

Losing Connection: the Modal Logic of Definable Link Deletion

Dazhu Li

Department of Philosophy, Tsinghua University, Beijing
 Institute for Logic, Language and Computation, University of Amsterdam, P.O. Box
 94242, 1090 GE Amsterdam, The Netherlands
 E-mail: lidazhu91@163.com

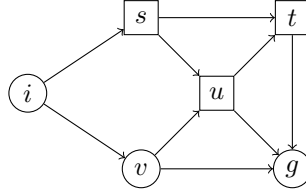
Abstract

In this article, we start with a two-player game that models communication under adverse circumstances in everyday life and study it from the perspective of a modal logic of graphs, where links can be deleted locally according to definitions available to the adversarial player. We first introduce a new language, semantics, and some typical validities. We then formulate a new type of first-order translation for this modal logic and prove its correctness. Then, a novel notion of bisimulation is proposed which leads to a characterization theorem for the logic as a fragment of first-order logic, and a further investigation is made of its expressive power against hybrid modal languages. Next, we discuss how to axiomatize this logic of link deletion, using dynamic-epistemic logics as a contrast. Finally, we show that our new modal logic lacks both the tree model property and the finite model property, and that its satisfiability problem is undecidable.

Keywords: Graph Game, Modal Logic, Dynamic Logic, Link Deletion, Undecidability

1 Introduction

In the graph of the World-Wide Web, to search for relevant and valuable information, a computer user usually clicks through consecutive hyperlinks passing through intermediate web pages. However, hyperlinks do not always work: say, because of technical malfunctions, or more interestingly, intentional obstruction. Such scenarios of search under adverse circumstances are quite common, and formally, they can be modeled as non-cooperative games played on graphs. For instance, consider the following web graph:



In this picture, nodes stand for web pages, directed arrows are hyperlinks, and the two kinds of shape, square and circle, denote two different properties of web pages. One player E , the user in the above scenarios, starts at point i , and tries to arrive at one of the goal points t and g . The other player A , say, Nature or some intentional opponent, tries to prevent this. The game goes in rounds: A first cuts one or more links in the graph, then E makes a step along some still available link. Since A can cut at most 8 links in all, the game is finite. E wins if she gets to the goal region, and loses if she cannot get there.

This description still leaves the game underspecified, since we must say more about how player A is allowed to cut before we can analyze the outcomes of the game. For concreteness, we start with a variant where the properties are not yet essential.

First Version Player A cuts one arrow from E 's current position to some reachable node.

In the resulting game on our graph, player E has a winning strategy: she is always able to find the information that she needs. Player A might start by deleting the link $\langle i, s \rangle$, then E moves to node v . In the second round, A must cut $\langle v, g \rangle$, and E goes to state u . Finally, player E can always arrive at t or g whatever link A deletes.

In this first version, the game is a local variant of the *sabotage game* (SG) in [10]. A sabotage game is played on a graph by two players: in each round, *Traveler* acts in the same way as E , while A 's counterpart *Demon* first cuts a link. However, *Demon*'s moves in sabotage games are global and allow cutting a link anywhere in the graph, not necessarily starting at the current position of *Traveler*. In contrast, our game restricts the moves available to *Demon*, giving him fewer winning strategies in general (cf. [5]).

However, the real-world scenarios that we considered suggest a more drastic deviation from existing sabotage games. In many cases of obstruction, the hostile opponent can cut more than one link, following a recipe rather than some arbitrary choice. For instance, blocking of links between computers is usually done by a program working on some explicit description of the targets to be blocked. Or for another concrete illustration of locality and definability, agents in a social network can cut friendship links starting with themselves, and they will often do that cutting according to some rule, such as 'delete all links to people that have proved to be dishonest'.

Our next game models such more realistic scenarios, taking care of both aspects.

Definitive Version In each round, player A chooses an available atomic property, and cuts all links from the position of E to nodes with the chosen property.

For example, in the above graph, when E is located at node s , A can cut both the links $\langle s, u \rangle$ and $\langle s, t \rangle$ if he chooses the definable property of nodes marked by the square.

Clearly, with this new version, A 's powers of blocking access to information have increased. Indeed, on the same graph as before, he now has a winning strategy. In the first

round, A cuts the link $\langle i, v \rangle$, and E 's only option is to move to node s . But then, A can cut both links $\langle s, u \rangle$ and $\langle s, t \rangle$ simultaneously, and E gets stuck and loses.

We will now focus on the logical analysis of our second more realistic game, calling it the *definable sabotage game*, denoted S_dG . Here existing modal logics for sabotage can serve as an inspiration, given the similarity of the games. But they must be modified, since we have made the obstructing player both less powerful (given the local nature of his choices) and more powerful (since he can remove more than one link in general). More concretely, to analyze the sabotage game, [5] presents a *sabotage modal logic (SML)* extending standard modal logic with a modality $\blacklozenge\varphi$ stating that φ is true at the evaluation point after removing some accessibility arrow from the model. But what is a suitable logic for S_dG ? The next section contains our proposal, called *definable sabotage modal logic (S_dML)*. We will study this logic in depth, not just for its connections to the above games, but also as a pilot study for throwing light on what is special and what is general about sabotage games, and the logical theory that already exists for them. In addition, our logic is a test case for how local sabotage, even though definable in ways reminiscent of dynamic-epistemic logics of information update, has its own behavior, including significantly higher complexity (cf. [9]).

Outline of the Paper. In Section 2, we present the syntax and semantics of S_dML (Section 2.1), and some typical logical validities (Section 2.2). In Section 3, we describe the non-trivial first-order translation for S_dML and check its correctness. In Section 4, we first introduce a notion of bisimulation for S_dML and investigate some of its model theory (Section 4.1), then we prove a characterization theorem for S_dML as a fragment of first-order logic that is invariant for the bisimulation introduced (Section 4.2), and finally we explore the expressive power of S_dML (Section 4.3). In Section 5, we provide some further analysis of an axiomatization of S_dML . In particular, we illustrate the relation between S_dML and hybrid logics (Section 5.1), and study recursion axioms (Section 5.2). Next, in Section 6, we show that S_dML lacks both the tree model property and the finite model property, and that the satisfiability problem for S_dML is undecidable. Finally, we discuss related work in Section 7, and conclude in Section 8 with a summary and outlook.

2 Language, Semantics and Logical Validities

In this section, we introduce the syntax and semantics of S_dML . After that, to understand the new device, we illustrate some properties of the logic by means of logical validities.

2.1 Language and Semantics

As mentioned above, the definable sabotage modal logic S_dML is intended to match S_dG . Therefore its language should be expressive enough to model the actions of the players. For player E , it is natural to think of the standard modality \Diamond , which characterizes the transition from a node to its successors (see [14]). However, to characterize the action of A , some dynamic operator is indispensable.

The language \mathcal{L}_d of $S_d\text{ML}$ is a straightforward extension of the standard modal language \mathcal{L}_\square . In addition to the modality \Diamond , it also includes a dynamic modal operator $[-]$. The formal definition is as follows:

Definition 1 (Language). *Let \mathbf{P} be a countable set of propositional atoms. The formulas of \mathcal{L}_d are defined by the following grammar in Backus-Naur Form:*

$$\mathcal{L}_d \ni \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid [-\varphi]\varphi$$

where $p \in \mathbf{P}$. Besides, notions \top , \perp , \vee , \rightarrow and \Diamond are as usual. For any $[-\varphi]\psi \in \mathcal{L}_d$, we define $\langle -\varphi \rangle\psi := \neg[-\varphi]\neg\psi$, i.e., $\langle - \rangle$ is the dual operator of $[-]$.

We will often omit parentheses when doing so ought not cause confusion. The operator $[-]$ is our device to model the action of A in $S_d\text{G}$. This can be clarified by the semantics of $S_d\text{ML}$. Formulas of \mathcal{L}_d are evaluated in standard relational models $\mathcal{M} = \langle W, R, V \rangle$, where W is the domain, a non-empty set of states, nodes or points, $R \subseteq W^2$ is the set of accessibility relations or links between points, and $V : \mathbf{P} \rightarrow 2^W$ is the valuation function. A pair $\mathcal{F} = \langle W, R \rangle$ is called a *frame*. For each $w \in W$, $\langle \mathcal{M}, w \rangle$ is a *pointed model*. For brevity, we usually write \mathcal{M}, w instead of $\langle \mathcal{M}, w \rangle$. For any $\langle w, v \rangle \in R$, we also write $\langle w, v \rangle \in \mathcal{M}$. Besides, we use $R(w)$ to denote the set $\{v \in W \mid \langle w, v \rangle \in R\}$ of successors of w . We now introduce the semantics, which is defined inductively by truth conditions.

Definition 2 (Semantics). *Given a pointed model $\langle \mathcal{M}, w \rangle$ and a formula φ of \mathcal{L}_d , we say that φ is true in \mathcal{M} at w , written as $\mathcal{M}, w \models \varphi$, when*

$$\begin{aligned} \mathcal{M}, w \models p & \text{ iff } w \in V(p) \\ \mathcal{M}, w \models \neg\varphi & \text{ iff } \mathcal{M}, w \not\models \varphi \\ \mathcal{M}, w \models \varphi \wedge \psi & \text{ iff } \mathcal{M}, w \models \varphi \text{ and } \mathcal{M}, w \models \psi \\ \mathcal{M}, w \models \Box\varphi & \text{ iff for each } v \in W, \text{ if } R w v, \text{ then } \mathcal{M}, v \models \varphi \\ \mathcal{M}, w \models [-\varphi]\psi & \text{ iff } \mathcal{M}|_{\langle w, \varphi \rangle}, w \models \psi \end{aligned}$$

where $\mathcal{M}|_{\langle w, \varphi \rangle} = \langle W, R \setminus (\{w\} \times V(\varphi) \cap R(w)), V \rangle = \langle W, R \setminus (\{w\} \times V(\varphi)), V \rangle$ is obtained by deleting all links from w to the nodes that are φ .

We say that formula φ is *satisfiable* if there exists a pointed model $\langle \mathcal{M}, w \rangle$ such that $\mathcal{M}, w \models \varphi$. By Definition 2, the truth conditions for Boolean and modal connectives \neg , \wedge , \Box are as usual, and $[-\varphi]\psi$ means that ψ is true at the evaluation point after deleting all accessibility relations from the current point to the nodes that are φ . Besides, we say that two pointed models $\langle \mathcal{M}_1, w \rangle$ and $\langle \mathcal{M}_2, v \rangle$ are *φ -sabotage-related* (notation, $\langle \mathcal{M}_1, w \rangle \xrightarrow{\varphi} \langle \mathcal{M}_2, v \rangle$) iff $\langle \mathcal{M}_2, v \rangle$ is $\langle \mathcal{M}_1|_{\langle w, \varphi \rangle}, w \rangle$. Intuitively, by the semantics, formula φ occurring in $[-]$ stands for a property of some successors of the current point, and $[-\varphi]$ is exactly an action of player A in $S_d\text{G}$.

Example Revisited. Recall the graph at the outset. Assume that the propositional atoms p and q refer to the properties denoted with circle and square respectively. Then we are able to express the facts of the game with formulas of \mathcal{L}_d . For instance, that ‘after A

deletes the links from v to the circle point, i.e., g , E still can move to a square node, i.e., u' can be expressed as the truth at v of the formula $[-p] \Diamond q$. Besides, \mathcal{L}_d can also define the existence of winning strategies for players. For example, the formula $[-p] \Box [-q] \Box \perp$ states that A can stop E successfully by removing the links from the position of E to the circle nodes in the first round, and cutting the links pointing to the square nodes in the second round. By our semantics for these formulas, $S_d\text{ML}$ captures $S_d\text{G}$ precisely.

