Jakub Michaliszyn

Decidability of modal logics with particular emphasis on the interval temporal logics

PhD Thesis
Supervisor: prof. dr hab. Jerzy Marcinkowski

Institute of Computer Science
University of Wrocław
Wrocław 2012

Jakub Michaliszyn

Badanie rozstrzygalności logik modalnych ze szczególnym uwzględnieniem logik operujących na przedziałach czasowych

Rozprawa doktorska napisana pod kierunkiem prof. dra hab. Jerzego Marcinkowskiego

Instytut Informatyki
Uniwersytet Wrocławski
Wrocław 2012

Abstract

Modal logic for almost a hundred year has been an important topic in many academic disciplines, including philosophy, mathematics, linguistics and computer science. Currently it seems to be most intensively investigated by computer scientists. Among numerous branches in which modal logic, sometimes in disguise, finds applications, are hardware and software verification, cryptography and knowledge representation. This thesis is about the decidability and the complexity of satisfiability problem of some variants of modal logic. The first part contains results concerning so-called elementary modal logics, while the second part is devoted to the Halpern–Shoham logic.

Elementary modal logics. In many practical applications of modal logic, it is natural to consider some restrictions of classes of admissible frames.

Traditionally classes of frames are defined by modal axioms. However, many important classes of frames can be defined, in a more natural way, by first-order formulas, e.g. the formula $\forall xyz.xRy \wedge yRz \rightarrow xRz$ defines the class of transitive frames, corresponding to modal logic K4.

We prove that there is an universal first-order formula with only three variables such that the local satisfiability problem and the global satisfiability problem are undecidable. The results hold also for the finite satisfiability. We also prove that universal Horn formulas may lead to undecidability if we use inequality or two modalities.

On the positive side, we prove that the local satisfiability problem and the global satisfiability problem for modal logic over the class of frames defined by any universally quantified, first-order Horn formula is decidable. Then we show that also the finite satisfiability problem for modal logic over such classes is decidable. This subsumes decidability results for many natural modal logics, including T, B, K4, S4, S5.

Halpern–Shoham Logic. The Halpern–Shoham logic is a modal logic of time intervals. Some effort has been put in last ten years to classify fragments of this beautiful logic with respect to decidability of its satisfiability problem. We complete this classification by showing that the logic of subintervals, the fragment of the Halpern–Shoham logic where only the operator "during", or D, is allowed, is undecidable over discrete structures. This is surprising as this, apparently very simple, logic is decidable over dense orders and its reflexive variant is known to be decidable over discrete structures. Our result subsumes a lot of previous negative results for the discrete case, like the undecidability for ABE, BD, ADB, $A\bar{A}D$, and so on.

Streszczenie

Badanie problemów decyzyjnych dotyczących różnych logik rozpoczęło się już w latach trzydziestych ubiegłego wieku. Początkowo traktowano je jako narzędzie czysto matematyczne, w ostatnich dekadach jednak okazało się, że istnieje dużo połączeń między tymi problemami a automatyczną weryfikacją, sztuczną inteligencją czy nawet przetwarzaniem dokumentów. Ta praca jest poświęcona klasyfikowaniu opisanych niżej logik modalnych ze względu na rozstrzygalność i złożoność problemu spełnialności ich formuł.

Elementarne logiki modalne. Klasycznym sposobem definiowania logik modalnych jest podanie ich aksjomatów - w ten sposób powstają najbardziej znane logiki, takie jak K, S4, S5 czy T. Okazuje się, że tak zdefiniowane logiki mają naturalne semantyki wyrażone za pomocą struktur Kripkego. W przypadku elementarnych logik modalnych ([18]) jest odwrotnie - logikę modalną tworzy się przez ograniczenie struktur Kripkego zadaną formułą pierwszego rzędu, używającą jednego predykatu (R) interpretowanego jako relacja przejścia struktury. Przykładowo, logikę S4 można zdefiniować formułą $\forall xyz.xRx \wedge (xRy \wedge yRz \Rightarrow xRz)$. Taki sposób definiowania logiki wydaje się bardziej naturalny, jednak czasem prowadzi do logik nierozstrzygalnych. Pierwsza część tej pracy zawiera dowód, że nawet bardzo prosta formuła $(\forall xyz.\neg xRy \vee \neg xRz \vee yRx \vee zRx \vee zRy \vee yRz)$ może prowadzić do nierozstrzygalnej logiki. Wynik ten jest dowiedziony za pomocą narzędzia, które wykorzystuję również do pokazania nierozstrzygalności innych interesujących logik.

Oprócz negatywnych wyników, niniejsza praca wskazuje również dużą klasę formuł definiujących rozstrzygalne logiki modalne, mianowicie uniwersalne formuły hornowskie. W przypadku, gdy rozpatrujemy lokalną spełnialność, złożoność takich logik okazuje się być zawsze w PSpace, podczas gdy w przypadku globalnej spełnialności może wzrosnąć do ExpTime. Dokładna złożoność zależy od pewnych własności formuły — niniejsza praca zawiera szczegółową analizę wszystkich przypadków. Okazuje się, że w zależności od formuły problem spełnialności może być w P, NP-zupełny, PSpace-zupełny lub ExpTime-zupełny.

Ponadto, rozpatruję również problem skończonej spełnialności, który często jest traktowany jako bardziej praktyczny od ogólnej spełnialności. Okazuje się, że również ten problem jest zawsze rozstrzygalny dla logik definiowanych uniwersalnymi formułami hornowskimi.

Logika Halperna–Shohama. Logika Halperna–Shohama [15] służy opisywaniu rzeczywistości w oparciu o zdarzenia które trwają, a więc zajmują pewien przedział czasu. Podejście to znacznie różni się od klasycznego, w którym czas traktuje się jako zbiór punktów. Logika ta, zaproponowana pod koniec lat osiemdziesiątych, składa się z dwunastu operatorów modalnych opisujących możliwe relacje między przedziałami w ustalonym porządku. Co ważne, logika sama w sobie nie czyni istotnych założeń dotyczących natury czasu — może on być dyskretny, gęsty, z końcami, bez końców, liniowy, drzewiasty itd.

Logika Halperna–Shohama przez wiele lat nie wzbudzała większego zainteresowania. Na początku XXI wieku jednak, dzięki motywacji płynącej od ludzi zajmującej się sztuczną inteligencją, badanie różnych fragmentów tej logiki stało się bardzo popularne. Niniejsza praca zawiera dowód, że bardzo mały fragment tej logiki, zawierający jedynie operator modalny D, odnoszący się do podprzedziałów, ma nierozstrzygalny problem spełnialności (lokalnej i globalnej) w sytuacji, gdy rozpatrujemy porządki dyskretne.

Acknowledgements

First of all, I would like to thank my advisor Jerzy Marcinkowski for his idea to investigate Halpern–Shoham logic, his cooperation on this topic and many valuable suggestions.

I wish to thank Emanuel Kieroński for introducing me into the concepts of modal logic and the fruitful cooperation during all these years.

I wish to express my gratitude to all my other co-authors: Jan Otop, Ian Pratt-Hartmann, and Lidia Tendera for all the time we spent writing, thinking and reading.

Finally, I would like to thank Polish Ministry of Science and Higher Education, for their generous support under grants N206 022 31/3660 and N N206 371339, European Union for the support under project POKL.04.01.01-00-054/10-00, and The European Science Foundation for organizing and sponsoring the GAMES workshop, during which I heard about the Halpern–Shoham logic for the first time.

