant for sabotage bisimulation.*

*Proof.* [Left to right] This direction follows from Proposition 2. [Right to left] We proceed as customary. Let $\varphi \in \mathcal{L}^1$ with one free variable $x$. Assume that $\varphi$ is invariant under s-bisimulation and consider the following set:

$$\mathbb{C}^s(\varphi) = \{ST_x^{\emptyset}(\psi) \mid \psi \in \mathcal{L}^s \text{ and } \varphi \models ST_x^{\emptyset}(\psi)\}.$$

The result follows from these two claims:

(i) If $\mathbb{C}^s(\varphi) \models \varphi$ then $\varphi$ is equivalent to the translation of an $\mathcal{L}^s$-formula.
(ii) It holds that $\mathbb{C}^s(\varphi) \models \varphi$, i.e., for any pointed model $\mathcal{M}, w$: if $\mathcal{M} \models \mathbb{C}^s(\varphi)[w]$ then $\mathcal{M} \models \varphi[w]$.

As to (i). Assume that $\mathbb{C}^s(\varphi) \models \varphi$. From the deduction and compactness theorems of FOL we have that $\models \bigwedge \Gamma \to \varphi$ for some finite $\Gamma \subset \mathbb{C}^s(\varphi)$. The converse holds by the definition of $\mathbb{C}^s(\varphi)$: $\models \varphi \to \bigwedge \Gamma$. We thus have that $\models \varphi \leftrightarrow \bigwedge \Gamma$ proving the claim.

As to (ii). Take a pointed model $\mathcal{M}, w$ such that $\mathcal{M} \models \mathbb{C}^s(\varphi)[w]$ and consider its sabotage modal theory $\mathbb{T}^s(\mathcal{M}, w)$. Now consider the set $\Sigma = ST_x^{\emptyset}(\mathbb{T}^s(\mathcal{M}, w)) \cup \{\varphi\}$. We proceed by showing that:

(a) $\Sigma$ is consistent;
(b) $\mathcal{M} \models \varphi[w]$, thus proving claim (ii).

To prove (a) assume, towards a contradiction, that $\Sigma$ is inconsistent. By the compactness of FOL we then obtain that $\models \varphi \to \neg \bigwedge \Gamma$ for some finite $\Gamma \in \Sigma$. But then, by the definition of $\mathbb{C}^s(\varphi)$, we have that $\neg \bigwedge \Gamma \in \mathbb{C}^s(\varphi)$, and hence $\neg \bigwedge \Gamma \in ST_x^{\emptyset}(\mathbb{T}^s(\mathcal{M}, w))$ which is impossible as $\Gamma \subset ST_x^{\emptyset}(\mathbb{T}^s(\mathcal{M}, w))$.

Now we will prove (b). As $\Sigma$ is consistent, it can be realized by a pointed model, which we call $\mathcal{M}', w'$. Observe, first of all, that $\mathcal{M}, w \leftrightsquigarrow_s \mathcal{M}', w'$ as they both have the same sabotage modal theory. Now take two $\omega$-saturated elementary extensions $(\mathcal{M}_\omega, w)$ and $(\mathcal{M}'_\omega, w')$ of $(\mathcal{M}, w)$ and $(\mathcal{M}', w')$. That such extensions exist can be proven by a chain construction argument (see [5, Proposition 3.2.6]). By the invariance of FOL under elementary extensions, since $\mathcal{M}' \models \varphi[w]$ (by the construction of $\Sigma$) we can conclude that $\mathcal{M}'_\omega \models \varphi[w]$. From this, by the assumption that $\varphi$ is invariant for s-bisimulation and Proposition 3, we conclude that $\mathcal{M}_\omega \models \varphi(x)[w]$ and again, by elementary extension, that $\mathcal{M} \models \varphi(x)[w]$, which establishes claim (ii) and completes the proof.      □

*Definable and undefinable properties in* SML. So which FOL properties belong to the fragment identified by Theorem 3 and which ones do not? We provide examples of SML-definable and undefinable (at model level) properties.
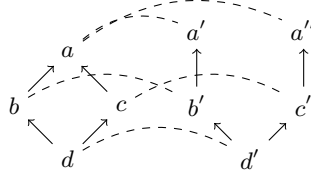
**Fig. 1.** Two s-bisimilar models (s-bisimulation rendered by the dashed lines). At state $d$ the property "all successors have one same successor" is true. It fails at state $d'$.
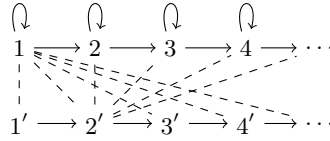


**Fig. 2.** Sabotage bisimulation between two frames (dashed lines). Only the part of the s-bisimulation relation originating in points 1 and $2'$ is depicted. The top frame is $\mathcal{F} = \langle \mathbb{N}, \geq \rangle$ (transitive edges are omitted) and the bottom model is $\mathcal{F}' = \langle \mathbb{N}, > \rangle$ (transitive edges are omitted).