2.2 Logical Validities

Although the language and semantics of $S_d\text{ML}$ look simple, there are some issues with the new operator $[-]$. To illustrate how it works, we explore some interesting validities of $S_d\text{ML}$. First of all, let us consider the following principle:

$$[-\varphi](\varphi_1 \rightarrow \varphi_2) \rightarrow ([-\varphi]\varphi_1 \rightarrow [-\varphi]\varphi_2) \quad (1)$$

which follows from the semantics of $S_d\text{ML}$ directly. The formula enables us to distribute $[-]$ over an implication. It is a common principle that applies to almost all modalities, e.g. the standard modality and the public announcement operator (see, e.g. [6]). However, operator $[-]$ also has some distinguishing features. For instance, the validity

$$[-\varphi]\psi \leftrightarrow \langle -\varphi \rangle \psi \quad (2)$$

illustrates that $[-]$ is self-dual and—less obviously—a model update function essentially. It is not hard to check that the validity of formulas (1) and (2) is closed under substitution. Interestingly, this is not a common feature of $S_d\text{ML}$. Some examples are as follows:

$$[-\varphi]p \leftrightarrow p \quad (3)$$

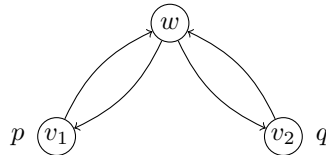
$$[-p] \Diamond q \leftrightarrow \Diamond(\neg p \wedge q) \quad (4)$$

$$[-p][-q]\varphi \leftrightarrow [-q][-p]\varphi \quad (5)$$

Principle (3) illustrates that operator $[-]$ does not change the truth value of propositional atoms. Formula (4) allows us to reduce a formula including $[-]$ to an \mathcal{L}_\Box -formula. By (5), when all formulas occurring in $[-]$ are propositional atoms, the order of different operators $[-]$ can be interchanged.

Actually each propositional atom occurring in formulas (3)-(5) can be replaced by any Boolean formula without affecting their validity. However, these schematic validities fail in general when we consider the deletions for complex properties. As an example, we show this phenomenon for principle (5).

Example 1. Consider the general schematic form $[-\varphi_1][- \varphi_2]\varphi \leftrightarrow [-\varphi_2][- \varphi_1]\varphi$ for the principle (5). Let $\varphi_1 := p$, $\varphi_2 := \Diamond \Diamond p$, and $\varphi := \Diamond q$. Define a model \mathcal{M} as follows:



By inspection, one sees that $\mathcal{M}, w \models [-p][-\Diamond p]\Diamond q$ and $\mathcal{M}, w \not\models [-\Diamond p][-\Diamond p]\Diamond q$. Therefore it holds that $\mathcal{M}, w \not\models [-p][-\Diamond p]\Diamond q \leftrightarrow [-\Diamond p][-\Diamond p]\Diamond q$.

Many instances of validity in $S_d\text{ML}$ are not straightforward, and require much more thought than the often rather obvious validities found in standard logical systems. In particular, the dynamic modality $[-]$ creates interesting complexity, since removing a link in a model can have side-effects for truth values of formulas at worlds throughout the model. Therefore, it is time to make a deeper technical investigation of our logic.

3 First-Order Translation for $S_d\text{ML}$

Given the semantics of $S_d\text{ML}$, a natural question is: is $S_d\text{ML}$ axiomatizable? Obviously the truth conditions for $S_d\text{ML}$ are first-order, so there must be a first-order translation like that for standard modal logic. In this section, we present a positive answer to the question by describing a recursive standard translation for $S_d\text{ML}$.

However we already know from SML that additional arguments may be needed in the translation: for SML, that extra argument was a finite set of links (see [5]). Interestingly, finding the translation here requires even more delicate analysis of the extra argument.

To do so, our method is to introduce a new device, being a sequence consisting of ordered pairs, e.g. $\langle v, \varphi \rangle$, to denote the occurrences of $[-]$ in a formula, where v is a variable and φ is a property of its successors. Let \mathcal{L}_1 be the first-order language consisting of countable unary predicates $P_{i \in \mathbb{N}}$, a binary relation R , and equivalence \equiv .

Definition 3 (Standard Translation for $S_d\text{ML}$). *Let x be a designated variable, and O be a finite sequence $\langle v_0, \psi_0 \rangle; \dots; \langle v_i, \psi_i \rangle; \dots; \langle v_n, \psi_n \rangle$ ($0 \leq i \leq n$), where $\psi_{0 \leq i \leq n}$ is an \mathcal{L}_d -formula and $v_{0 \leq i \leq n}$ is a variable. Then the translation $ST_x^O : \mathcal{L}_d \rightarrow \mathcal{L}_1$ is defined recursively as follows:*

$$\begin{aligned}
 ST_x^O(p) &= Px \\
 ST_x^O(\top) &= x \equiv x \\
 ST_x^O(\neg\varphi) &= \neg ST_x^O(\varphi) \\
 ST_x^O(\varphi_1 \wedge \varphi_2) &= ST_x^O(\varphi_1) \wedge ST_x^O(\varphi_2) \\
 ST_x^O(\Diamond\varphi) &= \exists y (Rxy \wedge \neg(x \equiv v_0 \wedge ST_y^{(x, \perp)}(\psi_0)) \wedge \\
 &\quad \bigwedge_{0 \leq i \leq n-1} \neg(x \equiv v_{i+1} \wedge ST_y^{(v_0, \psi_0); \dots; (v_i, \psi_i)}(\psi_{i+1})) \wedge ST_y^O(\varphi)) \\
 ST_x^O([- \varphi_1] \varphi_2) &= ST_x^{O; \langle x, \varphi_1 \rangle}(\varphi_2)
 \end{aligned}$$

The key inductive clauses in Definition 3 concern \Diamond -formulas and $[-]$ -formulas. Formula $\Diamond\varphi$ is translated as a first-order formula stating that the current point x has a successor y which is φ , and that this accessibility relation is not deleted by the operator $[-]$ indexed in the sequence O . The first-order translation for $[-\varphi_1]\varphi_2$ says that the translation of φ_2 is carried out with respect to the sequence $O; \langle x, \varphi_1 \rangle$, and that this translation is realized at the current point x .

According to Definition 3, the index sequence O may become longer and longer, but it is always finite. For each formula φ of \mathcal{L}_d , $ST_x^{\langle x, \perp \rangle}(\varphi)$ yields a first-order formula with only x free. Now we use an example to illustrate the translation.

Example 2. Consider formula $\Diamond[-\Diamond p_1]\Box p_2$. Its translation runs as follows:

$$\begin{aligned}
ST_x^{\langle x, \perp \rangle}(\Diamond[-\Diamond p_1]\Box p_2) &= \exists y(Rxy \wedge \neg(x \equiv x \wedge ST_y^{\langle x, \perp \rangle}(\perp)) \wedge \\
&\quad ST_y^{\langle x, \perp \rangle}([-\Diamond p_1]\Box p_2)) \\
&= \exists y(Rxy \wedge \neg(x \equiv x \wedge ST_y^{\langle x, \perp \rangle}(\perp)) \wedge \\
&\quad ST_y^{\langle x, \perp \rangle; \langle y, \Diamond p_1 \rangle}(\Box p_2)) \\
&= \exists y(Rxy \wedge \neg(x \equiv x \wedge ST_y^{\langle x, \perp \rangle}(\perp)) \wedge \\
&\quad \forall z(Ryz \wedge \neg(y \equiv x \wedge ST_z^{\langle x, \perp \rangle}(\perp)) \wedge \\
&\quad \neg(y \equiv y \wedge ST_z^{\langle x, \perp \rangle}(\Diamond p_1)) \rightarrow ST_z^{\langle x, \perp \rangle; \langle y, \Diamond p_1 \rangle}(p_2)) \\
&= \exists y(Rxy \wedge \neg(x \equiv x \wedge ST_y^{\langle x, \perp \rangle}(\perp)) \wedge \\
&\quad \forall z(Ryz \wedge \neg(y \equiv x \wedge ST_z^{\langle x, \perp \rangle}(\perp)) \wedge \\
&\quad \neg(y \equiv y \wedge \exists z'(Rzz' \wedge \neg(z \equiv x \wedge ST_{z'}^{\langle x, \perp \rangle}(\perp)) \wedge \\
&\quad ST_{z'}^{\langle x, \perp \rangle}(p_1)) \rightarrow ST_z^{\langle x, \perp \rangle; \langle y, \Diamond p_1 \rangle}(p_2)) \\
&= \exists y(Rxy \wedge \neg(x \equiv x \wedge \neg y \equiv y) \wedge \forall z(Ryz \wedge \\
&\quad \neg(y \equiv x \wedge \neg z \equiv z) \wedge \neg(y \equiv y \wedge \exists z'(Rzz' \wedge \\
&\quad \neg(z \equiv x \wedge \neg z' \equiv z') \wedge P_1 z') \rightarrow P_2 z))
\end{aligned}$$

The result is much complicated. Actually, it is equivalent to formula $\exists y(Rxy \wedge \forall z(Ryz \wedge \neg \exists z'(Rzz' \wedge P_1 z') \rightarrow P_2 z))$, which states that there exists a successor y of the current point x such that, for each successor z of y , if z does not has any P_1 -successors, then z is P_2 . Example 2 can be considered as a small case illustrating that S_dML is succinct notation for a complex part of first-order logic. In order to check the result, we now prove the correctness of Definition 3.

Theorem 1 (Correctness of the Standard Translation). *Let $\langle \mathcal{M}, w \rangle$ be a pointed model and φ be a formula of \mathcal{L}_d , then*

$$\mathcal{M}, w \models \varphi \text{ iff } \mathcal{M} \models ST_x^{\langle x, \perp \rangle}(\varphi)[w].$$

Proof. The proof is by induction on the structure of φ . The cases for Boolean and modal connectives are straightforward. When φ is $[-\varphi_1]\varphi_2$, the following equivalences hold:

$$\begin{aligned}
\mathcal{M}, w \models [-\varphi_1]\varphi_2 &\text{ iff } \exists \mathcal{M}' \text{ s.t. } \langle \mathcal{M}, w \rangle \xrightarrow{-\varphi_1} \langle \mathcal{M}', w \rangle \text{ and } \mathcal{M}', w \models \varphi_2 \\
&\text{ iff } \exists \mathcal{M}' \text{ s.t. } \langle \mathcal{M}, w \rangle \xrightarrow{-\varphi_1} \langle \mathcal{M}', w \rangle \text{ and } \mathcal{M}' \models ST_x^{\langle x, \perp \rangle}(\varphi_2)[w] \\
&\text{ iff } \mathcal{M} \models ST_x^{\langle x, \perp \rangle; \langle x, \varphi_1 \rangle}(\varphi_2)[w] \\
&\text{ iff } \mathcal{M} \models ST_x^{\langle x, \perp \rangle}(\varphi)[w]
\end{aligned}$$

The first equivalence follows from the semantics directly. By the inductive hypothesis, the second one holds. The last two equivalences hold by Definition 3. \square

Remark 1. *The first-order translation for S_dML is quite different from that for SML . To translate a SML formula, it suffices to maintain a finite set of ordered pairs of nodes encoding the links already deleted (cf. [5]). However it fails for S_dML , since the number of links cut by $[-]$ may be infinite. Besides, Example 1 shows that we should also take care of the order of $[-]$ in a formula. Our finite sequence of ordered pairs of nodes and properties solves these problems and yields a translation for S_dML .*

Finally, we end by answering the question stated at the outset of this section, which follows directly from Definition 3 and Theorem 1:

Corollary 1. *By the completeness theorem for first-order logic, S_dML is axiomatizable.*

4 Bisimulation and Expressivity for S_dML

Through the standard translation, we can translate a formula of S_dML into first-order logic syntactically. In this section, we investigate the other aspect, i.e., model theories, for its expressive power. Let us begin with considering the notion of bisimulation for S_dML .

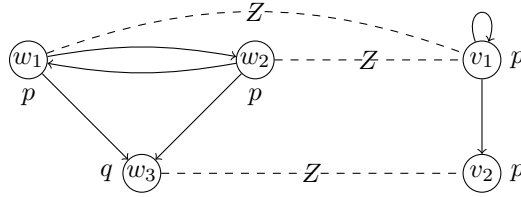
4.1 Bisimulation for S_dML

After expanding the standard modal language \mathcal{L}_\square with the operator $[-]$, formulas of \mathcal{L}_d are not invariant under the standard bisimulation any longer (cf. [14]).

To show this, we first introduce a notion of *definable sabotage modal equivalence* (notation, \rightsquigarrow_d) between pointed models: $\langle \mathcal{M}_1, w \rangle \rightsquigarrow_d \langle \mathcal{M}_2, v \rangle$ iff for each $\varphi \in \mathcal{L}_d$, $\mathcal{M}_1, w \models \varphi$ iff $\mathcal{M}_2, v \models \varphi$.