Contents

Part I.

Introduction

1. Modal logic

Modal logic was introduced by philosophers to study modes of truth. The idea was to extend propositional logic by some new constructions, of which two most important were $\Diamond\varphi$ and $\Box\varphi$, originally read as φ *is possible* and φ *is necessary*, respectively. A typical question was, given a set of axioms $\mathcal{A}$, corresponding usually to some intuitively acceptable aspects of truth, what is the logic defined by $\mathcal{A}$, i.e. which formulas are provable from $\mathcal{A}$ in a Hilbert-style system.

One of the most important steps in the history of modal logic was the invention in 1960s of a formal semantics based on the notion of the so-called Kripke structures. Basically, a Kripke structure is a directed graph, called a *frame*, together with a valuation of propositional variables. Vertices of this graph are called *worlds*. For each world truth values of all propositional variables can be defined independently. In this semantics, $\Diamond\varphi$ means *the current world is connected to some world in which* φ *is true*; and $\Box\varphi$, equivalent to $\neg\Diamond\neg\varphi$, means φ *is true in all worlds to which the current world is connected.*

It appeared that there is a beautiful connection between syntactic and semantic approaches to modal logic [38]: logics defined by axioms can be often equivalently defined by restricting classes of frames. E.g., the axiom $\Diamond\Diamond P \rightarrow \Diamond P$ (*if it is possible that* P *is possible, then* P *is possible*), is valid precisely in the class of transitive frames; the axiom $P \rightarrow \Diamond P$ (*if* P *is true, then* P *is possible*) – in the class of reflexive frames, $P \rightarrow \Box\Diamond P$ (*if* P *is true, then it is necessary that* P *is possible*) – in the class of symmetric frames, and the axiom $\Diamond P \rightarrow \Box\Diamond P$ (*if* P *is possible, then it is necessary that* P *is possible*) – in the class of Euclidean frames.

Thus we may think that every modal formula φ defines a class of frames, namely the class of those frames in which φ is valid. A formula φ is valid in a frame K if for any possible truth-assignment of propositional variables to the worlds of K, φ is true at every world. To express this definition we require second-order logic, since it involves quantification over sets of elements: for each variable P and a subset V of the set of worlds we have to consider the case in which P is true exactly in the worlds from V. Note however, that many important classes of frames, in particular all the classes we mentioned above, can be defined by simple first-order formulas. For a given first-order sentence Φ over the signature consisting of a single binary symbol R we define $\mathcal{K}_\Phi$ to be the set of those frames which satisfy Φ.

It is not hard to see that some modal logics defined by a first-order formula are undecidable. A stronger result was presented in [17]—it was shown that there exists a universal first-order formula with the equality such that the global satisfiability problem over the frames that satisfy this formula is undecidable. In [19], this result was improved — it was shown that the equality is not necessary. The proof from

[19] works also for local satisfiability. In this thesis, we show that even a very simple formula with three variables without the equality may lead to undecidability.

Decidability for various classes of frames can be shown by employing the so-called standard translation of modal logic to first-order logic. Indeed, the satisfiability of a modal formula φ in $\mathcal{K}_\Phi$ is equivalent to satisfiability of $st(\varphi) \wedge \Phi$, where $st(\varphi)$ is the standard translation of φ. In this way, we can show that even multimodal logic is decidable in any class defined by two-variable logic [33], even extended with linear order [35] or equivalence closures of two distinguished binary relations [21].

A number of decidability results may be obtained by adapting the results for the guarded fragment [12]. It has been shown that many interesting extensions of this logic are decidable, including some restricted application of fixed-points [13] and transitive closures [27] in guards. These results often can be extended for the finite satisfiability problem [2, 22]. The complexity bounds obtained this way, however, are high — usually between ExpTime and 2NExpTime.

The classes of frames we mentioned earlier, i.e. transitive, reflexive, symmetric and Euclidean are decidable. They can be defined by first-order sentences even if we further restrict the language to universal Horn formulas, UHF. Universal Horn formulas were considered in [18], where a dichotomy result was proved, that the satisfiability problem for modal logic over the class of structures defined by an UHF formula (with an arbitrary number of variables) is either in NP or PSpace-hard. The authors of [18] conjecture that the problem is decidable in PSpace for all universal Horn formulas. We confirm this conjecture in this thesis.

In case of some UHF formulas, decidability of corresponding modal logics is shown by demonstrating the finite model property, i.e. by proving that every modal formula satisfiable over $\mathcal{K}_\Phi$ has also a finite model in $\mathcal{K}_\Phi$. However, it is not hard to construct a UHF formula Φ, such that some modal formulas have only infinite models over $\mathcal{K}_\Phi$. Assume e.g. that Φ enforces irreflexivity and transitivity, and consider the following modal formula: $\Diamond p \wedge \Box\Diamond p$.

This naturally leads to the question, whether for any UHF formula Φ the finite satisfiability problem for modal logic over K_Φ is decidable. This question is particularly important, if one considers practical applications, in which the structures (corresponding e.g. to knowledge bases or descriptions of programs) are usually required to be finite.

Decision procedures for the finite satisfiability problem for modal and related logics are very often more complex than procedures for general satisfiability. As argued in [41], the model theoretic reason for the good behavior of modal logics is the tree model property. A standard technique is to unravel an arbitrary model into a (usually infinite) tree. Clearly such an approach is not sufficient if we are interested only in finite models. In this thesis we are, however, able to prove that also finite satisfiability problems are always decidable for logics over the classes defined by universal Horn formulas.

2. Halpern–Shoham logic

In classical temporal logic, structures are defined by assigning properties (propositional variables) to time points (time is an ordering, discrete or dense). However, not all phenomena can be well described by such logics. Sometimes, we need to talk about actions (processes) that take some time and we would like to be able to say that one such action takes place, for example, during or after another.

The Halpern–Shoham logic [15], which is the subject of the second part of this thesis, is one of the modal logics of time intervals. Judging by the number of papers published, and by the amount of work devoted to the research on it, this logic is probably the most influential time interval logic. But historically it was not the first one. Actually, the earliest papers about intervals in context of modal logic were written by philosophers, e.g., [16]. In computer science, the earliest attempts to formalize time intervals were process logic [36, 37] and interval temporal logic [34]. Relations between intervals in linear orders from an algebraic point of view were first studied systematically by Allen [1].

The Halpern–Shoham logic is a modal temporal logic, where the elements of a model are no longer — like in classical temporal logics — points in time, but rather pairs of points in time. Any such pair — call it $[p, q]$, where q is equal to or later than p — can be viewed as a (closed) time interval, that is, the set of all time points between p and q. HS logic does not assume anything about order — it can be discrete or continuous, linear or branching, complete or not.

Halpern and Shoham introduce six modal operators acting on intervals. Their operators are "begins" B, "during" D, "ends" E, "meets" A, "later" L, "overlaps" O and the six inverses of those operators: $\bar{B}, \bar{D}, \bar{E}, \bar{A}, \bar{L}$, $\bar{O}$. It is easy to see that the set of operators is redundant. For example, A, B and E can define D (B and E suffice for that – a prefix of my suffix is my infix) and L (here A is enough –"later" means "meets an interval that meets"). The operator O can be expressed using E and $\bar{B}$.

In their paper, Halpern and Shoham show that (satisfiability of formulae of) their logic is undecidable. Their proof requires logic with three operators (B, E and A are explicitly used in the formulae and, as we mentioned above, once B, E and A are allowed, D and L come for free) so they state a question about decidable fragments of their logic.