*Example 1 (Counting successors).* Consider the FOL property "there exist at most $n$ successors" (2). This property is not bisimulation invariant, but it is invariant with respect to sabotage bisimulation. It is therefore definable in SML (by formula (3)).

*Example 2 (Confluence).* Consider the FOL property "all successors have one same successor". The property is not invariant for sabotage bisimulation. It is therefore not definable in SML. See Figure 1 for an illustration.

*Example 3 (Reflexive states).* Consider the FOL property $xRx$. This property is not invariant with respect to sabotage bisimulation. To witness this fact take two pointed models built on the set of natural numbers (with 0) where the point of evaluation is set at 0 and the accessibility relations are: on the first model the *greater or equal* relation (hence reflexive), and on the second one the strictly greater relation (hence irreflexive). That is: $\mathcal{M} = \langle \mathbb{N}, \geq \rangle$ and $\mathcal{M}' = \langle \mathbb{N}, > \rangle$. We have that $\langle \mathcal{M}, 0 \rangle \leftrightarroweq_s \langle \mathcal{M}', 0 \rangle$. Figure 2 depicts (part of) a relation which is a (standard) bisimulation $Z$ between the two models and which in addition has the property that any edge deletion on one model can be 'mirrored' on the other model obtaining pointed models that are still connected by $Z$ (recall Definition 2). In particular observe that the deletion of a reflexive edge in $\mathcal{M}$ at point $i$ can be 'mirrored' by the deletion of edge $(i, i+1)$ in $\mathcal{M}'$ (note that the accessibility relations are transitive in both models). However, $\mathcal{M} \models xRx[0]$ and $\mathcal{M}' \not\models xRx[0]$. Property $xRx$ is therefore not definable in SML.

## 5  Tableau Method for SML

Since SML is not invariant under (standard) bisimulation [9], it is clear that the sort of reduction argument normally used to obtain sound and complete axiom systems for logics of model update (cf. [16]) can not be applied as that would imply an embedding of SML into logic K. It is therefore natural to attempt a semantics-driven approach to the proof theory of SML, like a tableau method.

Moreover, SML does not have the tree-model property: there are specific SML formulae satisfied in Kripke models whose underlying frames can *not* be trees. For example, the formula $\Diamond\top \wedge \Box\Diamond\top \wedge \blacksquare\Box\bot$ is true in a model if and only if its underlying frame consists of one reflexive point. Hence, the labeled tableau system for logic K has to be adapted for SML.

**Definition 4 (Label, labeled formula and relation term).** *Let $S$ be an infinite set whose elements are called* labels*. An* extended label *is an expression of the form $\ell^E$ where $\ell \in S$ and $E$ is a finite set of pairs of $S$. A* labeled formula *is an expression of the form $(\ell^E\ \varphi)$ where $\ell^E$ is a label and $\varphi \in \mathcal{L}^s$. A* relation term *is an expression of the form $(R\ \ell_1\ \ell_2)$ where $\ell_1, \ell_2 \in S$.*

---

**Input**: A formula $\varphi \in \mathcal{L}^s$.

**Output**: A tableau $\mathcal{T}$ for $\varphi$: each branch may be infinite, finite and labeled open, or finite and labeled closed.