Fact 1. *Formulas of \mathcal{L}_d are not invariant under the standard bisimulation.*

Proof. It suffices to give an example. Consider two models \mathcal{M}_1 and \mathcal{M}_2 that are defined as depicted in the following figure:



By the definition of standard bisimulation, we know that both $\langle \mathcal{M}_1, w_1 \rangle$ and $\langle \mathcal{M}_1, w_2 \rangle$ are bisimilar to $\langle \mathcal{M}_2, v_1 \rangle$, and that $\langle \mathcal{M}_1, w_3 \rangle$ is bisimilar to $\langle \mathcal{M}_2, v_2 \rangle$. However, we have $\mathcal{M}_1, w_1 \models [-q] \Diamond \Diamond q$ and $\mathcal{M}_2, v_1 \not\models [-q] \Diamond \Diamond q$. Therefore bisimulation does not imply definable sabotage modal equivalence. \square

What is a suitable notion of bisimulation for S_dML ? Now we introduce a new notion of *definable sabotage bisimulation* (*d-bisimulation*). Here is the formal definition.

Definition 4 (d-bisimulation). Let $\mathcal{M}_1 = \langle W_1, R_1, V_1 \rangle$ and $\mathcal{M}_2 = \langle W_2, R_2, V_2 \rangle$ be two models. A non-empty relation Z_d is a d-bisimulation between pointed models $\langle \mathcal{M}_1, w \rangle$ and $\langle \mathcal{M}_2, v \rangle$ (notation, $Z_d : \langle \mathcal{M}_1, w \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v \rangle$) if the following five conditions are satisfied:

Atom: If $\langle \mathcal{M}_1, w \rangle Z_d \langle \mathcal{M}_2, v \rangle$, then $\mathcal{M}_1, w \models p$ iff $\mathcal{M}_2, v \models p$, for each $p \in \mathbf{P}$.

Zig $_{\diamond}$: If $\langle \mathcal{M}_1, w \rangle Z_d \langle \mathcal{M}_2, v \rangle$ and there exists $w' \in W_1$ such that $R_1 ww'$, then there exists $v' \in W_2$ such that $R_2 vv'$ and $\langle \mathcal{M}_1, w' \rangle Z_d \langle \mathcal{M}_2, v' \rangle$.

Zag $_{\diamond}$: If $\langle \mathcal{M}_1, w \rangle Z_d \langle \mathcal{M}_2, v \rangle$ and there exists $v' \in W_2$ such that $R_2 vv'$, then there exists $w' \in W_1$ such that $R_1 ww'$ and $\langle \mathcal{M}_1, w' \rangle Z_d \langle \mathcal{M}_2, v' \rangle$.

Zig $_{[-]}$: For each $\varphi \in \mathcal{L}_d$, if $\langle \mathcal{M}_1, w \rangle Z_d \langle \mathcal{M}_2, v \rangle$ and there exists \mathcal{M}'_1 such that $\langle \mathcal{M}_1, w \rangle \xrightarrow{-\varphi} \langle \mathcal{M}'_1, w \rangle$, then there exists \mathcal{M}'_2 such that $\langle \mathcal{M}_2, v \rangle \xrightarrow{-\varphi} \langle \mathcal{M}'_2, v \rangle$ and $\langle \mathcal{M}'_1, w \rangle Z_d \langle \mathcal{M}'_2, v \rangle$.

Zag $_{[-]}$: For each $\varphi \in \mathcal{L}_d$, if $\langle \mathcal{M}_1, w \rangle Z_d \langle \mathcal{M}_2, v \rangle$ and there exists \mathcal{M}'_2 such that $\langle \mathcal{M}_2, v \rangle \xrightarrow{-\varphi} \langle \mathcal{M}'_2, v \rangle$, then there exists \mathcal{M}'_1 such that $\langle \mathcal{M}_1, w \rangle \xrightarrow{-\varphi} \langle \mathcal{M}'_1, w \rangle$ and $\langle \mathcal{M}'_1, w \rangle Z_d \langle \mathcal{M}'_2, v \rangle$.

For brevity, we write $\langle \mathcal{M}_1, w \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v \rangle$ if there exists a d-bisimulation Z_d such that $\langle \mathcal{M}_1, w \rangle Z_d \langle \mathcal{M}_2, v \rangle$.

Here the conditions for \diamond are as usual, and they do not change the model but change the evaluation point along the accessibility relation. While, the conditions for $[-]$ keep the evaluation point fixed but remove some links from the model. In the standard modal logic, given any two models \mathcal{M} and \mathcal{N} , there always exists a bisimulation called *largest bisimulation*, i.e., the set-theoretic union of all bisimulation relations between \mathcal{M} and \mathcal{N} (see [8]). By Definition 4, it is not hard to see that this also holds for the new notion: for any two models, there is a largest d-bisimulation between them. This result is useful in various aspects, say, it can help us to simplify given models to smaller equivalent ones.

As a concrete illustration of the notion introduced here, it is easy to see that the pointed models $\langle \mathcal{M}_1, w_1 \rangle$ and $\langle \mathcal{M}_2, v_1 \rangle$ in the proof of Fact 1 are not d-bisimilar.

Next we show that formulas of $S_d\text{ML}$ are invariant for d-bisimulation:

Theorem 2 ($\xleftrightarrow{d} \subseteq \rightsquigarrow_d$). For any $\langle \mathcal{M}_1, w \rangle$ and $\langle \mathcal{M}_2, v \rangle$, if $\langle \mathcal{M}_1, w \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v \rangle$, then $\langle \mathcal{M}_1, w \rangle \rightsquigarrow_d \langle \mathcal{M}_2, v \rangle$.

Proof. We prove it by induction on the syntax of φ . Let $\langle \mathcal{M}_1, w \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v \rangle$.

(1). $\varphi \in \mathbf{P}$. By Definition 4, it holds directly that $\mathcal{M}_1, w \models \varphi$ iff $\mathcal{M}_2, v \models \varphi$.

(2). φ is $\neg\psi$. By the inductive hypothesis, $\mathcal{M}_1, w \models \psi$ iff $\mathcal{M}_2, v \models \psi$. Consequently, we know that $\mathcal{M}_1, w \models \varphi$ iff $\mathcal{M}_2, v \models \varphi$.

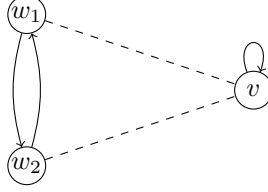
(3). φ is $\varphi_1 \wedge \varphi_2$. By the inductive hypothesis, for each $i \in \{1, 2\}$, $\mathcal{M}_1, w \models \varphi_i$ iff $\mathcal{M}_2, v \models \varphi_i$. Thus it holds that $\mathcal{M}_1, w \models \varphi$ iff $\mathcal{M}_2, v \models \varphi$.

(4). φ is $\diamond\psi$. If $\mathcal{M}_1, w \models \varphi$, then there exists $w_1 \in W_1$ such that $R_1 ww_1$ and $\mathcal{M}_1, w_1 \models \psi$. By **Zig $_{\diamond}$** , there exists $v_1 \in W_2$ s.t. $R_2 vv_1$ and $\langle \mathcal{M}_1, w_1 \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v_1 \rangle$. By the inductive hypothesis, $\mathcal{M}_1, w_1 \models \psi$ iff $\mathcal{M}_2, v_1 \models \psi$. It is followed by $\mathcal{M}_2, v_1 \models \psi$ immediately. Consequently it holds that $\mathcal{M}_2, v \models \varphi$. Similarly, we can obtain $\mathcal{M}_1, w \models \varphi$ from $\mathcal{M}_2, v \models \varphi$ by **Zag $_{\diamond}$** .

(5). φ is $[-\varphi_1]\varphi_2$. If $\mathcal{M}_1, w \models \varphi$, then there is a \mathcal{M}'_1 s.t. $\langle \mathcal{M}_1, w \rangle \xrightarrow{-\varphi_1} \langle \mathcal{M}'_1, w \rangle$ and $\mathcal{M}'_1, w \models \varphi_2$. By **Zig** $[-]$, there is some \mathcal{M}'_2 such that $\langle \mathcal{M}_2, v \rangle \xrightarrow{-\varphi_1} \langle \mathcal{M}'_2, v \rangle$ and $\langle \mathcal{M}'_1, w \rangle \xleftrightarrow{d} \langle \mathcal{M}'_2, v \rangle$. By the inductive hypothesis, $\mathcal{M}'_1, w \models \varphi_2$ iff $\mathcal{M}'_2, v \models \varphi_2$. Hence it holds that $\mathcal{M}_2, v \models \varphi$. Similarly, by **Zag** $[-]$, $\mathcal{M}_1, w \models \varphi$ follows from $\mathcal{M}_2, v \models \varphi$. \square

As an application of Theorem 2, let us consider a simple example:

Example 3. Consider two models \mathcal{M}_1 and \mathcal{M}_2 defined respectively as follows:



By Definition 4, it holds that $\langle \mathcal{M}_1, w_1 \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v \rangle$ and $\langle \mathcal{M}_1, w_2 \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v \rangle$ (the d -bisimulation runs via the dashed lines). From Theorem 2, we know that $\langle \mathcal{M}_1, w_1 \rangle \rightsquigarrow_d \langle \mathcal{M}_2, v \rangle$ and $\langle \mathcal{M}_1, w_2 \rangle \rightsquigarrow_d \langle \mathcal{M}_2, v \rangle$. Therefore, S_dML cannot distinguish between nodes $w_{1(2)}$ and v .

Furthermore, for ω -saturated models, the converse of Theorem 2 holds as well. For each finite set Y , we denote the expansion of \mathcal{L}_1 with a set Y of constants with \mathcal{L}_1^Y , and denote the expansion of \mathcal{M} to \mathcal{L}_1^Y with \mathcal{M}^Y .

Definition 5 (ω -saturation). A model $\mathcal{M} = \langle W, R, V \rangle$ is ω -saturated if, for every finite subset Y of W , the expansion \mathcal{M}^Y realizes every set $\Gamma(x)$ of \mathcal{L}_1^Y -formulas whose finite subsets $\Gamma'(x)$ are all realized in \mathcal{M}^Y .

Not all models are ω -saturated, but every model can be extended to an ω -saturated model with the same first-order theory (see [16]). From Definition 3, we know that each model \mathcal{M} has an ω -saturated extension with the same theory of S_dML . For brevity, we use the set $\mathbb{T}^d(\mathcal{M}, w) = \{\varphi \in \mathcal{L}_d \mid \mathcal{M}, w \models \varphi\}$ of \mathcal{L}_d -formulas to denote the theory of w in \mathcal{M} . By Definition 5, we have the following result.

Theorem 3 ($\rightsquigarrow_d \subseteq \xleftrightarrow{d}$). For any two ω -saturated pointed models $\langle \mathcal{M}_1, w \rangle$ and $\langle \mathcal{M}_2, v \rangle$, if $\langle \mathcal{M}_1, w \rangle \rightsquigarrow_d \langle \mathcal{M}_2, v \rangle$, then $\langle \mathcal{M}_1, w \rangle \xleftrightarrow{d} \langle \mathcal{M}_2, v \rangle$.

Proof. We prove this by showing that \rightsquigarrow_d satisfies the definition of d -bisimulation.

(1). For each $p \in \mathbf{P}$, by the definition of \rightsquigarrow_d , it holds that $\mathcal{M}_1, w \models p$ iff $\mathcal{M}_2, v \models p$. This satisfies the condition of **Atom**.

(2). Let $w_1 \in W_1$ such that $R_1 w w_1$. We show that point v has a successor v_1 with $\langle \mathcal{M}_1, w_1 \rangle \rightsquigarrow_d \langle \mathcal{M}_2, v_1 \rangle$. For each finite subset Γ of $\mathbb{T}^d(\mathcal{M}_1, w_1)$, it holds that:

$$\begin{aligned} \mathcal{M}_1, w \models \bigwedge \Gamma &\text{ iff } \mathcal{M}_2, v \models \bigwedge \Gamma \\ &\text{ iff } \mathcal{M}_2 \models ST_x^{(x, \perp)}(\bigwedge \Gamma)[v] \\ &\text{ iff } \mathcal{M}_2 \models \exists y(Rxy \wedge ST_y^{(x, \perp)}(\bigwedge \Gamma))[v] \end{aligned}$$

Therefore every finite subset Γ of $\mathbb{T}^d(\mathcal{M}_1, w_1)$ is satisfiable in the set of successors of node v . From Definition 5, we know that v has a successor v_1 where $\mathbb{T}^d(\mathcal{M}_1, w_1)$ is true. Thus, $\langle \mathcal{M}_1, w_1 \rangle \rightsquigarrow_d \langle \mathcal{M}_2, v_1 \rangle$. The proof of the **Zig** $_{\diamond}$ clause is completed.

(3). Similar to (2), we can prove that the condition of **Zag** $_{\diamond}$ is satisfied.