Considerable effort has been put since then to settle this question. First, it was shown [24] that the BE fragment is undecidable. Recently, negative results were also given for the classes $B\bar{E}$, $\bar{B}\bar{E}$, $\bar{B}E$, $A\bar{A}D, \bar{A}D\bar{B}, \bar{A}DB, \bar{A}\bar{D}\bar{B}, \bar{A}\bar{D}B$ [5, 8], and BD [26]. Another elegant negative result was that $O\bar{O}$ is undecidable over discrete orders [6].

On the positive side, it was shown that some small fragments, like $B\bar{B}$ or $E\bar{E}$, are decidable and easy to translate into standard, point-based modal logic [10]. The fragment using only A and $\bar{A}$ is harder and its decidability was only recently shown [8, 9]. Obviously, this last result implies decidability of $L\bar{L}$ as L is expressible by A. Another fragment known to be decidable is $AB\bar{B}\bar{L}$ [32].

A very simple, interesting fragment of the Halpern and Shoham logic of unknown status was the fragment with the single operator D ("during"), which we call here *the logic of subintervals*. Since D does not seem to have much expressive power (a natural language account of a D-formula would be "each morning I spend a while thinking of you" or "each nice period of my life contains an unpleasant fragment"), logic of subintervals was widely believed to be decidable. A number of decidability results concerning variants of this logic has been published. For example, it was shown in [7, 31] that satisfiability of formulae of logic of subintervals is decidable over dense structures. In [30] decidability is proved for the (slightly less expressive) "reflexive D". The results in [40] imply that D (as well as some richer fragments of the HS logic) is decidable if we allow models in which not all the intervals defined by the ordering are elements of the Kripke structure. On the negative side, no nontrivial lower bound was known for satisfiability of this logic.

In the second part of this thesis, we show that satisfiability of formulae from the D fragment is undecidable over the class of finite orderings as well as over the class of all discrete orderings. Our result subsumes the negative results for the discrete case for ABE [15], BD [26] and ADB, $A\bar{A}D$ [5, 8]. The logic of subintervals for finite orderings is so simple that we are tempted to write that it is one of the simplest known undecidable logics.

Part II.

Elementary modal logics

3. Overview

The authors of [18] conjecture that the satisfiability problem of the modal logic over any class definable by a UHF sentence is decidable in PSpace. We confirm the conjecture from [18] with the following theorem.

Theorem 3.1. *Let Φ be a* UHF *sentence. Then the local and the global satisfiability problems for unimodal logic over $\mathcal{K}_\Phi$ are in PSPACE and EXPTIME, resp.*

This theorem extends the decidability results for some well-known modal logics, e.g., T, B, K4, S4, and S5. It also works for some interesting classes of frames, for which, up to our knowledge, decidability has not been established so far. An example is the class defined by $\forall xyzv(xRy \wedge yRz \wedge zRv \to xRv)$.

Then, we extend the result to cover the Horn formulas with equality. We prove the following.

Theorem 3.2. *Let Φ be a universal Horn formula with equality. Then the local and the global satisfiability problems for unimodal logic over $\mathcal{K}_\Phi$ are in PSPACE and EXPTIME, resp.*

We also show decidability of the finite satisfiability.

Theorem 3.3. *Let Φ be a universal Horn formula. Then the finite local and the finite global satisfiability problems for modal logic over $\mathcal{K}_\Phi$ are decidable.*

We also show that those results are optimal in many ways, showing that some richer classes contain formulas which define undecidable logics.

Theorem 3.4. *There exist three-variable universal formulas Γ, Γ', universal Horn formulas Γ_i, Γ_i' with inequality, and universal Horn formulas Γ_b, Γ_b' (all six formulas are without equality) such that the following problems are undecidable.*

3.4.1 The **global** *satisfiability problem for* **unimodal** *logic over $\mathcal{K}_\Gamma$.*

3.4.2 The **local** *satisfiability problem for* **unimodal** *logic over $\mathcal{K}_{\Gamma'}$.*

3.4.3 The **global** *satisfiability problem for* **unimodal** *logic over $\mathcal{K}_{\Gamma_i}$.*

3.4.4 The **local** *satisfiability problem for* **unimodal** *logic over $\mathcal{K}_{\Gamma_i'}$.*

3.4.5 The **global** *satisfiability problem for* **bimodal** *logic over $\mathcal{K}_{\Gamma_b}$.*

3.4.6 The **local** *satisfiability problem for* **bimodal** *logic over $\mathcal{K}_{\Gamma_b'}$.*

The same holds if we consider only finite structures.

Related work. The results presented in this part of the thesis come from three published papers [20, 29, 28]. The only actual exception is in Section 5 — Lemma 5.1 is a generalization of the technique used in [20] and [29]. Also, the undecidability result for the bimodal case is not published.

4. Preliminaries

As we work with both first-order logic and modal logic we will help the reader to distinguish them in our notation: we denote first-order formulas with Greek capital letters, and modal formulas with Greek lower-case letters. We assume that the reader is familiar with first-order logic and propositional logic.

Except for Section 5.4, we consider only unimodal logics. Modal logic extends propositional logic with the operator $\Diamond$ and its dual $\Box$. Formulas of modal logic are interpreted in Kripke structures, which are triples of the form $\langle W, R, \pi \rangle$, where W is a set of worlds, $\langle W, R \rangle$ is a directed graph called a *frame*, and π is a function that assigns to each world a set of propositional variables which are true at this world. We say that a structure $\langle W, R, \pi \rangle$ is *based* on the frame $\langle W, R \rangle$. For a given class of frames $\mathcal{K}$, we say that a structure is $\mathcal{K}$-based if it is based on some frame from $\mathcal{K}$. We will use calligraphic letters $\mathcal{M}, \mathcal{N}$ to denote frames and Fraktur letters $\mathfrak{M}, \mathfrak{N}$ to denote structures. To keep the notation light, we identify a structure $\langle W, R, \pi \rangle$ with $\langle \langle W, R \rangle, \pi \rangle$.

For a frame $\langle W, R \rangle$ and a subset $W' \subseteq W$, we define $R_{\restriction W'} = R \cap (W' \times W')$. Similarly, for a labeling function π, we define $\pi_{\restriction W'}$ to be such that $\pi_{\restriction W'}(w) = \pi(w)$ for all $w \in W'$ and $\pi_{\restriction X}$ to be such that $\pi_{\restriction X}(w) = \pi(w) \cap X$. We define the restriction of a frame $\langle W, R \rangle_{\restriction W'}$ for $W' \subseteq W$ as $\langle W', R_{\restriction W'} \rangle$.

The semantics of modal logic is defined recursively. A modal formula φ is (locally) *satisfied* in a world w of a model $\mathfrak{M} = \langle W, R, \pi \rangle$, denoted as $\mathfrak{M}, w \models \varphi$ if

(i) $\varphi = p$ where p is a variable and $\varphi \in \pi(w)$,

(ii) $\varphi = \neg p$ where p is a variable and $\varphi \notin \pi(w)$,

(iii) $\varphi = \varphi_1 \vee \varphi_2$ and $\mathfrak{M}, w \models \varphi_1$ or $\mathfrak{M}, w \models \varphi_2$,

(iv) $\varphi = \varphi_1 \wedge \varphi_2$ and $\mathfrak{M}, w \models \varphi_1$ and $\mathfrak{M}, w \models \varphi_2$,

(v) $\varphi = \Diamond \varphi'$ and there exists a world $v \in W$ such that $(w, v) \in R$ and $\mathfrak{M}, v \models \varphi'$,

(vi) $\varphi = \Box \varphi'$ and for all worlds $v \in W$ such that $(w, v) \in R$ we have $\mathfrak{M}, v \models \varphi'$.