1. Initially, $\mathcal{T}$ is a tree consisting of a single root node labeled with $(\ell^\emptyset\ \varphi)$.
2. Repeat the following steps as long as possible:
   (a) *Choose* a branch which is neither closed nor open and choose a labeled formula $(\ell^E\ \psi)$ (or a pair of labeled formula $(\ell^E\ \psi)$ and relation term $(R\ \ell_1\ \ell_2)$) not selected before on this branch.
   (b) *Apply* the appropriate tableau rule of Figure 4 to $(\ell^E\ \psi)$ (or the pair $(\ell^E\ \psi)$, $(R\ \ell_1\ \ell_2)$):
       − if the tableau rule is rule $\neg\wedge$ (or rules $\Diamond$, $\blacklozenge$), add two successor nodes (resp. $n + 1$, $n$ successor nodes) to the branch labeled with the instantiations of the denominators of that rule,
       − otherwise, add a unique successor node labeled with the instantiation of the denominator(s) of that rule.
   (c)  i. *Label* by $\times$ (*closed*) the (new) branches which contain two labeled formulae $(\ell^E\ p)$ and $(\ell^F\ \neg p)$ (where $E$ and $F$ may possibly be different sets) or two labeled formulae $(\ell^E\ \varphi)$ and $(\ell^E\ \neg\varphi)$.
       ii. *Label* by $\odot$ (*open*) the (new) branches where there are no more formulae to decompose.

---

**Fig. 3.** Construction of a tableau.

**Definition 5 (Tableau).** *A* (labeled) tableau *is a tree whose nodes are labeled with labeled formulae or relation terms. The tableau tree for a formula is constructed as shown in the algorithm of Figure 3. In the tableau rules of Figure 4,*

$$\frac{(\ell^E\ \varphi \wedge \psi)}{(\ell^E\ \varphi)\ (\ell^E\ \psi)}\ \wedge \qquad \frac{(\ell^E\ \neg(\varphi \wedge \psi))}{(\ell^E\ \neg\varphi)\mid(\ell^E\ \neg\psi)}\ \neg\wedge \qquad \frac{(\ell^E\ \neg\neg\varphi)}{(\ell^E\ \varphi)}\ \neg\neg$$

$$\frac{(\ell_1^E\ \neg\Diamond\varphi)\quad(R\ \ell_1\ \ell_2)}{(\ell_2^E\ \neg\varphi)}\ \neg\Diamond \qquad \frac{(\ell^E\ \neg\blacklozenge\varphi)\quad(R\ \ell_1\ \ell_2)}{(\ell^{E\cup\{(\ell_1,\ell_2)\}}\ \neg\varphi)}\ \neg\blacklozenge$$

where $(\ell_1, \ell_2) \notin E$ in both rules above.

$$\frac{(\ell^E\ \Diamond\varphi)}{(R\ \ell\ \ell_1)(\ell_1^E\ \varphi)\mid \ldots \mid (R\ \ell\ \ell_n)(\ell_n^E\ \varphi)\mid(R\ \ell\ \ell_{n+1})(\ell_{n+1}^E\ \varphi)}\ \Diamond$$

where $\{\ell_1, \ldots, \ell_n\}$ are all the labels occurring in the current branch such that $(\ell, \ell_i) \notin E$ for all $i \in \{1, \ldots, n\}$ and $\ell_{n+1}$ is a 'fresh' label not occurring in the current branch.

$$\frac{(\ell^E\ \blacklozenge\varphi)}{(R\ \ell_1\ \ell_1')(\ell^{E\cup\{(\ell_1,\ell_1')\}}\ \varphi)\mid \ldots \mid (R\ \ell_n\ \ell_n')(\ell^{E\cup\{(\ell_n,\ell_n')\}}\ \varphi)}\ \blacklozenge$$

where $\{(\ell_1, \ell_1'), \ldots, (\ell_n, \ell_n')\} := (M \times M) \cup \{(\ell_+, \ell_{++})\} \setminus E$, with $M$ the set of labels occurring in the current branch to which we add a 'fresh' label $\ell_*$, and $(\ell_+, \ell_{++})$ is a pair of 'fresh' and distinct labels.

**Fig. 4.** Tableau rules.

*the formulae above the horizontal lines are called* numerators *and those below are called* denominators. *A tableau* closes *when all its branches are closed. A branch is* open *when it is infinite or it terminates in a leaf labeled open.*

The construction of a tableau may not necessarily terminate (see Example 4). This is in line with the fact that the satisfiability problem of SML is undecidable. Nevertheless, a tableau closes only if the construction terminates. Note that if we remove the rules for sabotage we obtain a sound and complete tableau method for logic K which is somewhat non-standard (and computationally demanding).