(4). Let $\langle \mathcal{M}'_1, w \rangle$ be a pointed model and $\varphi \in \mathcal{L}_d$ such that $\langle \mathcal{M}_1, w \rangle \xrightarrow{-\varphi} \langle \mathcal{M}'_1, w \rangle$. We prove the **Zig** $_{[-]}$ clause by showing there exists \mathcal{M}'_2 with $\langle \mathcal{M}_2, v \rangle \xrightarrow{-\varphi} \langle \mathcal{M}'_2, v \rangle$ and $\langle \mathcal{M}'_1, w \rangle \rightsquigarrow_d \langle \mathcal{M}'_2, v \rangle$. For each finite subset Γ of $\mathbb{T}^d(\mathcal{M}'_1, w)$, the following sequence of equivalences holds:

$$\begin{aligned} \mathcal{M}_1, w \models [-\varphi] \bigwedge \Gamma &\text{ iff } \mathcal{M}_2, v \models [-\varphi] \bigwedge \Gamma \\ &\text{ iff } \mathcal{M}_2 \models ST_x^{\langle x, \perp \rangle}([- \varphi] \bigwedge \Gamma)[v] \\ &\text{ iff } \mathcal{M}_2 \models ST_x^{\langle x, \perp \rangle; \langle x, \varphi \rangle}(\bigwedge \Gamma)[v] \end{aligned}$$

Hence each finite subset of $\mathbb{T}^d(\mathcal{M}'_1, w)$ is true at $\langle \mathcal{M}'_2, v \rangle$, where $\langle \mathcal{M}_2, v \rangle \xrightarrow{-\varphi} \langle \mathcal{M}'_2, v \rangle$. By Definition 5, $\mathbb{T}^d(\mathcal{M}'_1, w)$ is true at $\langle \mathcal{M}'_2, v \rangle$. It is followed by $\langle \mathcal{M}'_1, w \rangle \rightsquigarrow_d \langle \mathcal{M}'_2, v \rangle$.

(5). Similar to (4), we can show that the condition of **Zag** $_{[-]}$ is satisfied.

Thus, we conclude that $\langle \mathcal{M}_1, w \rangle \xleftrightarrow{\pm}_d \langle \mathcal{M}_2, v \rangle$. The proof is completed. \square

4.2 Characterization of $S_d\text{ML}$

By the notion of d-bisimulation, we can characterize $S_d\text{ML}$ as the one-free-variable fragment of FOL that is invariant for d-bisimulation, where a first-order formula $\alpha(x)$ is invariant for d-bisimulation means that for all pointed models $\langle \mathcal{M}_1, w_1 \rangle$ and $\langle \mathcal{M}_2, w_2 \rangle$ such that $\langle \mathcal{M}_1, w_1 \rangle \xleftrightarrow{\pm}_d \langle \mathcal{M}_2, w_2 \rangle$, it holds that $\mathcal{M}_1 \models \alpha(x)[w_1]$ iff $\mathcal{M}_2 \models \alpha(x)[w_2]$.

Theorem 4 (Characterization of $S_d\text{ML}$ by d-bisimulation Invariance). *An \mathcal{L}_1 -formula is equivalent to the translation of an \mathcal{L}_d -formula iff it is invariant for d-bisimulation.*

Proof. The direction from left to right holds directly by Theorem 2. For the converse direction, let α be an \mathcal{L}_1 -formula with one free variable x . Assume that α is invariant for d-bisimulation. Now we consider the following set:

$$\mathbb{C}_d(\alpha) = \{ST_x^{\langle x, \perp \rangle}(\varphi) \mid \varphi \in \mathcal{L}_d \text{ and } \alpha \models ST_x^{\langle x, \perp \rangle}(\varphi)\}.$$

The result holds from the following two claims:

- (i). If $\mathbb{C}_d(\alpha) \models \alpha$, then α is equivalent to the translation of an \mathcal{L}_d -formula.
- (ii). $\mathbb{C}_d(\alpha) \models \alpha$, i.e., for any pointed model $\langle \mathcal{M}, w \rangle$, $\mathcal{M} \models \mathbb{C}_d(\alpha)[w]$ entails $\mathcal{M} \models \alpha[w]$.

We show (i) first. Suppose that $\mathbb{C}_d(\alpha) \models \alpha$. From the compactness and deduction theorems of first-order logic, it holds that $\models \bigwedge \Gamma \rightarrow \alpha$ for some finite subset Γ of $\mathbb{C}_d(\alpha)$. The converse can be shown by the definition of $\mathbb{C}_d(\alpha)$: $\models \alpha \rightarrow \bigwedge \Gamma$. Thus it holds that $\models \alpha \leftrightarrow \bigwedge \Gamma$ proving the claim.

As to the claim (ii), let $\langle \mathcal{M}, w \rangle$ be a pointed model such that $\mathcal{M} \models \mathbb{C}_d(\alpha)[w]$. Consider the set $\Sigma = ST_x^{\langle x, \perp \rangle}(\mathbb{T}^d(\mathcal{M}, w)) \cup \{\alpha\}$. We now show that:

- (a). The set Σ is consistent.
- (b). $\mathcal{M} \models \alpha[w]$, thus proving claim (ii).

Suppose that Σ is not consistent. By the compactness of first-order logic, it follows that $\models \alpha \rightarrow \neg \bigwedge \Gamma$ for some finite subset Γ of Σ . But then, by the definition of $\mathbb{C}_d(\alpha)$, we obtain $\neg \bigwedge \Gamma \in \mathbb{C}_d(\alpha)$, which is followed by $\neg \bigwedge \Gamma \in ST_x^{(x, \perp)}(\mathbb{T}^d(\mathcal{M}, w))$. However, it contradicts to $\Gamma \subseteq ST_x^{(x, \perp)}(\mathbb{T}^d(\mathcal{M}, w))$. Hence (a) holds.

Now we show that (b) holds as well. Since Σ is consistent, it can be realized by some pointed model, say, $\langle \mathcal{M}', w' \rangle$. Note that both the pointed models have same theories, thus $\langle \mathcal{M}, w \rangle \rightsquigarrow_d \langle \mathcal{M}', w' \rangle$. Now take two ω -saturated elementary extensions $\langle \mathcal{M}_\omega, w \rangle$ and $\langle \mathcal{M}'_\omega, w' \rangle$ of $\langle \mathcal{M}, w \rangle$ and $\langle \mathcal{M}', w' \rangle$ respectively. It can be shown that such extensions always exist (see [16]). By the invariance of first-order logic under elementary extensions, from $\mathcal{M}' \models \alpha[w']$ we know $\mathcal{M}'_\omega \models \alpha[w']$. Moreover, by Theorem 3 and the assumption that α is invariant for d-bisimulation, we have $\mathcal{M}_\omega \models \alpha[w]$. By the elementary extension, we obtain $\mathcal{M} \models \alpha[w]$ that entails the claim (ii). Consequently, the proof is completed. \square

Just as with SML, the key model-theoretic argument using saturation needed special care, but now with new modifications matching the above translation of $S_d\text{ML}$ (cf. [5]).

4.3 Exploring Expressive Power

So far, we have already been able to show whether or not a first-order property belongs to the fragment identified by Theorem 4. In this section, we show several concrete examples, which will also present a comparison between $S_d\text{ML}$ and SML with respect to their expressive power on models.

Example 4. Consider the first-order property $\alpha_1(x)$ ‘The current point is irreflexive and not a dead end. Each of its successors only has access to it’, i.e., $\alpha_1(x) := \neg Rxx \wedge \exists y Rxy \wedge \forall y (Rxy \rightarrow Ryx \wedge \forall z (Ryz \rightarrow z \equiv x))$. From Example 3, we know that this property is not invariant for d-bisimulation. For instance, formula $\alpha_1(x)$ is true at state w_1 in \mathcal{M}_1 but fails at v in \mathcal{M}_2 . Thus this property is not definable in $S_d\text{ML}$.

Interestingly, the result may be quite different if we change the first-order property in Example 4 slightly, say,

Fact 2. The first-order property $\alpha_1^+(x)$ ‘The current point is irreflexive and not a dead end. Some of its successors are dead ends, the others only have access to dead ends and the current point’, i.e., $\alpha_1^+(x) := \neg Rxx \wedge \exists y (Rxy \wedge \neg \exists z Ryz) \wedge \exists y (Rxy \wedge \exists z Ryz) \wedge \forall y (Rxy \rightarrow \neg \exists z Ryz \vee (Ryx \wedge \exists z (Ryz \wedge \neg \exists u Rzu)) \wedge \forall z (Ryz \rightarrow z \equiv x \vee \neg \exists u Rzu))$, is definable in $S_d\text{ML}$.

Proof. Consider the following formulas of $S_d\text{ML}$:

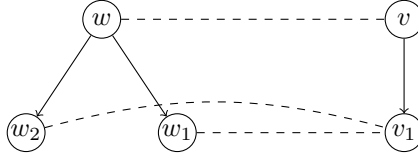
- (B₁) $\Diamond \Box \perp \wedge \Diamond \Diamond \top$
- (B₂) $\Box (\Diamond \top \rightarrow \Diamond \Box \perp \wedge \Diamond (\Diamond \Box \perp \wedge \Diamond \Diamond \top)) \wedge \Box (\Box \perp \vee (\Diamond \Box \perp \wedge \Diamond \Diamond \top))$
- (B₃) $[\neg \Box \perp] \Box (\Diamond \Box \perp \wedge \Box (\neg \Box \perp \rightarrow \neg \Diamond \Box \perp))$

Let $\varphi_1^+ := (B_1 \wedge B_2 \wedge B_3)$. This formula is satisfiable, say, it is true at $\langle \mathcal{M}_1, w_1 \rangle$ in the proof of Fact 1. Let $\langle \mathcal{M}, u \rangle$ be a pointed model. It is not hard to see that $\mathcal{M}, u \models \varphi_1^+$ if $\mathcal{M} \models \alpha_1^+(x)[u]$. Now assume that $\mathcal{M}, u \models \varphi_1^+$. Formula (B_1) states that, the current point u has some successors u_1 that are dead ends, and some successors u_2 which have successors. By (B_2) , each u_2 reaches some dead end u_3 , and some point u_4 which is similar to u : it has some successors which are dead ends, and some successors that also have successors. After cutting the links from node u to the dead ends, from (B_3) it holds that u_2 still can see some dead ends, and that u_4 cannot reach dead ends any longer. Therefore we obtain $u_2 \neq u$ and $u_4 = u$, consequently, $\mathcal{M} \models \alpha_1^+(x)[u]$. So we conclude that $\mathcal{M} \models \alpha_1^+(x)[u]$ iff $\mathcal{M}, u \models \varphi_1^+$ for any pointed model $\langle \mathcal{M}, u \rangle$. \square

Through observation, we can find that the property $\alpha_1^+(x)$ expands the current point and its successors in $\alpha_1(x)$ with some successors that are dead ends. But the former one is definable in $S_d\text{ML}$ and the latter one is not. What is the reason for this?

Suppose that $\langle \mathcal{M}, u \rangle$ be a pointed model that is d-bisimilar to $\langle \mathcal{M}_1, w_1 \rangle$ in the proof of Fact 1. By Definition 4, we know that u can reach some dead end u_1 , and some u_2 that has access to some dead ends. Except those dead ends, u_2 can also see some point u_3 that is similar to u : u_3 can reach some dead end and some node that has successors. Further more, after cutting the links from u to the dead ends, u_2 still can see some dead ends, but u_3 cannot reach any dead ends now. So we have $u_2 \neq u$ and $u_3 = u$. In such a way, we conclude that the property $\alpha_1^+(x)$ is invariant under d-bisimulation.

Example 5. Consider the FOL property ‘There exist n successors of the current point’. This property is not invariant for d-bisimulation. For instance, in the following models:



the property ‘there exist 2 successors’ is true at point w in the model to the left, but it fails at v to the right. Hence it is not definable in $S_d\text{ML}$.

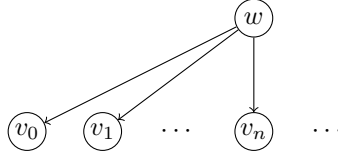
In contrast, as noted in [5], SML can count successors of the current state, and it can also define the length of a cycle. That is, for each positive natural number n , there exists a SML formula φ such that, for any $\mathcal{M} = \langle W, R, V \rangle$ and $w \in W$, $\mathcal{M}, w \models \varphi$ iff $\langle W, R \rangle$ is a cycle of length n . Is this property definable in $S_d\text{ML}$?