Note that in this paper all formulas are in the negation normal form. By $|\varphi|$ denote the length of φ. We say that a formula φ is *globally* satisfied in $\mathfrak{M}$, denoted as $\mathfrak{M} \models \varphi$, if for all worlds w of $\mathfrak{M}$, we have $\mathfrak{M}, w \models \varphi$.

For a given class of frames $\mathcal{K}$, we say that a formula φ is *locally* (resp. *globally*) $\mathcal{K}$-*satisfiable* if there exists a $\mathcal{K}$-based structure $\mathfrak{M}$, and a world $w \in W$ such that $\mathfrak{M}, w \models \varphi$ (resp. $\mathfrak{M} \models \varphi$).

We employ a standard notion of a type. For a given formula φ, a Kripke structure $\mathfrak{M}$, and a world $w \in W$ we define the *type* of w (with respect to φ) in $\mathfrak{M}$ as $tp^{\varphi}_{\mathfrak{M}}(w) = \{\psi : \mathfrak{M}, w \models \psi$ and ψ is subformula of $\varphi\}$. We write $tp_{\mathfrak{M}}(w)$ if the formula is clear from context. Note that $|tp^{\varphi}_{\mathfrak{M}}(w)| \leq |\varphi|$, where $|\varphi|$ denotes the length of φ.

In our constructions we use the following terminology. A world w is *k-followed* (*k-preceded*) in a frame $\mathcal{M}$, if there exists a directed path $(w, u_1, u_2, \ldots, u_k)$ (resp. $(u_1, u_2, \ldots, u_k, w)$) in $\mathcal{M}$. Note that we do not require this path to consist of distinct elements. We say that a world w is *k-inner* in $\mathcal{M}$ if it is k-proceeded and k-followed. We use also naturally defined notions of ∞-preceded, ∞-followed, and ∞-inner worlds. In particular, a world on a cycle is ∞-inner.

The set of *universal Horn formulas*, UHF, is defined as the set of those Φ over the language $\{R\}$ which are of the form $\forall\vec{x}.\Phi_1 \wedge \Phi_2 \wedge ... \wedge \Phi_i$, where each Φ_i is a Horn clause. A Horn clause is a disjunctions of literals of which at most one is positive. We usually present Horn clauses as implications. For example, the formula $\forall xyz.(xRy \wedge yRz \Rightarrow xRz) \wedge (xRx \Rightarrow \bot)$ defines the set of transitive and irreflexive frames. We often skip the quantifiers and represent such formulas as a set of clauses, e.g.: $\{xRy \wedge yRz \Rightarrow xRz, xRx \Rightarrow \bot\}$. We assume without loss of generality that each Horn clause consists of variables x, y and $z_1, z_2, \ldots,$ and is of the form $\Psi \Rightarrow \bot$, $\Psi \Rightarrow xRx$, or $\Psi \Rightarrow xRy$. We define $\Psi(v_x, v_y, v_1, \ldots, v_k)$ as the instantiation of Ψ with $x = v_x$, $y = v_y$, $z_1 = v_1$, $z_2 = v_2$, and so on, e.g. $(xRz_1 \wedge z_1Rz_2 \wedge z_2Ry \Rightarrow xRy)(a, b, c, d) = aRc \wedge cRd \wedge dRb \Rightarrow aRb$. We consider also the set of *universal Horn formulas with equality*, $\mathsf{UHF}^{=}$, and the set of *universal Horn formulas with inequality*, $\mathsf{UHF}^{\neq}$, defined in a similar way.

We define the *local* (resp. *global*) *satisfiability problem* $\mathcal{K}$-SAT (resp. global $\mathcal{K}$-SAT) as follows. For a given modal formula, is this formula locally (resp. globally) $\mathcal{K}$-satisfiable? For a given $\Phi \in \mathsf{UHF}$, we define $\mathcal{K}_\Phi$ as the class of frames satisfying Φ. We will be interested in local and global $\mathcal{K}_\Phi$-SAT problems.

When consider problems $\mathcal{K}_\Phi$-SAT and global $\mathcal{K}_\Phi$-SAT, formula Φ is fixed and does not depend on the input. However, the complexity depends on this formula. To hide unnecessary details, we often use a function $\mathfrak{g}$. Please keep in mind that once Φ is fixed, $\mathfrak{g}(|\Phi|)$ can be treated as a constant, and, while considering complexity, the precise value of $\mathfrak{g}$ is not important (it will follow from the proofs).

To keep the notation light, we abbreviate $\mathbb{N} \cup \{\infty\}$ by $\mathbb{N}_\infty$ and assume that $n \bmod \infty = n$ for any n.

The following fact will prove useful.

Fact 4.1. *Assume that X is a (possibly infinite) set of positive numbers that is closed under addition. Then, there exists a finite subset X' of X such that $gcd(X) = gcd(X')$. Moreover, for each $x > lcm(X')$, $gcd(X')$ divides x if and only if $x \in X$.*

The proof is straightforward, by employing Euclidian algorithm.

4.1. Domino systems

By $\mathbb{Z}_k$ we denote the set $\{0, 1, \ldots, k-1\}$.

Definition 4.2. *A* domino system *is a tuple* $\mathcal{D} = (D, H_{\mathcal{D}}, V_{\mathcal{D}})$*, where* D *is a set of domino pieces and* $H_{\mathcal{D}}, V_{\mathcal{D}} \subseteq D \times D$ *are binary relations specifying admissible horizontal and vertical adjacencies. We say that* $\mathcal{D}$ tiles $\mathbb{N} \times \mathbb{N}$ *if there exists a function* $t : \mathbb{N} \times \mathbb{N} \to \mathcal{D}$ *such that* $\forall i, j \in \mathbb{N}$ *we have* $(t(i,j), t(i+1,j)) \in H_{\mathcal{D}}$ *and* $(t(i,j), t(i,j+1)) \in V_{\mathcal{D}}$*. Similarly,* $\mathcal{D}$ *tiles* $\mathbb{Z}_k \times \mathbb{Z}_l$*, for* $k, l \in \mathbb{N}$*, if there exists* $t : \mathbb{Z}_k \times \mathbb{Z}_l \to \mathcal{D}$ *such that* $(t(i,j), t(i+1 \mod k, j)) \in H_{\mathcal{D}}$ *and* $(t(i,j), t(i, j+1 \mod l)) \in V_{\mathcal{D}}$*.*

The following lemma comes from [3, 14].

Lemma 4.3. *The following problems are undecidable:*

(i) For a given domino system $\mathcal{D}$ *determine if* $\mathcal{D}$ *tiles* $\mathbb{N} \times \mathbb{N}$*.*

(ii) For a given domino system $\mathcal{D}$ *determine if there exists* $k \in \mathbb{N}$ *such that* $\mathcal{D}$ *tiles* $\mathbb{Z}_k \times \mathbb{Z}_k$*.*

The bounded-space domino problem is defined as follows. For a given triple $\langle \mathcal{D}, V_{\mathcal{D}}, H_{\mathcal{D}} \rangle$, where $V_{\mathcal{D}}, H_{\mathcal{D}} \subseteq \mathcal{D} \times \mathcal{D}$, and $n = O(|D|)$, is there a tiling $t : \mathbb{Z}_n \times \mathbb{N} \to \mathcal{D}$ such that for all $k < n$ and $l \in \mathbb{N}$, $(t(k,l), t(k,l+1)) \in V_{\mathcal{D}}$ and if $k < n-1$, then $(t(k,l), t(k+1,l)) \in H_{\mathcal{D}}$? It is an easy exercise that this problem is PSpace-complete.