**Theorem 4 (Soundness and completeness).** *Let $\varphi \in \mathcal{L}^s$. If $\varphi$ is unsatisfiable, then the tableau for $\varphi$ closes (completeness). If the tableau for $\varphi$ closes then $\varphi$ is unsatisfiable (soundness).*

*Example 4.* In Figure 5, on the right, we display the execution of the tableau method of Figure 3 on the formula $\Diamond\top \wedge \Box\Diamond\top \wedge \blacksquare\Box\bot$. We obtain a single open branch (labeled with $\odot$) from which we can extract a model whose frame is a single reflexive point. This formula is thus satisfiable, and in fact only in this frame. In Figure 5, on the left, we show that the tableau construction may not necessarily terminate by exhibiting an infinite branch in the tableau for the formula $\Diamond\top \wedge \blacksquare\Diamond\top$. Even if the formula holds in pointed models having at least
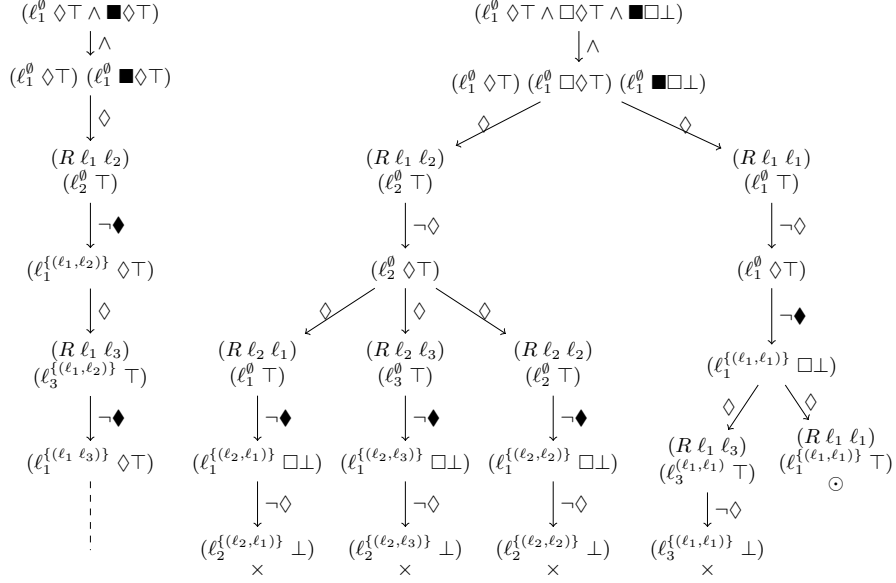
**Fig. 5.** An infinite branch in the tableau for $\Diamond\top \wedge \blacksquare\Diamond\top$ (*left*), and tableau for $\Diamond\top \wedge \square\Diamond\top \wedge \blacksquare\square\bot$ (*right*).

two successors, our tableau method does not terminate with this formula as input and produces a pointed model with infinitely many successors.

## 6   Conclusions and Future Work

We have touched upon some model theoretic aspects of SML and fleshed out the theory of a standard translation for SML, which was only sketched in [10]. We have studied such translation together with a notion of bisimulation tailored to SML thereby establishing a novel characterization theorem for the logic. We have also provided the first proof system for SML in the form of a sound and complete tableau method.

SML remains a rather under-investigated formalism and many natural questions are still open. We conclude by mentioning a few. First, it is unclear to what extent standard techniques of modal correspondence theory (see [4, Ch. 3]) are applicable to SML. In particular, can the Sahlqvist theorem be extended to SML? Second, the set of valid formulae of SML is not closed under uniform substitution (e.g., $p \leftrightarrow \blacksquare p$). Is the set of schematic validities of SML decidable? Is it axiomatizable? Third, SML is not a well-behaved logic (recall Theorem 1). The fact that edge deletion is arbitrary seems to be the key feature that sets SML apart from better behaved logics in the dynamic epistemic logic landscape where deletions, even of a very general kind, are definable (e.g., [15]). Are there natural

restrictions on the semantics of SML (e.g., 'localized' edge deletion) which yield better behaved variants? Finally, the notion of sabotage bisimulation suggests a natural operationalization of equivalence in terms of model comparison games. How such games relate to the original sabotage game of [13] and what further insights they can give into SML are also worthwhile lines of research.