Example 6. Recall the two models displayed in Example 3. The underlying frame of \mathcal{M}_1 is a cycle of length 2, while that of \mathcal{M}_2 is a cycle of length 1. So $S_d\text{ML}$ cannot define the length of a cycle.

Intuitively, these differences between $S_d\text{ML}$ and SML stem from the features of $[-]$ and the standard sabotage modality \blacklozenge . In SML, each occurrence of \blacklozenge in a formula deletes exactly one link. While, in $S_d\text{ML}$, $[-]$ operates uniformly, which blocks the logic to define the first-order properties in Example 5-6. However, this does not mean that $S_d\text{ML}$ is less expressive than SML with respect to models. Actually the notion of bisimulation

for sabotage modal logic is not an extension of d-bisimulation. When tackling with cases involving infinite, operator $[-]$ may show more strength. Here is an example.

Example 7. *We establish a model in the following way:*



As we can see, node w has countable successors. By the truth condition for $[-]$, formula $[-\top]\Box\perp$ is true at node w , which says that all links starting at w are cut by the operator $[-]$. However there is no such a formula φ of SML that can do this: each occurrence of \blacklozenge cuts one link, but the number of \blacklozenge occurring in a formula is always finite.

From Examples 5-7, we then get the following conclusion.

Fact 3. *S_dML and SML are not comparable in their expressive power on models.*

5 From S_dML to Hybrid Logics

While an effective first-order translation shows that validity in S_dML is effectively axiomatizable, it gives no concrete information about a more ‘modal’ complete set of proof principles. In this section, following the techniques developed by some dynamic-epistemic logics (cf. e.g. [6]), we try to axiomatize S_dML by means of recursion axioms.

The principles for Boolean cases are as usual. However, as for $[-\varphi]\Box\psi$, there is a problem. From the typical method of recursion axioms used in dynamic-epistemic logic, we know that dynamic operators can be pushed inside through standard modalities. But it fails for S_dML , since that after pushing $[-]$ under a standard modality over successors of the current world, the model change is not local in the successors any longer and it takes place somewhere else (cf. [5]).

Hence the principle for $[-\varphi]\Box\psi$ should illustrate the position where the change happens. To do so, a natural method is to seek help from hybrid logics, which enable us to name nodes in a model. Consider the hybrid logic with *nominals*, *at operator* $@$ and *down-arrow operator* \downarrow , which is denoted by $\mathcal{H}(\downarrow)$. With its formulas of the form $\downarrow x\Box\downarrow y\varphi$, we can manipulate links by naming pairs of points (see [2]).

5.1 S_dML and Hybrid Logics

As a warm-up, we briefly discuss the relation between S_dML and hybrid logics. In particular, the following translation illustrates that S_dML can be reduced to $\mathcal{H}(\downarrow)$. Similar to the standard translation, a finite sequence O will be used.

Definition 6 (The Hybrid Translation for S_dML). *Let O be a finite sequence of pairs of variables of nominals and properties, denoted with $\langle x_0, \psi_0 \rangle; \dots; \langle x_i, \psi_i \rangle; \dots; \langle x_n, \psi_n \rangle$ ($0 \leq i \leq n$). The translation $T^O : \mathcal{L}_d \rightarrow \mathcal{H}(\downarrow)$ is recursively defined in the following way:*

$$T^O(p) = p$$

$$\begin{aligned}
 T^O(\top) &= \top \\
 T^O(\neg\varphi) &= \neg T^O(\varphi) \\
 T^O(\varphi_1 \wedge \varphi_2) &= T^O(\varphi_1) \wedge T^O(\varphi_2) \\
 T^O(\Diamond\varphi) &= \downarrow x \Diamond (\neg (@_x x_0 \wedge T^{\langle x_0, \perp \rangle}(\psi_0)) \wedge \\
 &\quad \bigwedge_{0 \leq i \leq n-1} \neg (@_x x_{i+1} \wedge T^{\langle x_0, \psi_0 \rangle; \dots; \langle x_i, \psi_i \rangle}(\psi_{i+1})) \wedge T^O(\varphi)) \\
 T^O([\neg\psi]\varphi) &= \downarrow x T^{O; \langle x, \psi \rangle}(\varphi)
 \end{aligned}$$

In fact, the truth value of a $\mathcal{H}(\downarrow)$ -formula in some model may depend on the valuation of nominals occurring in it. However, this is not problematic: by Definition 6, for each $\varphi \in \mathcal{L}_d$, $T^{\langle x, \perp \rangle}(\varphi)$ yields a $\mathcal{H}(\downarrow)$ -formula with no free variables of nominals. For brevity, we will leave out the assignment of values to variables in models of $\mathcal{H}(\downarrow)$ if there is no ambiguity. Now we show the correctness of Definition 6.

Theorem 5 (Correctness of the Hybrid Translation). *Let $\langle \mathcal{M}, w \rangle$ be a pointed model and φ be a formula of \mathcal{L}_d , then*

$$\mathcal{M}, w \models \varphi \text{ iff } \mathcal{M}, w \models T^{\langle x, \perp \rangle}(\varphi).$$

Proof. The proof is by induction on the structure of φ . The Boolean cases are straightforward, and we only show the non-trivial cases.

(1). When φ is $\Diamond\psi$, the following equivalences hold:

$$\begin{aligned}
 \mathcal{M}, w \models \varphi &\text{ iff there exists } v \in W \text{ s.t. } R w v \text{ and } \mathcal{M}, v \models \psi \\
 &\text{ iff there exists } v \in W \text{ s.t. } R w v \text{ and } \mathcal{M}, v \models T^{\langle x, \perp \rangle}(\psi) \\
 &\text{ iff } \mathcal{M}, w \models \Diamond T^{\langle x, \perp \rangle}(\psi) \\
 &\text{ iff } \mathcal{M}, w \models \downarrow x \Diamond (\neg (@_x x \wedge T^{\langle x, \perp \rangle}(\perp)) \wedge T^{\langle x, \perp \rangle}(\psi)) \\
 &\text{ iff } \mathcal{M}, w \models T^{\langle x, \perp \rangle}(\varphi)
 \end{aligned}$$

The first equivalence holds by the semantics of \mathcal{L}_d . The second one follows from the inductive hypothesis. The third and fourth equivalences follow by the semantics of $\mathcal{H}(\downarrow)$. The last one holds by Definition 6.

(2). When φ is $[-\varphi_1]\varphi_2$, we have the following equivalences:

$$\begin{aligned}
 \mathcal{M}, w \models [-\varphi_1]\varphi_2 &\text{ iff } \exists \mathcal{M}' \text{ s.t. } \langle \mathcal{M}, w \rangle \xrightarrow{-\varphi_1} \langle \mathcal{M}', w \rangle \text{ and } \mathcal{M}', w \models \varphi_2 \\
 &\text{ iff } \exists \mathcal{M}' \text{ s.t. } \langle \mathcal{M}, w \rangle \xrightarrow{-\varphi_1} \langle \mathcal{M}', w \rangle \text{ and } \langle \mathcal{M}', w \rangle \models T^{\langle x, \perp \rangle}(\varphi_2) \\
 &\text{ iff } \langle \mathcal{M}, w \rangle \models T^{\langle x, \perp \rangle; \langle x, \varphi_1 \rangle}(\varphi_2) \\
 &\text{ iff } \langle \mathcal{M}, w \rangle \models T^{\langle x, \perp \rangle}(\varphi)
 \end{aligned}$$

The first equivalence follows directly from the semantics of \mathcal{L}_d . The second one holds by the inductive hypothesis. The last two equivalences follow by Definition 6.

Therefore, for each $\varphi \in \mathcal{L}_d$, it holds that $\mathcal{M}, w \models \varphi$ iff $\mathcal{M}, w \models T^{\langle x, \perp \rangle}(\varphi)$. \square

In the way described, we can reduce $S_d\text{ML}$ to $\mathcal{H}(\downarrow)$. But, does the converse direction hold? First note that the following property is definable in $\mathcal{H}(\downarrow)$:

Fact 4. *The property ‘there exist n successors of the current point’ is definable in $\mathcal{H}(\downarrow)$.*

Proof. We prove it by building the desired formula. Let n be a positive natural number. Consider the following $\mathcal{H}(\downarrow)$ -formula:

$$\downarrow x(\diamond \downarrow x_1(@_x \diamond \downarrow x_2(...(@_x \diamond \downarrow x_n(@_x \Box (\bigvee_{0 \leq i \leq n} x_i \wedge \bigwedge_{1 \leq i < j \leq n} \neg @_x x_i x_j) \underbrace{\dots}_{n+2})))$$

The formula states that the current point x has successors x_1, \dots, x_n , that each node reachable from x must be some x_i , where $1 \leq i \leq n$, and that for any different i and j such that $1 \leq i, j \leq n$, x_i is distinct from x_j . Thus, there exist n successors of the current point iff the stated hybrid formula holds at that point. \square

But Example 5 showed that this property is not definable in $S_d\text{ML}$.

Fact 5. *$\mathcal{H}(\downarrow)$ is more expressive than $S_d\text{ML}$ over models.*

Therefore $S_d\text{ML}$ can be viewed as a fragment of $\mathcal{H}(\downarrow)$. Any hybrid logic at least as expressive as $\mathcal{H}(\downarrow)$ is more expressive than $S_d\text{ML}$. Even so, the hybrid translation described in Definition 6 suggests that it may be viable to analyze validity in the logic $S_d\text{ML}$ with expressive resources similar to those of $\mathcal{H}(\downarrow)$.

5.2 Digression on Recursion Axioms

One attractive format for axiomatizing logics of model change are recursion axioms in the style of dynamic-epistemic logic (see [9]). As mentioned already, Boolean cases are available for $S_d\text{ML}$ as well. We begin with the principle for $[-]$:¹

Fact 6. *Let φ, ψ and χ be \mathcal{L}_d -formulas. Then it holds that*

$$[-\varphi][-\psi]\chi \leftrightarrow \downarrow x[-\downarrow y(\varphi \vee @_x[-\varphi]@_y\psi)]\chi \quad (6)$$

where x and y are new nominal variables.

Proof. Let $\langle \mathcal{M}, w \rangle$ be a pointed model. We prove it by showing that $\mathcal{M}|_{\langle w, \varphi \rangle | \langle w, \psi \rangle}$ and $\mathcal{M}|_{\langle w, \downarrow y(\varphi \vee @_x[-\varphi]@_y\psi) \rangle}$ are same, where $w \in V(x)$. Suppose not, then there must be some $v \in W$ such that $\langle w, v \rangle \in \mathcal{M}|_{\langle w, \varphi \rangle | \langle w, \psi \rangle}$ and $\langle w, v \rangle \notin \mathcal{M}|_{\langle w, \downarrow y(\varphi \vee @_x[-\varphi]@_y\psi) \rangle}$, or that $\langle w, v \rangle \in \mathcal{M}|_{\langle w, \downarrow y(\varphi \vee @_x[-\varphi]@_y\psi) \rangle}$ and $\langle w, v \rangle \notin \mathcal{M}|_{\langle w, \varphi \rangle | \langle w, \psi \rangle}$.

Now consider the first case. From $\langle w, v \rangle \notin \mathcal{M}|_{\langle w, \downarrow y(\varphi \vee @_x[-\varphi]@_y\psi) \rangle}$, we know that $\mathcal{M}, v \models \varphi \vee @_x[-\varphi]@_y\psi$ where $v \in V(y)$. By $\langle w, v \rangle \in \mathcal{M}|_{\langle w, \varphi \rangle | \langle w, \psi \rangle}$, it follows that $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \psi$. Since $\mathcal{M}|_{\langle w, \varphi \rangle | \langle w, \psi \rangle}$ is a submodel of $\mathcal{M}|_{\langle w, \varphi \rangle}$, we obtain $\langle w, v \rangle \in \mathcal{M}|_{\langle w, \varphi \rangle}$. Consequently, it holds that $\mathcal{M}, v \not\models \varphi$, thus, $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \psi$. So we have arrived at a contradiction.

Next we consider the second case. By $\langle w, v \rangle \in \mathcal{M}|_{\langle w, \downarrow y(\varphi \vee @_x[-\varphi]@_y\psi) \rangle}$, it holds that $\mathcal{M}, v \models \neg \varphi \wedge @_x[-\varphi]@_y\neg \psi$ where $v \in V(y)$. Then we know $\langle w, v \rangle \in \mathcal{M}|_{\langle w, \varphi \rangle}$. Besides, by $\langle w, v \rangle \notin \mathcal{M}|_{\langle w, \varphi \rangle | \langle w, \psi \rangle}$, we obtain $\mathcal{M}|_{\langle w, \varphi \rangle}, v \not\models \psi$ that entails a contradiction. \square

¹Actually, the principle for $[-]$ is not necessary to show a complete set of recursion axioms, cf. [10].