5. Undecidability

In this section we work with signatures consisting of a single binary symbol R (except subsection 5.4), and a number of unary symbols, including P_{ij}, for $0 \leq i, j \leq 2$. Structures over such signatures can be naturally viewed as Kripke structures in which R is the accessibility relation, and unary relations describe valuations of propositional variables. To simplify our notation we assume that subscripts in P_{ij} are always taken modulo 3, e.g. if $i = 2, j = 2$, then $P_{i+1,j+1}$ denotes P_{00}. We define $\mathcal{P} = \{P_{ij} | i, j \in \{0, 1, 2\}\}$.

This section is organized as follows. Subsection 5.1 provides a tool for all the proofs. Subsections 5.2, 5.3, and 5.4 contain proofs of undecidability of global satisfiability problems for three modal logics. Finally, in Subsection 5.5 we discuss local satisfiability problems.

5.1. Key tool

For $l \in \mathbb{N}_\infty$ we define the grid $\mathfrak{G}_l$ as $\langle W_l, R_l, \pi_l \rangle$ where

- $W_l = \{a_{ij} | 0 \leq i, j < l\}$;
- $a_{ij} R a_{i'j'}$ iff $i' = i$ and $j' = j + 1 \mod l$ or $i' = i + 1 \mod l$ and $j' = j$;
- $\pi(a_{ij}) = \{P_{ij}\}$ for all i, j.

We say that a structure $\langle W, R, \pi \rangle$ is $\mathfrak{G}_l$*-like* if $W = W_l$, $R_l \subseteq R$, and $a_{ij} R a_{i'j'}$ implies that $|i - i'| \mod l \leq 1$ and $|j - j'| \mod l \leq 1$, and $\pi = \pi_l$. Roughly speaking, $\mathfrak{G}_l$-like structure contains $\mathfrak{G}_l$ and, perhaps, some additional edges connecting worlds that are close in grid. Figure 5.1 contains an example of such a structure.

We say that a structure $\langle W, R, \pi \rangle$ is an *extension of* $\mathfrak{G}_l$ if $\langle W_\pi, R_{\restriction W_\pi}, (\pi_{\restriction W_\pi})_{\restriction \mathcal{P}} \rangle$ is $\mathfrak{G}_l$-like, where $W_\pi = \{w \in W | \exists i, j. P_{ij} \in \pi(w)\}$ is a substructure consisting of all elements satisfying variables P_{ij}. Figures 5.2 and 5.3 contain examples of such extensions.

We are ready to define our tool for proving undecidability.

Lemma 5.1. *Let Φ be a first order formula and φ be a modal logic formula such that*

(a) there exists a $\mathcal{K}_\Phi$-based model of φ which is an extension of $\mathfrak{G}_\infty$;

(b) for all k there exists a $\mathcal{K}_\Phi$-based model of φ which is an extension of $\mathfrak{G}_{3k}$;

(c) for each $\mathcal{K}_\Phi$-based model $\langle W, R, \pi\rangle$ of φ there is a homomorphism from $\mathfrak{G}_\infty$ into $\langle W, R, \pi_{\restriction\mathcal{P}}\rangle$.

(d) for each finite $\mathcal{K}_\Phi$-based model $\langle W, R, \pi\rangle$ of φ there is a homomorphism from $\mathfrak{G}_l$ into $\langle W, R, \pi_{\restriction\mathcal{P}}\rangle$ for some $l \in \mathbb{N}$.

Then the global satisfiability problem and the finite global satisfiability problem over $\mathcal{K}_\Phi$ are undecidable.

Proof. Let Φ and φ be a formulas that satisfy the assumptions of Lemma 5.1. For a given domino system $\mathcal{D} = (D, D_H, D_V)$ we define

$$\lambda^{\mathcal{D}} = \lambda_0 \wedge \bigwedge_{0\le i,j\le 2} (\lambda^H_{ij} \wedge \lambda^V_{ij}).$$

For every $d \in D$ we use a fresh propositional letter P_d. λ_0 says that each world contains a domino piece, λ^H_{ij} and λ^V_{ij} say that pairs of elements satisfying horizontal and vertical adjacency relations respect D_H and D_V, respectively.

$$\lambda^H_{ij} = \bigwedge_{d\in D} ((P_d \wedge P_{ij}) \to \Box(P_{i+1,j} \to \bigvee_{d':(d,d')\in D_H} P_{d'})),$$

$$\lambda^V_{ij} = \bigwedge_{d\in D} ((P_d \wedge P_{ij}) \to \Box(P_{i,j+1} \to \bigvee_{d':(d,d')\in D_V} P_{d'})).$$

Lemma 5.1 is a straightforward consequence of Lemma 4.3 and the following facts.

(i) $\mathcal{D}$ tiles $\mathbb{N}\times\mathbb{N}$ iff there exists a $\mathcal{K}_\Phi$-based model of $\varphi \wedge \lambda^{\mathcal{D}}$.

(ii) $\mathcal{D}$ tiles some $\mathbb{Z}_k \times \mathbb{Z}_k$ iff there exists a finite $\mathcal{K}_\Phi$-based model of $\varphi \wedge \lambda^{\mathcal{D}}$.

Proof of (i), $\Rightarrow$. Let t be a tiling of $\mathbb{N}\times\mathbb{N}$ and $\mathfrak{M}$ be a $\mathcal{K}_\Phi$-based model of φ which is an extension of $\mathfrak{G}_\infty$. We construct $\mathfrak{M}'$ by extending the labeling of $\mathfrak{M}$ in such a way that for every $i, j \in \mathbb{N}$ the element $a_{i,j}$ satisfies $P_{t(i,j)}$. It is readily checked that $\mathfrak{M}'$ is as required.

Proof of (ii), $\Rightarrow$. If $\mathcal{D}$ tiles $\mathbb{Z}_k \times \mathbb{Z}_k$ then it also tiles $\mathbb{Z}_{3k} \times \mathbb{Z}_{3k}$. Let t be a tiling of $\mathbb{Z}_{3k} \times \mathbb{Z}_{3k}$. Let $\mathfrak{M}$ be a $\mathcal{K}_\Phi$-based model of φ which is an extension of $\mathfrak{G}_{3k}$. We construct $\mathfrak{M}'$ by extending the labeling of $\mathfrak{M}$ in such a way that for every $i, j \in \mathbb{N}$ the element $a_{i,j}$ satisfies $P_{t(i,j)}$. Again, checking that $\mathfrak{M}'$ is as required is straightforward.