## References

1. C. Areces, R. Fervari, and G. Hoffmann. Moving arrows and four model checking results. In L. Ong and R. Queiroz, editors, *Logic, Language, Information and Computation*, volume 7456, pages 145–153, 2012.
2. C. Areces, R. Fervari, and G. Hoffmann. Swap logic. *Logic Journal of the IGPL*, 22(2):309–332, 2014.
3. G. Aucher, P. Balbiani, L. Fariñas del Cerro, and A. Herzig. Global and local graph modifiers. *Electronic Notes in Theoretical Computer Science*, 231:293–307, 2009.
4. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, Cambridge, 2001.
5. C. C. Chang and H. J. Keisler. *Model Theory*. Studies in Logic and the Foundations of Mathematics. North-Holland, 1973.
6. N. Gierasimczuk, L. Kurzen, and F. Velázquez-Quesada. Learning and teaching as a game: a sabotage approach. In X. He, J. Horty, and E. Pacuit, editors, *Proceedings of LORI-2*, number 5834 in LNAI, 2009.
7. S. Gruener, F. Radmacher, and W. Thomas. Connectivity games over dynamic networks. *Theoretical Computer Science*, 498:46–65, 2013.
8. B. Kooi and B. Renne. Arrow update logic. *Review of Symbolic Logic*, 4(4), 2011.
9. C. Löding and P. Rohde. Model checking and satisfiability for sabotage modal logic. In P. K. Pandya and J. Radhakrishnan, editors, *FSTTCS 2003*, volume 2914 of *LNCS*, pages 302–313. Springer, 2003.
10. C. Löding and P. Rohde. Solving the sabotage game is PSPACE-hard. Technical report, Department of Computer Science RWTH Aachen, 2003.
11. F. Radmacher and W. Thomas. A game theoretic approach to the analysis of dynamic networks. *Electronic Notes in Theoretical Computer Science*, 200(2):21–37, 2008.
12. J. van Benthem. *Modal Logic and Classical Logic*. Monographs in Philosophical Logic and Formal Linguistics. Bibliopolis, 1983.
13. J. van Benthem. An essay on sabotage and obstruction. In D. Hutter and W. Stephan, editors, *Mechanizing Mathematical Reasoning*, volume 2605 of *LNCS*, pages 268–276. Springer, 2005.
14. J. van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011.
15. J. van Benthem and F. Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logic*, 17(2), 2007.
16. H. van Ditmarsch, B. Kooi, and W. van der Hoek. *Dynamic Epistemic Logic*, volume 337 of *Synthese Library Series*. Springer, 2007.

# A  Soundness and Completeness of the Tableau Method

## A.1  Soundness

The soundness is proved by resorting to the notion of *interpretability*. A set of labeled formulae $L$ is *interpretable* if there is a Kripke model $\mathcal{M} = (W, R, V)$ and a mapping $f : S \longrightarrow W$ such that for all $(\ell^E \; \varphi) \in L$, we have that $(W, R \setminus f(E), V), f(\ell) \models \varphi$, where $f(E) = \{(f(\ell_1), f(\ell_2)) \mid (\ell_1, \ell_2) \in E\}$. Then, we prove two facts:

**Fact 1** *If $\varphi$ is satisfiable, then, at any step of the construction of the tableau for $\varphi$, the set of labeled formulae of some branch is interpretable.*

**Fact 2** *If $\varphi$ is satisfiable, then any branch whose set of labeled formulae is interpretable cannot close. That is, there is an extension of this interpretable branch which does not close (this extended branch is therefore either open or infinite).*

These two facts combined together prove that if $\varphi$ is satisfiable, then the tableau for $\varphi$ cannot close. That is, by contraposition, if the tableau for $\varphi$ closes then $\varphi$ is unsatisfiable (soundness).

*Proof (Fact 1).* We prove the first fact by induction on the number of times $n$ we use inference rules in the construction of the tableau for $\varphi$. The case $n = 0$ holds trivially: in that case $L$ is a singleton $\{(\ell \; \varphi)\}$ and it suffices to define $f$ so that it assigns to $\ell$ the world of the Kripke model where $\varphi$ is satisfiable. The induction step $n + 1$ is proved by examining each rule on a case by case basis and by showing that for each rule we can extend the mapping $f$ associated to an interpretable branch to assign world(s) to the new label(s) created by the application of the rule and we can also extend the range of the accessibility relation $R$ to assign a pair of worlds to the new relation term created by the rule. Indeed, we know by the Induction Hypothesis that there is a branch of the tableau whose terms are all interpretable by this mapping $f$. The key steps to consider are the rules $\Diamond$ and $\blacklozenge$.