Note that some operators of $\mathcal{H}(\downarrow)$ occur in (6). From Definition 6, we know that it is equivalent with some formula of $\mathcal{H}(\downarrow)$. Consider formula $\downarrow x[-\downarrow y(\varphi \vee @_x[-\varphi]@_y\psi)]\chi$. By the semantics, that it is true at a pointed model $\langle \mathcal{M}, w \rangle$ means that w is χ in the model $\mathcal{M}|_{\langle w, \downarrow y(\varphi \vee @_x[-\varphi]@_y\psi) \rangle}$, where $V(x) = \{w\}$. Intuitively, the new model is obtained by removing all links from w to the points that are φ , and to the points which are ψ after removing the links from w to φ -points. This is exactly what $[-\varphi][-\psi]\chi$ states.

We now move to the case for \Box . It seems like that the following result will work:

Fact 7. *For each $[-\varphi]\Box\psi \in \mathcal{L}_d$, the following equivalence holds:*

$$[-\varphi]\Box\psi \leftrightarrow \downarrow x\Box \downarrow y(\neg\varphi \rightarrow @_x[-\varphi]@_y\psi) \quad (7)$$

where x and y are new nominal variables.

Proof. Let $\langle \mathcal{M}, w \rangle$ be a pointed model. For the direction from left to right, we suppose that $\mathcal{M}, w \models [-\varphi]\Box\psi$ and $\mathcal{M}, w \not\models \downarrow x\Box \downarrow y(\neg\varphi \rightarrow @_x[-\varphi]@_y\psi)$. Then it holds that $w \in V(x)$ has a successor $v \in V(y)$ such that $\mathcal{M}, v \models \neg\varphi \wedge @_x[-\varphi]@_y\neg\psi$. From $\mathcal{M}, w \models [-\varphi]\Box\psi$, it follows that $\mathcal{M}|_{\langle w, \varphi \rangle}, w \models \Box\psi$. Since $\mathcal{M}, v \models \neg\varphi$, we obtain $\langle w, v \rangle \in \mathcal{M}|_{\langle w, \varphi \rangle}$. Thus it holds that $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \psi$. Besides, $\mathcal{M}, v \models \neg\varphi \wedge @_x[-\varphi]@_y\neg\psi$ entails $\mathcal{M}, w \models [-\varphi]@_y\neg\psi$. Consequently, it holds that $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \neg\psi$, which entails a contradiction.

For the converse direction, we assume that $\mathcal{M}, w \models \downarrow x\Box \downarrow y(\neg\varphi \rightarrow @_x[-\varphi]@_y\psi)$ and $\mathcal{M}, w \not\models [-\varphi]\Box\psi$. Then there exists $v \in W$ such that $\langle w, v \rangle \in R \setminus (\{w\} \times V(\varphi))$ and $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \neg\psi$. Consider the case where w and v are named as x and y respectively. It holds that $\mathcal{M}|_{\langle w, \varphi \rangle}, w \models @_y\neg\psi$. So we obtain $\mathcal{M}|_{\langle w, \varphi \rangle}, w \models @_x[-\varphi]@_y\neg\psi$. Further more, from $\langle w, v \rangle \in R \setminus (\{w\} \times V(\varphi))$, we know $\langle w, v \rangle \in R$ and $\mathcal{M}, v \models \neg\varphi$. Thus it is conclude that $\mathcal{M}, w \not\models \downarrow x\Box \downarrow y(\neg\varphi \rightarrow @_x[-\varphi]@_y\psi)$. \square

In formula (7), that $\downarrow x\Box \downarrow y(\neg\varphi \rightarrow @_x[-\varphi]@_y\psi)$ is true at $\langle \mathcal{M}, w \rangle$ says that for each point v , if $v \in R(w)$ and v is not φ , then v is ψ after deleting all links from w to the φ -points. However, although formula (7) is valid, it is not the solution: the formula of the form $@_x[-\varphi]@_y\psi$ blocks the recursion format, even though we have that

Fact 8. *For any $p \in \mathbf{P}$, \mathcal{L}_d -formulas φ , ψ and χ , and nominal variable x , the following equivalences hold:*

$$[-\varphi]@_xp \leftrightarrow @_xp \quad (8)$$

$$[-\varphi]@_x\neg\psi \leftrightarrow \neg[-\varphi]@_x\psi \quad (9)$$

$$[-\varphi]@_x(\psi \wedge \chi) \leftrightarrow [-\varphi]@_x\psi \wedge [-\varphi]@_x\chi \quad (10)$$

$$[-\varphi]@_x\Box\psi \leftrightarrow \downarrow y@_x\Box \downarrow z(\neg(\varphi \wedge @_xy) \rightarrow @_y[-\varphi]@_z\psi) \quad (11)$$

where y and z are new nominal variables.

Proof. The validity of (8)-(10) is straightforward. We now consider (11). Let $\langle \mathcal{M}, w \rangle$ be a pointed model. From left to right. Suppose that $\mathcal{M}, w \models [-\varphi]@_x\Box\psi$ and $\mathcal{M}, w \not\models \downarrow y@_x\Box \downarrow z(\neg(\varphi \wedge @_xy) \rightarrow @_y[-\varphi]@_z\psi)$. Let u be a point such that $V(x) = \{u\}$. Then it holds that $\mathcal{M}, u \models \Box \downarrow z(\neg(\varphi \wedge @_xy) \wedge @_y[-\varphi]@_z\neg\psi)$ where $w \in V(y)$. Therefore there

exists some point v such that Ruv , $v \in V(z)$ and $\mathcal{M}, v \models \neg(\varphi \wedge @_x y) \wedge @_y [-\varphi]@_z \neg\psi$. By $\mathcal{M}, v \models \neg(\varphi \wedge @_x y)$, it holds that $\langle u, v \rangle \in \mathcal{M}|_{\langle w, \varphi \rangle}$. From $\mathcal{M}, v \models @_y [-\varphi]@_z \neg\psi$, we obtain $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \neg\psi$, which contradicts to $\mathcal{M}, w \models [-\varphi]@_x \Box\psi$.

From right to left. Suppose that $\mathcal{M}, w \models \downarrow y@_x \Box \downarrow z(\neg(\varphi \wedge @_x y) \rightarrow @_y [-\varphi]@_z \psi)$ and $\mathcal{M}, w \not\models [-\varphi]@_x \Box\psi$. Let u be a point such that $V(x) = \{u\}$. Then there exists some point v such that $\langle u, v \rangle \in \mathcal{M}|_{\langle w, \varphi \rangle}$ and $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \neg\psi$. From $\mathcal{M}, w \models \downarrow y@_x \Box \downarrow z(\neg(\varphi \wedge @_x y) \rightarrow @_y [-\varphi]@_z \psi)$, it holds that $\mathcal{M}, v \models @_y [-\varphi]@_z \psi$ where $w \in V(y)$ and $v \in V(z)$. Consequently, we have $\mathcal{M}|_{\langle w, \varphi \rangle}, v \models \psi$ that entails a contradiction. \square

In the rest of this section, we are not going to present a solution for this issue. Actually we conjecture that there exists no a recursion axiom for $[-\varphi]@_x \Box\psi$ in $\mathcal{H}(\downarrow)$, which is contrasted with our initial intuition. However, given Corollary 1, there must be some sort of recursion axioms for it. Thus a question arises:

Open Problem. Could there be a complete set of recursion axioms for $S_d\text{ML}$?

Through the above considerations, we understand why $\mathcal{H}(\downarrow)$ fails to do the job. In fact, there may be no easy solution, short of going to full first-order logic. All this suggests that, despite the axiomatizability in principle (as observed in Section 3), the structure of the logical validities of $S_d\text{ML}$ is computationally complex. This suspicion will be confirmed in the next section, where we prove the undecidability of the logic.

6 Undecidability of $S_d\text{ML}$

Up to now, we have already shown that $S_d\text{ML}$ is more expressive than the standard modal logic. Meanwhile, it is also a fragment of the hybrid logic $\mathcal{H}(\downarrow)$. It is well-known that the satisfiability problem for the standard modal logic is decidable. While, as noted in [13], $\mathcal{H}(\downarrow)$ is undecidable. So, is $S_d\text{ML}$ decidable or not?

Actually, there are some fragments of $\mathcal{H}(\downarrow)$ that are decidable. For instance, [15] shows that after removing all formulas containing a nesting of \Box, \downarrow and \Box , $\mathcal{H}(\downarrow)$ becomes decidable. But in this section, we will present a negative answer to the question above, i.e., the satisfiability problem for $S_d\text{ML}$ is undecidable. Moreover, we will identify the source of its high complexity. Before these results, we first show that $S_d\text{ML}$ lacks both the tree model property and the finite model property.

Theorem 6. *The logic $S_d\text{ML}$ does not have the tree model property.*

Proof. Consider the following formulas:

$$\begin{aligned} (R_1) \quad & p \wedge \Diamond p \wedge \Diamond \neg p \\ (R_2) \quad & \Box(p \rightarrow \Diamond p \wedge \Diamond \neg p) \\ (R_3) \quad & [-\neg p] \Box \Box p \end{aligned}$$

Let $\varphi_r := (R_1 \wedge R_2 \wedge R_3)$. We now show that, for any $\mathcal{M} = \{W, R, V\}$ and $w \in W$, if $\mathcal{M}, w \models \varphi_r$, then the evaluation point w is reflexive. By (R_1) , w has some p -successor(s) and some $\neg p$ -successor(s). Formula (R_2) states that each its p -successor w_1 also has at

least one p -successor w_2 and at least one $\neg p$ -successor w_3 . From (R_3) we know that, after deleting all links from w to the $\neg p$ -points, w_1 does not have $\neg p$ -successors any longer. If node w_1 is not w , then φ_r cannot be true at w . That is to say, for each $v \in W$, if Rwv and $\mathcal{M}, v \models p$, then $v = w$, i.e., $R(w) \cap V(p) = \{w\}$. So if formula φ_r is true, the evaluation point must be reflexive (with at least one $\neg p$ -successor). A model for φ_r is the \mathcal{M}_2 in the proof of Fact 1, and φ_r is true at the point v_1 . \square

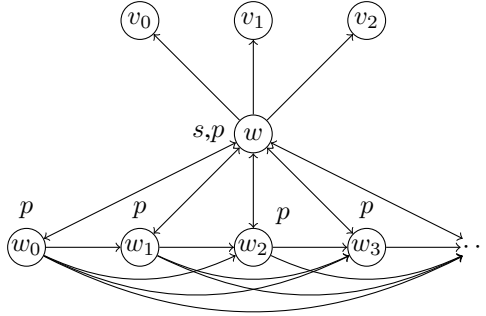
In addition, S_dML also lacks the finite model property. To show this, inspired by the methods of [13], we will construct a ‘spy point’, i.e., a special point which has access in one step to any reachable point in the model.

Theorem 7. *The logic S_dML does not have the finite model property.*

Proof. Let φ_∞ be the conjunction of the following formulas:

$$\begin{aligned}
 (F_1) \quad & s \wedge p \wedge \Box \neg s \wedge \Diamond p \wedge \Diamond \neg p \wedge \Box (\neg p \rightarrow \Box \perp) \\
 (F_2) \quad & \Box (p \rightarrow \Diamond s \wedge \Diamond \neg s \wedge \Box p) \\
 (F_3) \quad & \Box (p \rightarrow \Box (s \rightarrow \Box \neg s \wedge \Diamond \neg p)) \\
 (F_4) \quad & [-\neg p] \Box \Box (s \rightarrow \neg \Diamond \neg p) \\
 (F_5) \quad & \Box (p \rightarrow \Box (\neg s \rightarrow \Diamond s \wedge \Diamond \neg s \wedge \Box p)) \\
 (F_6) \quad & \Box (p \rightarrow \Box (\neg s \rightarrow \Box (s \rightarrow \Box \neg s \wedge \Diamond \neg p))) \\
 (F_7) \quad & [-\neg p] \Box \Box (\neg s \rightarrow \Box (s \rightarrow \neg \Diamond \neg p)) \\
 (Spy) \quad & \Box (p \rightarrow \Box (\neg s \rightarrow [-\neg s] \Box \Diamond (p \wedge \Box s))) \\
 (Irr) \quad & \Box (p \rightarrow [-s] \Box \Diamond s) \\
 (No-3cyc) \quad & \neg \Diamond (p \wedge [-s] \Diamond [-s] \Diamond \Diamond (\neg s \wedge \Box \neg s)) \\
 (Trans) \quad & \Box (p \rightarrow [-s] \Box \Box (\neg s \rightarrow [-\neg s] \Box \Diamond (\Box \neg s \wedge \Diamond \Box s)))
 \end{aligned}$$

First, we show that the formula φ_∞ is satisfiable. Consider the following model \mathcal{M} :



clearly, $\mathcal{M}, w \models \varphi_\infty$. Thus there exists at least one model satisfying formula φ_∞ .