Proofs of (i) and (ii), $\Leftarrow$. Let $\mathfrak{M} = \langle W, R, \pi\rangle$ be a (finite) $\mathcal{K}_\Phi$-based model of $\varphi \wedge \lambda^{\mathcal{D}}$ and f be a homomorphism from $\mathfrak{G}_\infty$ into $\langle W, R, \pi_{\mathcal{P}}\rangle$. We define a tiling $t : \mathbb{N}\times\mathbb{N} \to D$ ($t : Z_l \times Z_l \to D$, resp.) by setting $t(i, j) = d$ for such d that $f(i, j)$ satisfies P_d (there is at least one such d owing to λ_0). The formulas $\lambda^H_{ij}, \lambda^V_{ij}$ imply that t is a correct tiling. $\Box$

5.2. First-order formulas with three variables

We define formulas Γ and τ satisfying the conditions of Lemma 5.1. Let

$$\Gamma = \forall xyz.(xRy \land \neg yRx \land xRz \land \neg zRx) \rightarrow (yRz \lor zRy)$$

It says, that if there are one-way connections from a world x to worlds y, z, then there is also a connection (not necessarily one-way) between y and z. The $\mathfrak{G}_\infty$-like structure illustrated in Fig. 5.1 (we assume that this structure is reflexive) is a model of Γ. Note that it is important that some connections are two-way.

A modal formula τ says that every element satisfying P_{ij} has three R-successors: one in $P_{i+1,j}$, one in $P_{i,j+1}$, and one in $P_{i+1,j+1}$, and forbids connections from $P_{i+1,j+1}$ to $P_{i,j+1}$, $P_{i+1,j}$, and P_{ij}. If we consider now any element a in a model, we see that τ enforces the existence of its horizontal successor a_h, its vertical successor a_v and its upper-right diagonal successor a_d (see Fig. 5.1). By τ, the connections to these successors are one-way, so we need, by Γ, connections between a_h and a_d, and a_v and a_d. Again, by τ, these connections have to go from a_h to a_d, and from a_v to a_d, so a_d is indeed a horizontal successor of a_v, and a vertical successor of a_h. Formally

$$\tau = \tau_0 \land \bigwedge_{0 \leq i,j \leq 2} (\tau_{ij}^{\Diamond} \land \tau_{ij}^{\Box}),$$

where τ_0 says that each element satisfies one of P_{ij}, $\tau_{ij}^{\Diamond}$ ensures that all elements have appropriate horizontal, vertical and upper-right diagonal successors, and $\tau_{ij}^{\Box}$ forbids reversing the horizontal, vertical and upper-right diagonal arrows.

$$\tau_{ij}^{\Diamond} = P_{ij} \rightarrow (\Diamond P_{i+1,j} \land \Diamond P_{i,j+1} \land \Diamond P_{i+1,j+1}),$$

$$\tau_{ij}^{\Box} = P_{ij} \rightarrow \Box(\neg P_{i-1,j} \land \neg P_{i,j-1} \land \neg P_{i-1,j-1}).$$

Note that $\tau_{ij}^{\Box}$ allow for reflexive edges.

It is not hard to see that Γ and τ satisfy the requirements (a) (see Fig. 5.1) and (b) (similarly) of Lemma 5.1. We prove that Γ and τ also satisfy (c) and (d).

Proof that Γ and τ satisfy (c). As in the case of the symbols P_{ij}, when referring to $\tau_{ij}^{\Box}$ or $\tau_{ij}^{\Diamond}$ we assume that subscripts are taken modulo 3.

First we show how to define the homomorphism f on $\mathbb{N} \times \{0\}$. Let $f(0,0) = c$ for an arbitrary element c of M satisfying P_{00}. Such c exists owing to τ_0 and $\tau_{ij}^{\Diamond}$. Assume that for some $i > 0$ we have defined $f(i-1, 0) = a$, and let a_h be an R-successor of a satisfying P_{i0}. Such a_h exists thanks to $\tau_{i-1,0}^{\Diamond}$. Define $f(i, 0) = a_h$.

Assume now that f is defined for $\mathbb{N} \times \{0, \ldots, j-1\}$ for some $j > 0$. We extend this definition to $\mathbb{N} \times \{j\}$. Let $f(0, j-1) = a$. By the inductive assumption a satisfies $P_{0,j-1}$. Choose a_v to be an R-successor of a satisfying P_{0j}. Such a_v exists by $\tau_{0,j-1}^{\Diamond}$. Set $f(0, j) = a_v$.

Assume that we have defined $f(i-1, j-1) = a$, $f(i-1, j) = a_v$, and $f(i, j-1) = a_h$. By the inductive assumptions $\mathfrak{M} \models P_{i-1,j-1}(a) \land P_{i-1,j}(a_v) \land P_{i,j-1}(a_h) \land aRa_h \land aRa_v$. Choose a_d to be an R-successor of a satisfying P_{ij}. Such a_d exists by $\tau_{i-1,j-1}^{\Diamond}$. We

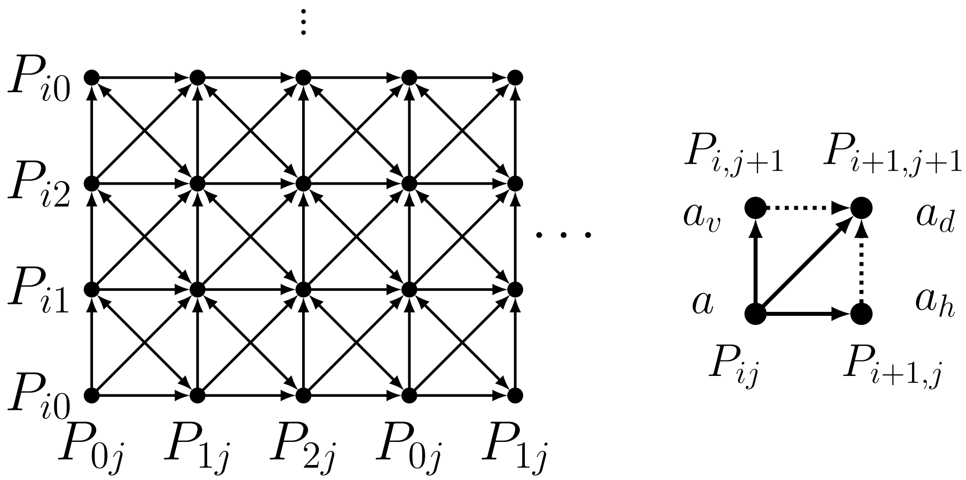

Figure 5.1.: A $\mathfrak{G}_\infty$-like structure. Reflexive arrows are omitted for clarity.

put $f(i,j) = a_d$. By $\tau^{\Box}_{ij}$, a_h, a_v and a_d cannot be connected to a, so Γ enforces R-connections between a_h and a_d, and between a_v and a_d. Since τ^{A}_{ij} forbids connection from a_d to a_h, and from a_d to a_v, it has to be that $\mathfrak{M} \models a_h R a_d \wedge a_v R a_d$. This finishes definition of f with the desired properties.

Proof that Γ and τ satisfy (d). We want to define for some $k, l \in \mathbb{Z}$ a function $f : \mathbb{Z}_k \times \mathbb{Z}_l \to M$ satisfying:

(a) $\mathfrak{M} \models P_{ij}(f(i,j))$,
(b) $\mathfrak{M} \models f(i,j) R f(i+1 \mod k, j)$,
(c) $\mathfrak{M} \models f(i,j) R f(i, j+1 \mod l)$.