- Rule $\Diamond$: Assume that the interpretable branch contains a labeled formula $(\ell^E \; \Diamond \varphi)$ (not chosen before in the execution of the tableau method). Applying the rule $\Diamond$ to this interpretable branch, we obtain $n + 1$ extended branches. We must show that one of them is interpretable. By assumption, we have that $(W, R \setminus f(E), V), f(\ell) \models \Diamond \varphi$. Therefore, there is $w \in W$ such that $f(\ell) R w$ and $\mathcal{M}, w \models \varphi$. If $w$ already corresponds to a label $\ell'$ such that $f(\ell') = w$, then we are in one of the first $n$ extensions of the interpretable branch. In that case, the mapping $f$ does not need to be extended and the label $\ell'$ has already been introduced in a rule earlier in the execution of the tableau method. Otherwise, we are in the last case of the rule $\Diamond$ and we need to extend the mapping $f$ and assign to the 'fresh' label $\ell_{n+1}$ the possible world $w$: we set $f(\ell_{n+1}) := w$.

– Rule ♦: Assume that the interpretable branch contains a labeled formula $(\ell^E \,\blacklozenge\varphi)$ (not chosen before in the execution of the tableau method). Applying the rule ♦ to this interpretable branch, we obtain $n$ extended branches corresponding to the $n$ elements of $(M \times M) \cup \{(\ell_+, \ell_{++})\} \setminus E$ (see the rule ♦). We must show that one of them is interpretable. By assumption, we have that $(W, R \setminus f(E), V), f(\ell) \models \blacklozenge\varphi$. Therefore, there is a pair $(w, v) \in R \setminus f(E)$ such that $(W, R \setminus (f(E) \cup \{(w, v)\}), V), f(\ell) \models \varphi$. This pair of worlds $(w, v)$ is either of the form $(f(\ell_0), f(\ell'_0))$, $(f(\ell_0), f(\ell_0))$, $(f(\ell_0), v)$, $(w, f(\ell_0))$, $(w, w)$ or simply $(w, v)$, for some labels $\ell_0, \ell'_0$ already introduced in this interpretable branch. The first five cases are covered by some of the cases corresponding to the elements of $M \times M$ of rule ♦ and the last case is covered by the case corresponding to $(\ell_+, \ell_{++})$ of rule ♦. So, at least one of the extended branches is interpretable and we can extend the mapping $f$ accordingly.

*Proof (Fact 2).* We prove the second fact by contraposition. Assume that any extension of the initial interpretable branch closes. Then, any extension of this initial branch is such that it contains a labeled formula $(\ell^E \,\varphi)$ and its negation $(\ell^E \,\neg\varphi)$ or two labeled formulae $(\ell^E \,p)$ and $(\ell^F \,\neg p)$. Then, this entails that the set of labeled formulae of any extended branch is *not* interpretable. So, by Fact 1, since $\varphi$ is satisfiable by assumption, this entails that the set of labeled formulae of the *initial* branch is also not interpretable, which is impossible by assumption. So, we have proved the second fact by contraposition.

## A.2 Completeness

We prove completeness by contraposition. Assume that the tableau for $\varphi$ does not close. Then, there is an open branch in the tableau for $\varphi$. Let $L$ be the set of labeled formulae appearing on this open branch and let $T$ be the set of relation terms appearing on this open branch. We build the Kripke model $\mathcal{M} := (W, R, V)$ as follows:

– $W := \{\ell \mid (\ell^E \,\varphi) \in L$ for some $\varphi \in \mathcal{L}^s$ and $E \in 2^{S \times S}\}$;
– $R := \{(\ell_1, \ell_2) \mid (R \,\ell_1 \,\ell_2) \in T\}$ and
– $V(p) := \{\ell \in W \mid (\ell^E \,p) \in L\}$, for all $p \in \mathbf{P}$.

Then, we have the following fact:

**Fact 3** *For all labeled formulae* $(\ell^E \,\chi)$,

$$\text{if } (\ell^E \,\chi) \in L \text{ then } (W, R \setminus E, V), \ell \models \chi. \tag{4}$$

*Proof.* We prove Expression (4) by induction on the size of $\chi$. The base case $\chi := p$ holds by definition of $V$. We prove the induction steps:

– $\chi := \varphi \wedge \psi$: Assume that $(\ell^E \,\varphi \wedge \psi) \in L$. Then, by saturation of the tableau rules, we also have that $(\ell^E \,\varphi)$ and $(\ell^E \,\psi)$ are in $L$. Then, by Induction Hypothesis, we must have that $(W, R \setminus E, V), \ell \models \varphi$ and $(W, R \setminus E, V), \ell \models \psi$. Hence, we obtain that $(W, R \setminus E, V), \ell \models \varphi \wedge \psi$.