Next, we show that for any $\mathcal{M} = \{W, R, V\}$ and $w \in W$, if $\mathcal{M}, w \models \varphi_\infty$, then W is infinite. For brevity, define that $B = \{v \in W \mid v \in R(w) \cap V(p)\}$, i.e., B is the set of the p -successors of w . In the following proof, we assume that all previous conjuncts hold.

By (F_1) , the evaluation point w is $(s \wedge p)$, and it cannot see any s -points. In particular, w cannot see itself. Besides, w has some p -successor(s) (i.e., $B \neq \emptyset$) and some $\neg p$ -successor(s) (i.e., $R(w) \setminus B \neq \emptyset$). In addition, each point in $R(w) \setminus B$ is a dead end.

From formula (F_2) , we know that each element in B can see some $(s \wedge p)$ -point(s) and $(\neg s \wedge p)$ -point(s), but cannot see any $\neg p$ -points. Hence each point in B has a successor distinct from itself.

According to formula (F_3) , for any $w_1 \in B$, each its s -successor can see some $\neg p$ -point(s), but cannot see any s -points.

By (F_4) , after removing all links from w to $\neg p$ -points, for each $w_1 \in B$, each of its s -successors w_2 has no $\neg p$ -successors. Thus (F_4) shows that each $w_1 \in B$ can see point w , and that for each s -point $w_2 \in W$, if w_2 is a successor of w_1 , then w_2 must be w .

Formulas (F_2) – (F_4) show the properties of the $(\neg s \wedge p)$ -points which are accessible from the point w in one step. Similarly, formulas (F_5) , (F_6) and (F_7) play the same role as (F_2) , (F_3) and (F_4) respectively, but focusing on showing the properties of the $(\neg s \wedge p)$ -points that are accessible from w in two steps. In particular, (F_7) guarantees that every $(\neg s \wedge p)$ -point w_1 which is accessible from w in two steps can also see w , and that for each s -point $w_2 \in W$, if w_2 is a successor of w_1 , then w_2 must be w .

Formula (Spy) says that, for each $(\neg s \wedge p)$ -point w_1 that is accessible from w in two steps, after removing the links from w_1 to the $\neg s$ -points, each successor w_2 of w_1 has a p -successor w_3 that only has s -successors. Besides, point w_2 must be s . By (F_7) , we know that $w_2 = w$. In addition, by (F_2) , w_3 should have some $\neg s$ -successor(s) if the cut induced by $[-s]$ does not take place at w_3 . So it holds that $w_3 = w_1$. In such a way, (Spy) makes the evaluation point w be a spy-point, and it ensures that each $(\neg s \wedge p)$ -point w_1 which is accessible from w in two steps is also accessible from w in one step. By (Irr) , for each $w_1 \in B$, after removing the link from w_1 to w , each its successor still can see w . Therefore each $w_1 \in B$ is irreflexive. Besides, $(No-3cyc)$ disallows cycles of length 2 or 3 in B , and $(Trans)$ forces the accessibility relation R to transitively order B .

Hence B is an unbounded strict partial order, thus it is infinite and so is W . Now we have already shown that φ_∞ is satisfiable, and that for each pointed model $\langle \mathcal{M}, w \rangle$, if $\mathcal{M}, w \models \varphi_\infty$, then \mathcal{M} is an infinite model. This completes the proof. \square

Now, by encoding the $N \times N$ tiling problem, we show that S_dML is undecidable. A tile t is a 1×1 square, of fixed orientation, with colored edges $right(t)$, $left(t)$, $up(t)$ and $down(t)$. The $N \times N$ tiling problem is: given a finite set of tile types T , is there a function $f : N \times N \rightarrow T$ such that $right(f(n, m)) = left(f(n+1, m))$ and $up(f(n, m)) = down(f(n, m+1))$? This problem is known to be undecidable (see [20]).

Following the ideas in [13], we will use three modalities \diamond_s , \diamond_u and \diamond_r . Correspondingly, a model $\mathcal{M} = \{W, R_s, R_u, R_r, V\}$ now has three kinds of accessibility relations. We will construct a spy point over the relation R_s . The relations R_u and R_r represent moving up and to the right, respectively, from one tile to the other. Besides, the operator $[-]$ will work in the usual way, i.e., all of the three kinds of relations should be cut if the current point have some particular successors via them.² Let us see the details.

²There is also no problem if we use three kinds of dynamic operators that correspond to the three kinds of accessibility relations respectively. In the proof of Theorem 8, these three kinds of links are disjoint.

Theorem 8. *The satisfiability problem for S_dML is undecidable.*

Proof. Let $T = \{T_1, \dots, T_n\}$ be a finite set of tile types. For each $T_i \in T$, we use $u(T_i)$, $d(T_i)$, $l(T_i)$, $r(T_i)$ to represent the colors of its up, down, left and right edges respectively. Besides, we code each tile type with a fixed propositional atom t_i . Now we will define a formula φ_T such that φ_T is satisfiable iff T tiles $N \times N$. Consider the following formulas:

$$\begin{aligned}
(M_1) \quad & s \wedge p \wedge \Box_s \neg s \wedge \Diamond_s p \wedge \Diamond_s \neg p \wedge \Box_s (\neg p \rightarrow \Box_s \perp) \\
(M_2) \quad & \Box_s (p \rightarrow \Diamond_s \top \wedge \Box_s (s \wedge \Diamond_s \neg p)) \\
(M_3) \quad & [\neg p] \Box_s \Box_s (s \wedge \neg \Diamond_s \neg p) \\
(M_4) \quad & \Box_s (p \rightarrow \Diamond_{\dagger} \top \wedge \Box_{\dagger} (\neg s \wedge p \wedge \Diamond_s \top \wedge \Box_s (s \wedge \Diamond_s \neg p))) \quad \dagger \in \{u, r\} \\
(M_5) \quad & [\neg p] \Box_s \Box_{\dagger} \Box_s \neg \Diamond_s \neg p \quad \dagger \in \{u, r\} \\
(M_6) \quad & \Box_s (p \rightarrow \Box_{\dagger} (\Diamond_u \top \wedge \Diamond_r \top \wedge \Box_u (\neg s \wedge p) \wedge \Box_r (\neg s \wedge p))) \quad \dagger \in \{u, r\} \\
(M_7) \quad & \Box_s (p \rightarrow [\neg s] \Box_{\dagger} (\Diamond_s s \wedge \neg \Diamond_{\dagger} \neg \Diamond_s s)) \quad \dagger \in \{u, r\} \\
(Spy) \quad & \Box_s (p \rightarrow \Box_{\dagger} [\neg s] \Box_s \Diamond_s (p \wedge \Box_u \perp \wedge \Box_r \perp)) \quad \dagger \in \{u, r\} \\
(Func) \quad & \Box_s (p \rightarrow [\neg s] \Box_{\dagger} [\neg s] \Diamond_s \Diamond_s (p \wedge \neg \Diamond_s s \wedge \Diamond_{\dagger} \top \wedge \\
& \Box_{\dagger} (\Box_u \perp \wedge \Box_r \perp)) \quad \dagger \in \{u, r\} \\
(No-UR) \quad & \Box_s (p \rightarrow [\neg s] \Box_u \Box_r \Diamond_s s \wedge [\neg s] \Box_r \Box_u \Diamond_s s) \\
(No-URU) \quad & \Box_s (p \rightarrow [\neg s] \Box_u \Box_r \Box_u \Diamond_s s) \\
(Conv) \quad & \Box_s (p \rightarrow [\neg s] \Diamond_u [\neg s] \Diamond_r [\neg s] \Diamond_s \Diamond_s (p \wedge \neg \Diamond_s s \wedge \\
& \Box_r (\Diamond_u \top \wedge \Diamond_r \top) \wedge \Diamond_u \neg \Diamond_s s \wedge \Diamond_r \Diamond_u (\Box_u \perp \wedge \Box_r \perp))) \\
(Unique) \quad & \Box_s (p \rightarrow \bigvee_{1 \leq i \leq n} t_i \wedge \bigwedge_{1 \leq i < j \leq n} (t_i \rightarrow \neg t_j)) \\
(Vert) \quad & \Box_s (p \rightarrow \bigwedge_{1 \leq i \leq n} (t_i \rightarrow \Diamond_u \bigvee_{1 \leq j \leq n, u(T_i)=d(T_j)} t_j)) \\
(Horiz) \quad & \Box_s (p \rightarrow \bigwedge_{1 \leq i \leq n} (t_i \rightarrow \Diamond_r \bigvee_{1 \leq j \leq n, r(T_i)=l(T_j)} t_j))
\end{aligned}$$

Define φ_T as the conjunction of the formulas above. Let $\mathcal{M} = \{W, R_s, R_u, R_r, V\}$ be an arbitrary model and $w \in W$. We first analyze the effects of the stated formulas on what this model must look like.

More concretely, suppose that $\mathcal{M}, w \models \varphi_T$. We show that \mathcal{M} is a tiling of $N \times N$. For brevity, define $G := \{v \in W \mid v \in R_s(w) \cap V(p)\}$ where $R_s(w) = \{v \in W \mid R_s w v\}$, and we will use its elements to represent the tiles. In the following proof, we also assume that all previous conjuncts hold.

Formula (M_1) is similar to (F_1) occurring in the proof of Theorem 7, except that (M_1) focuses on the relation R_s only.

By (M_2) , each tile w_1 has some successor(s) via the relation R_s , and each such successor w_2 is $(s \wedge p)$ and also has some $(\neg s \wedge \neg p)$ -successor(s) via R_s . It is worthy to note that formulas (M_1) and (M_2) illustrate that R_s is irreflexive.

Formula (M_3) ensures that each tile w_1 can see w via R_s , and that for each $(s \wedge p)$ -point $w_2 \in W$, if w_2 is accessible from w_1 via R_s , then $w_2 = w$.

(M_4) states that each tile has some successor(s) via R_u and some successor(s) via R_r . Besides, each point that is accessible from a tile via R_u or R_r is very similar to a tile: it is $(\neg s \wedge p)$, and has some $(s \wedge p)$ -successor(s) w_1 via relation R_s where each w_1 can see some $(\neg s \wedge \neg p)$ -point(s) via R_s .

By formula (M_5) , each $w_1 \in W$ accessible from a tile via R_u or R_r can see w by R_s . Also, for each $(s \wedge p)$ -point $w_2 \in W$, if it is accessible from w_1 via R_s , then $w_2 = w$.

Formula (M_6) ensures that each $w_1 \in W$ that is accessible from some tile via R_u or R_r also has some successor(s) via R_u and some successor(s) via R_r . Besides, each its successor via R_u or R_r is $(\neg s \wedge p)$.

From formula (M_7) , it follows that both R_u and R_r are irreflexive and asymmetric.

By (Spy) , we know that the evaluation point w is a spy point via the relation R_s .

Note that formula (M_4) says that each tile has some tile(s) above it and some tile(s) to its right. Now, with $(Func)$, we have that each tile has exactly one tile above it and exactly one tile to its right.

By $(No-UR)$, no tile can be above/below as well as to the left/right of another tile. Formula $(No-URU)$ disallows cycles following successive steps of the R_u , R_r , and R_u relations, in this order. Further more, $(Conv)$ ensures that the tiles are arranged as a grid.

Formula $(Unique)$ guarantees that each tile has a unique type. $(Vert)$ and $(Horiz)$ force the colors of the tiles to match properly.

Thus we conclude that \mathcal{M} is indeed a tiling of $N \times N$.