We define f as a partial function on $\mathbb{N} \times \mathbb{N}$ and then restrict it to an appropriate domain. We first define f on $\mathbb{N} \times \{0\}$, exactly as in the proof of Part (i), $\Leftarrow$. Since $\mathfrak{M}$ this time is finite, it has to be that $f(k,0) = f(k',0)$ for some $k > k'$. To simplify the presentation we assume $k' = 0$, but this assumption is not relevant. Observe that for $i \in [0,k)$ we have $\mathfrak{M} \models f(i,0) R f(i+1 \mod k, 0)$. We extend the definition of f to $[0,k) \times \mathbb{N}$ inductively. Assume that f is defined on $[0,k) \times \{0, \ldots, j-1\}$. We define it on $[0,k) \times \{j\}$. For each $i \in [0,k)$ we find an element a^i_d in M such that $\mathfrak{M} \models P_{i+1,j}(a^i_d) \wedge f(i,j-1) R a^i_d$. Such a^i_d exists owing to $\tau^{\Diamond}_{i,j-1}$. We set $f(i+1 \mod k, j) = a^i_d$. Now Γ and formulas of the type $\tau^{\Box}$ enforce for all $i \in [0,k)$ that $\mathfrak{M} \models f(i,j-1) R f(i,j)$, and $\mathfrak{M} \models f(i,j) R f(i+1 \mod k, j)$.

Finiteness of $\mathfrak{M}$ implies now that for some $l > l'$ we have $f \restriction [0,k) \times \{l\} = f \restriction [0,k) \times \{l'\}$. Again for simplicity we assume that $l' = 0$. Observe that at this moment f is as desired on $\mathbb{Z}_k \times \mathbb{Z}_l$.

Finally, we extend f to $f' : \mathbb{Z}_m \times \mathbb{Z}_m \to M$ for $m = gcd(k,l)$ in the obvious way. The function f' is the required homomorphism.

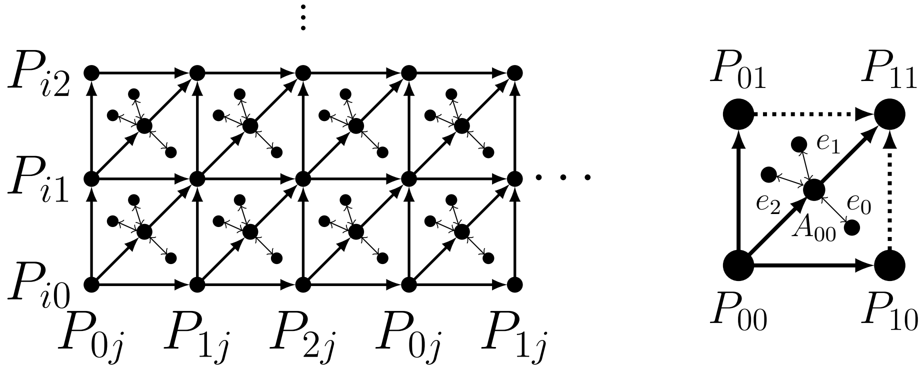

Figure 5.2.: An extension of $\mathfrak{G}_\infty$ for the inequality case.

5.3. Undecidability for unimodal logic defined by Horn formulas with inequality

Now we define a formula Γ_i and we prove that global $\mathcal{K}_{\Gamma_i}$-SAT is undecidable. In the proof we use the inequality only to say that the out-degree of a vertex is large. That is, we define an abbreviation $deg_{\geq k}(v)$ that uses the fresh variables $u_1^v, \ldots, u_k^v$ as follows.

$$deg_{\geq k}(v) = \bigwedge_{1 \leq i \leq k} (vRu_i^v) \wedge \bigwedge_{1 \leq i < j \leq k} u_i^v \neq u_j^v$$

For example, the formula $deg_{\geq 5}(v) \Rightarrow vRz$ says that all the worlds with out-degree greater than five are connected to all worlds.

Now, we are ready to define the formula Γ_i that gives us undecidability.

$$\Gamma_i = xRy \wedge xRu \wedge uRz \wedge deg_{\geq 2}(x) \wedge deg_{\geq 4}(u) \wedge deg_{\geq 2}(z) \Rightarrow yRz$$

The formula Γ_i contains only one Horn clause. Note that the structure illustrated in Fig. 5.2 is a model of Γ_i.

To get the undecidability we construct a modal formula τ_i such that Γ_i, τ_i satisfy the requirements of Lemma 5.1. Namely, τ_i says that:

(i) each world is labeled with exactly one of the variables from the set $\{P_{ij} | i, j \in \{0, 1, 2\}\} \cup \{A_{ij} | i, j \in \{0, 1, 2\}\} \cup \{e_{ij}^k | i, j, k \in \{0, 1, 2\}\}$.

(ii) every element satisfying P_{ij} has (at least) three R-successors: one in $P_{(i+1)j}$, one in $P_{i(j+1)}$, and one in A_{ij}, and each of its successors satisfies $P_{(i+1)j}$, $P_{i(j+1)}$, or A_{ij};

(iii) every element satisfying A_{ij} has four successors: one in $P_{(i+1 \mod 3)(j+1 \mod 3)}$, one in e^0_{ij}, one in e^1_{ij}, and one in e^2_{ij}, and each of its successors satisfies $P_{(i+1)(j+1)}$, e^0_{ij}, e^1_{ij}, or e^2_{ij}.

(iv) every element satisfying e^k_{ij} has a successor satisfying A_{ij} and and each of its successors satisfies A_{ij}.

All those properties are easy to express in modal logic. It is not hard to see that Γ, τ satisfy the requirements (a) (see Fig. 5.2) and (b) of Lemma 5.1.

Proof that Γ_i and τ_i satisfy (c). First we define the homomorphism f on $\mathbb{N} \times \{0\}$ exactly as in the three-variables case.

Assume now that f is defined for $\mathbb{N} \times \{0, \ldots, j-1\}$ for some $j > 0$. We extend this definition to $\mathbb{N} \times \{j\}$. Let $f(0, j-1) = a$. By the inductive assumption a satisfies $P_{0,j-1}$. Choose a_v to be an R-successor of a satisfying P_{0j}. Such a_v exists by $\tau^{\Diamond}_{0,j-1}$. Set $f(0, j) = a_v$.

Assume that we have defined $f(i-1, j-1) = a$, $f(i-1, j) = a_v$, and $f(i, j-1) = a_h$. By the inductive assumptions $\mathfrak{M} \models P_{i-1,j-1}(a) \wedge P_{i-1,j}(a_v) \wedge P_{i,j-1}(a_h) \wedge aRa_h \wedge aRa_v$. Choose a'_d to be an R-successor of a satisfying $A_{i-1,j-1}$ and a_d to be an R-successor of a'_d satisfying P_{ij}. Such a'_d, a_d exist by properties (ii) and (iii) of τ_i. Moreover, (iii) guarantees that a'_d has the out-degree grater than 3, an therefore Γ_i enforces R-connections from a_h to a_d, and from a_v to a_d. We put $f(i, j) = a_d$. This finishes definition of f with the desired properties.

Proof that Γ_i and τ_i satisfy (d) is similar to the corresponding proof for three variables case and the above proof and therefore we skip it.

5.4. Undecidability for bimodal logic defined by Horn formulas

We start from the semantics of the bimodal logic. The frame of bimodal logic is a triple $\langle W, R, R' \rangle$, where $R, R' \subseteq W^2$. The semantics of this logic is defined as for unimodal logic, but we have two more symbols $(\Diamond', \Box')$ and two more rules:

(vii) $\varphi = \Diamond' \varphi'$ and there is a world $v \in W$ such that $(w, v) \in R'$ and $\mathfrak{M}, v \models \varphi'$,

(viii) $\varphi = \Box' \varphi'$ and for all worlds $v \in W$ such that $(w, v) \in R'$ we have $\mathfrak{M}, v \models \varphi'$.