- $\chi := \Diamond\varphi$: Assume that $(\ell^E \ \Diamond\varphi) \in L$. Then, by saturation of the tableau rules, there is $(R \ \ell \ \ell') \in T$ such that $(\ell, \ell') \notin E$ and $(\ell'^E \ \varphi) \in L$. Then, by Induction Hypothesis, $(W, R \setminus E, V), \ell' \models \varphi$ and $(\ell, \ell') \in R \setminus E$. Hence, we have that $(W, R \setminus E, V), \ell \models \Diamond\varphi$.
- $\chi := \blacklozenge\varphi$. Assume that $(\ell^E \ \blacklozenge\varphi) \in L$. Then, by saturation of the tableau rules, there is $(R \ \ell \ \ell') \in T$ such that $(\ell, \ell') \notin E$ and $(\ell^{E \cup \{(\ell, \ell')\}} \ \varphi) \in L$. Then, by Induction Hypothesis, $(W, R \setminus (E \cup \{(\ell, \ell')\}), V), \ell \models \varphi$ and $(\ell, \ell') \in R \setminus E$. Hence, we have that $(W, R \setminus E, V), \ell \models \blacklozenge\varphi$.
- $\chi := \neg p$: Assume that $(\ell^E \ \neg p) \in L$ and assume towards a contradiction that $(W, R \setminus E, V), \ell \models p$. Then, by definition of $\mathcal{M}$, there is a set of pairs of labels $F$ such that $(\ell^F \ p) \in L$. However, if both $(\ell^E \ \neg p)$ and $(\ell^F \ p)$ belong to the same branch, the branch cannot be open, which is impossible by assumption.
- $\chi := \neg(\varphi \wedge \psi)$: Assume that $(\ell^E \ \neg(\varphi \wedge \psi)) \in L$. Then, by saturation of the tableau rules, we must have either that $(\ell^E \ \neg\varphi) \in L$ or $(\ell^E \ \neg\psi) \in L$. Then, by Induction Hypothesis, we must have either that $(W, R \setminus E, V), \ell \models \neg\varphi$ or $(W, R \setminus E, V), \ell \models \neg\psi$. In both cases, we obtain that $(W, R \setminus E, V), \ell \models \neg(\varphi \wedge \psi)$.
- $\chi := \neg\neg\varphi$: Assume that $(\ell^E \ \neg\neg\varphi) \in L$. Then, by saturation of the tableau rules, we obtain that $(\ell^E \ \varphi) \in L$. So, by Induction Hypothesis, we have that $(W, R \setminus E, V), \ell \models \varphi$, and therefore also $(W, R \setminus E, V), \ell \models \neg\neg\varphi$.
- $\chi := \neg\Diamond\varphi$: Assume that $(\ell^E \ \neg\Diamond\varphi) \in L$. Then, for all $(R \ \ell \ \ell') \in T$ such that $(\ell, \ell') \notin E$, $(\ell'^E \ \neg\varphi) \in L$ by saturation of the tableau rules. Therefore, by Induction Hypothesis, $(W, R \setminus E, V), \ell' \models \neg\varphi$, for all $(\ell, \ell') \in R \setminus E$. So, $(W, R \setminus E, V), \ell \models \neg\Diamond\varphi$.
- $\chi := \neg\blacklozenge\varphi$. Assume that $(\ell^E \ \neg\blacklozenge\varphi) \in L$. Then, by saturation of the rule $\neg\blacklozenge$ of the tableau rules, for all $(R \ \ell \ \ell') \in T$ such that $(\ell, \ell') \notin E$, we must have that $(\ell^{E \cup \{\ell, \ell')\}} \ \neg\varphi) \in L$. So, by Induction Hypothesis, this entails that $(W, R \setminus (E \cup \{(\ell, \ell')\}), V), \ell \models \neg\varphi$. That is, $(W, R \setminus E, V), \ell \models \neg\blacklozenge\varphi$.

Thus, in particular, since $(\ell^\emptyset \ \varphi) \in L$ is the root of the tableau for $\varphi$, we have from Expression (4) that $\mathcal{M}, \ell \models \varphi$. Hence, $\varphi$ is satisfiable. This proves the completeness of our tableau method.