Next we show the other direction required for our proof. Suppose the function $f : N \times N \rightarrow T$ is a tiling of $N \times N$. Define a model $\mathcal{M} = \{W, R_s, R_u, R_r, V\}$ as follows:

$$\begin{aligned}
 W &= (N \times N) \cup \{w, v\} \\
 R_s &= \{\langle w, v \rangle\} \cup \{\langle w, n \rangle \mid n \in N\} \cup \{\langle n, w \rangle \mid n \in N\} \\
 R_u &= \{\langle \langle n, m \rangle, \langle n, m+1 \rangle \rangle \mid n, m \in N\} \\
 R_r &= \{\langle \langle n, m \rangle, \langle n+1, m \rangle \rangle \mid n, m \in N\} \\
 V(s) &= \{w\} \\
 V(p) &= \{w\} \cup N \\
 V(t_i) &= \{\langle n, m \rangle \in N \times N \mid f(\langle n, m \rangle) = T_i\}, \text{ for each } i \in \{1, \dots, n\} \\
 V(q) &= \emptyset, \text{ for any other propositional atoms } q
 \end{aligned}$$

In particular, w is a spy point in \mathcal{M} . By construction, we know that $\mathcal{M}, w \models \varphi_T$. □

Thus, perhaps surprisingly, given the simple-looking syntax and semantics of S_dML , the complexity of its logic is high. What is the reason for this high complexity, as contrasted with decidability of dynamic-epistemic logics of link deletion [11]? For SML, the reason offered by [5] is the stepwise nature of link deletion, and this is confirmed by the result in [1] showing how a very simple stepwise variant of public announcement logic is undecidable. However, our case is different, since links are cut in a uniform definable way: the only remaining potential culprit is then the locality.

To see the effects of this feature, recall the above formula (7). We already saw in Section 5.2 that a formula of the form $@_x[-\varphi]@_y\psi$ blocks the recursion format. In contrast,

consider a global version $S_d^g\text{ML}$ of $S_d\text{ML}$. The truth condition for $[-]$ now reads:

$$\langle W, R, V \rangle, w \models [-\varphi]\psi \text{ iff } \langle W, R \setminus \{ \langle s, t \rangle \in R \mid \mathcal{M}, t \models \varphi \}, V \rangle, w \models \psi.$$

Given the global change made in this semantics, here is a valid recursion axiom for \Box :

$$[-\varphi]\Box\psi \leftrightarrow \Box(\neg\varphi \rightarrow [-\varphi]\psi).$$

Indeed, following the general method for modal logics of definable model change presented in [11], one can find a complete set of recursion axioms for $S_d^g\text{ML}$:

Fact 9. *The logic $S_d^g\text{ML}$ is axiomatizable and decidable.*

The complexity effect of the local behavior of $S_d\text{ML}$ also show at a crucial step in our proof of undecidability. In the proof of Theorem 8, formula (Conv) forces the tiles to satisfy a first-order convergence property, i.e.,

$$\forall t \forall t_1 \forall t_2 (R_u t t_1 \wedge R_r t_1 t_2 \rightarrow \exists t_3 (R_r t t_3 \wedge R_u t_3 t_2)).$$

As noted in [8], this property can give logics high complexity.

By contrast, convergence is not definable in $S_d^g\text{ML}$, even though we expand the model with some extra tools, e.g. a spy point. Roughly speaking, given two tiles t_1 and t_2 that have same properties, we still can distinguish between them with $S_d\text{ML}$, say, their properties will be different after cutting some links starting from t_1 ; however, we cannot do this with $S_d^g\text{ML}$, since links are cut in a global way.³

The more general issue arising here goes beyond our specific logics of sabotage.

Open Problem. Does making update operations local (world-relative) generate undecidability in general for decidable dynamic-epistemic logics?

This would provide an alternative diagnosis to the comparison of sabotage and update offered in [5], closer to the modified dynamic-epistemic logics studied in [7].

7 Related Work

This article is primarily inspired by existing work on sabotage games [10] and their applications. A good source for the latest developments in sabotage modal logics is [5], which also has extensive references to current work on related modal logics for definable graph change. Meanwhile, a number of authors have studied other graph games using matching modal logics. For instance, in poison games, originating in graph theory, instead of deleting links, a player can poison a node, to make it inaccessible to her opponent. Poison games have been recently studied in the modal logics of [23], using the close similarities between these systems and variants of so-called memory logics [22] in the hybrid tradition. In another tradition, that of Boolean network games, [25] has proposed a logic of local fact change which can characterize Nash equilibria, providing a new way of looking at the interaction between graph games, network games and logics of control.

³From a technical point of view, to show that $S_d^g\text{ML}$ cannot define the convergence property, we need its notion of bisimulation, which is easily defined.

Throughout the paper, dynamic-epistemic logic [6, 9] has been used as a decidable contrasting design to our systems. Technically, our logic S_dML has resemblances to several recent logics for local announcements. [7] introduces a logic to characterize both global and local announcements. Similar to our set-up, it has definable updates of links, but there is also a difference. Although more expressive than public announcement logic, this logic is decidable. Moreover, we are inspired by other logics for local graph modifiers, too. For instance, [3] investigates a special type of local SML, whose dynamic operator refers to a model transition that cuts a link from the current state and then evaluates a formula at the target of the deleted arrow. Finally, more akin to the above-mentioned [25], [4] studies local modifiers that update the valuation at the evaluation point, and shows that adding those modifiers dramatically increases the expressive power of the logic.

Next, a highly relevant line of research for this article is hybrid logic, an area from which we have taken several basic techniques. As far as we know, [13] is the first to present the method of constructing a spy point, the main tool that was used to prove the undecidability of our logic S_dML . [2] shows how relation-changing logics such as SML can be seen as fragments of hybrid logics, and identifies various decidable fragments of those logics with the help of hybrid translations. This fits with our findings in Section 5.1. Finally, [19] merges hybrid logic with public announcement logic. Differently from the operator $[-]$ in S_dML , the announcement modality there operates in a global way, making it possible to axiomatize the logic by means of recursion axioms.

It remains to note that this article fits with the general program recently proposed in [12] for a much broader study of analysis and design for graph games in tandem with matching modal logics. In particular, it proposed various meaningful new games, and identified general questions behind the match between logic and game.

8 Summary and Further Directions

In this article, we started with a definable sabotage game S_dG that models some interesting phenomena in everyday life, and explored a matching logical system, definable sabotage modal logic S_dML . We presented a first-order translation for the logic, showed a characterization theorem with regard to a novel notion of definable sabotage bisimulation, probed an axiomatization for S_dML using recursion axioms in an extended hybrid language, and finally, we proved its undecidability.

Immediate technical open problems for our logic S_dML resemble those in the literature for SML. For instance, we would like to have a good Hilbert-style proof theory, which may perhaps be found by analyzing semantic tableaux for S_dML . Another open problem is the complexity of the schematic validities of our language.

Next, while our language can define winning positions for players in given finite graphs, it cannot express generic winning conditions across models. To obtain the latter, we need a modal μ -calculus enriched with our local definable deletion modality, whose behavior shows the complexities already noted for sabotage μ -calculus in [5].

In terms of generality, one would like to establish the precise connections between our logic S_dM and other modal logics for graph games in the cited literature. For instance, the

difference in expressive power that we noted in Section 4 between $S_d\text{ML}$ and SML does not preclude the existence of faithful embeddings either way.

As a final technical issue, we mentioned the contrast between locality and stepwise link deletion as sources of undecidability, discussed in Section 6. One could also merge these in a stepwise version of our logic, denoted $S_d^s\text{ML}$. Clearly, its validities are different from those of $S_d\text{ML}$: for instance, $[-]$ is no longer self-dual. Our methods from Section 6 should also be able to prove its undecidability, but we have not yet been able to do so.

We end by stepping back to reality. In our introduction, we mentioned social networks [21], where adding links (gaining friends or neighbors) is as important as deleting links (losing friends or neighbors). A connection between our logic and existing logics for social networks, and games played over these, would be a natural next step.

Another such step toward greater realism would arise when making connections to more elaborate versions of our game scenarios, for instance involving more complex independent goals for players than we have considered, or imperfect information when players cannot perfectly observe each other's moves. In general, such games may have probabilistic equilibria, and our logics would have to acquire interfaces with probability.

Acknowledgements

I wish to thank Johan van Benthem and Fenrong Liu for their generous help and patient guidance throughout all stages of this project. I would like to thank Alexandru Baltag for his inspiring suggestions. This work was supported by the Major Program of the National Social Science Foundation of China [17ZDA026].

References

- [1] C. Areces and J. van Benthem. The Logic of Stepwise Removal. Working manuscript. Institute for Logic, Language and Computation, University of Amsterdam and Department of Informatics, University of Cordoba.
- [2] C. Areces, R. Fervari, G. Hoffmann and M. Martel. Relation-Changing Logics as Fragments of Hybrid Logics. *Electronic Proceedings in Theoretical Computer Science*, **226**, pp. 16-29, 2016.
- [3] C. Areces, R. Fervari and G. Hoffmann. Relation-Changing Modal Operators. *Logic Journal of the IGPL*, **23**, pp. 601-627, 2015.
- [4] G. Aucher, P. Balbiani, L. Fariñas del Cerro and A. Herzig. Global and Local Graph Modifiers. *Electronic Notes in Theoretical Computer Science*, **231**, pp. 293-307, 2009.
- [5] G. Aucher, J. van Benthem and D. Grossi. Modal Logics of Sabotage Revisited. *Journal of Logic and Computation*, 2016.

-
- [6] A. Baltag, L. Moss and L. Solecki. The Logic of Public Announcements and Common Knowledge and Private Suspicions. In *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 1998)*, Evanston, IL, USA, July 22-24, 1998, I. Gilboa, ed., pp. 43–56. Morgan Kaufmann, 1998.
 - [7] F. Belardinelli, H. van Ditmarsch and W. van der Hoek. A Logic for Global and Local Announcements. In *Proceedings Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2017)*, J. Lang, ed., pp. 28-42, 2017.
 - [8] J. van Benthem. *Modal Logic for Open Minds*. CSLI Publications, 2010.
 - [9] J. van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, Cambridge UK, 2011.
 - [10] J. van Benthem. *Logic in Games*. The MIT Press, 2013.
 - [11] J. van Benthem and F. Liu. Dynamic Logic of Preference Upgrade. *Journal of Applied Non-Classical Logics*, **17**, pp. 157-182, 2007.
 - [12] J. van Benthem and F. Liu. Graph Games and Logic Design. Manuscript, 2018.
 - [13] P. Blackburn and J. Seligman. Hybrid Languages. *Journal of Logic, Language and Information*, **4**, pp. 251-272, 1995.
 - [14] P. Blackburn, M. de Rijke and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.
 - [15] B. ten Cate and M. Franceschet. On the Complexity of Hybrid Logics with Binders. In *Proceedings of Computer Science Logic 2005*, L. Ong, ed., Vol. 3634 of *Lecture Notes in Computer Science*, pp. 339-354, Springer Verlag, 2005.
 - [16] C. C. Chang and H. J. Keisler. *Model Theory*. Studies in Logic and the Foundations of Mathematics. North-Holland, 1973.
 - [17] P. Duchet and H. Meyniel. Kernels in Directed Graphs: A Poison Game. *Discrete Mathematics*, **115**, pp. 273-276, 1993.
 - [18] N. Gierasimczuk, L. Kurzen and F. Velázquez-Quesada. Learning and Teaching as a Game: A Sabotage Approach. In *Proceedings of LORI 2009*, X. He, J. Horty, and E. Pacuit, eds, Vol. 5834 of *Lecture Notes in Computer Science*, 2009.
 - [19] J. U. Hansen. A Hybrid Public Announcement Logic with Distributed Knowledge. *Electronic Notes in Theoretical Computer Science*, **273**, pp. 33-50, 2011.
 - [20] D. Harel. Recurring Dominoes: Making the Highly Undecidable Highly Understandable. *Annals of Discrete Mathematics*, **24**, pp. 51-72, 1985.
 - [21] F. Liu, J. Seligman and P. Girard. Logical Dynamics of Belief Change in the Community. *Synthese*, **191**, pp. 2403-2431, 2014.

-
- [22] S. Mera. *Modal Memory Logics*. PhD Thesis, Universidad de Buenos Aires and Université Henri Poincaré, 2009.
 - [23] K. Mierzewski and F. Zaffora Blando. The Modal Logic(s) of Poison Games. Working manuscript. Department of Philosophy, Stanford University, 2018.
 - [24] J. Seligman, F. Liu and P. Girard. Facebook and Epistemic Logic of Friendship. In *Proceedings of the 14th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2013)*, pp. 229-238, 2013.
 - [25] D. Thompson. Local Fact Change Logic. Working manuscript. Department of Philosophy, Stanford University, 2018.