Observe that the proof of Lemma (5.1) works well for the case of bimodal logic. We define

$$\Gamma_b = zRx \wedge xRs \wedge zRu \wedge uR'y \Rightarrow xRy$$

Actually, for the global satisfiability case it is enough to consider simpler formula, namely $zRx \wedge zR'y \Rightarrow xRy$, but it is harder to extend the result for this formula for the local satisfiability case.

The formula τ_b says that:

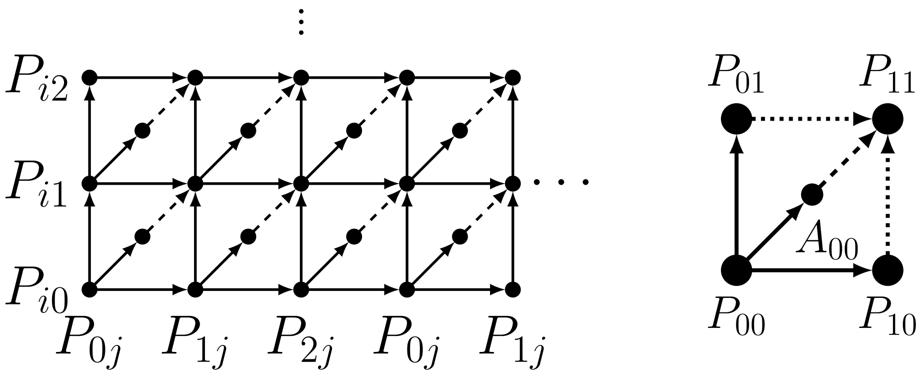

Figure 5.3.: An extension of $\mathfrak{G}_\infty$ for the bimodal case. Relation R' is marked by dashed lines.

(i) each element is labeled with exactly one of the variables from the set $\{P_{ij}|i,j \in \{0,1,2\}\} \cup \{A_{ij}|i,j \in \{0,1,2\}\}$.

(ii) every element satisfying P_{ij} has three R-successors: one in $P_{(i+1)j}$, one in $P_{i(j+1)}$, and one in A_{ij};

(iii) every element satisfying A_{ij} has no R successors and one R' successor in $P_{(i+1)(j+1)}$.

It is now not hard to see that Fig. 5.3 contains a model of Γ_i and τ_i and so that Γ_i and τ_i satisfy the assumption of Lemma 5.1.

5.5. Local satisfiability

Three variables case

Observe that our proof of the undecidability of global satisfiability over $\mathcal{K}_\Gamma$ works for the subclass of reflexive models. This allows us to use the trick from [19] to cover also the case of local satisfiability. We enforce by a modal formula the existence of an irreflexive world and, by a first-order formula, we make it connected to all reflexive worlds. Such a *universal world* can be then used to reach all relevant elements in the model. The class of structures is defined by a formula Γ', which says that each world with an incoming edge is reflexive and has an incoming edge from all irreflexive worlds, and enforces Γ for all reflexive worlds:

$$\Gamma' = \forall xyz.((xRy \land \neg zRz) \to (yRy \land zRy))\land \\ ((xRx \land yRy \land zRz) \to (\neg xRy \lor yRx \lor \neg xRz \lor zRx \lor yRz \lor zRy)).$$

In the modal formula we use a fresh symbol P_U to distinguish an irreflexive world. Now, for a given domino system $\mathcal{D}$ we can show that $P_U \wedge \Box\neg P_U \wedge \Diamond\top \wedge \Box(\tau \wedge \lambda^D)$ is locally (finitely) satisfiable over $\mathcal{K}_{\Gamma'}$ iff $\mathcal{D}$ covers $\mathbb{N} \times \mathbb{N}$ (some $\mathbb{Z}_k \times \mathbb{Z}_k$). This proves Theorem 3.4.

Inequality case

The trick from [19], that reduces the local satisfiability problem to the global one, requires a formula which is not a Horn formula, so we cannot use it. It turns out, however, that only a slight modification is needed. Observe that our proof of the undecidability of global satisfiability over $\mathcal{K}_{\Gamma_i}$ works for the subclass of models such that the out-degree of each world is bounded by four. Now, we enforce by a modal formula the existence of a universal world with out-degree (at least) 5 and, by a first-order formula, we make it connected to all worlds. As before, we used it to reach all relevant elements in the model.

$$\Gamma_i' = (deg_{\geq 5}(u) \wedge u \neq v \Rightarrow uRv) \wedge \Gamma_i$$

In the modal formula we use a fresh symbols $f_1, \ldots, f_5$ to guarantee that a world with the degree at least 5 exists. Now, for each modal formula φ we define its local version φ^l by $\bigwedge_{i\in\{1,\ldots,5\}} \Diamond f_i \wedge \bigwedge_{1\leq i<j\leq 5} \neg\Diamond(f_i \wedge f_j) \wedge \Box\varphi$ such that φ^l is locally satisfiable over $\mathcal{K}_{\Gamma_i'}$ iff φ is globally (finitely) satisfiable over $\mathcal{K}_{\Gamma_i}$.

Bimodal case

In this case, the proof for the local satisfiability bases on the fact that in figure presented in Fig. 5.3 no world has both R successors and R' successors. Now we require that all worlds with both R-successors and R'-successors is connected to all worlds with at lest one predecessor.

$$\Gamma_b' = \Gamma_b \wedge (xRu_1 \wedge xR'u_2 \wedge vRy \Rightarrow xRy) \wedge (xRu_1 \wedge xR'u_2 \wedge vR'y \Rightarrow xRy)$$

Note that Fig. 5.3 is a model of Γ'. To reduce $\mathcal{K}_\Gamma'$-SAT to $\mathcal{K}_\Gamma$-SAT, we simply replace a formula φ by $\varphi^l = \Diamond\top \wedge \Diamond'\top \wedge \Box\varphi$. Clearly, φ' can be satisfied only in a world that has both R successors and R' successors. Such a world has to be connected to all worlds which have at least one predecessors, and therefore φ has to be satisfied in all such worlds.

6. Model properties

6.1. Minimal tree-based models

In this section, we show that for every UHF formula Φ and every modal formula φ, if φ is $\mathcal{K}_\Phi$-satisfiable then it has a "nice" model. We start from an arbitrary $\mathcal{K}_\Phi$-based model $\mathfrak{M} \models \varphi$ and unravel it (using standard unraveling technique, as in [38] and [4]) into a model $\mathfrak{M}_0$ whose frame is a tree with the degree of its nodes bounded by $|\varphi|$. Clearly the frame of $\mathfrak{M}_0$ is not necessarily a member of $\mathcal{K}_\Phi$. In the next step, we add to $\mathfrak{M}_0$ the edges implied by the Horn clauses of Φ. This is performed in countably many stages, until the least fixed point is reached. We observe that the resulting structure, $\mathfrak{M}_\infty$, is still a model of φ, and its frame belongs to $\mathcal{K}_\Phi$.

Formally, we say that an edge (w, w') is a *consequence* of Φ in $\mathcal{M} = \langle W, R\rangle$, if for some worlds $v_1, \ldots, v_k \in W$ and $\Psi_1 \Rightarrow \Psi_2 \in \Phi$ we have $\mathcal{M} \models \Psi_1(w, w', v_1, \ldots, v_k)$, and $\Psi_2(w, w', v_1, \ldots, v_k) = wRw'$. We denote the set of all consequences of Φ in $\mathcal{M}$ by $\mathrm{C}^{\Phi}_{\leadsto}(\mathcal{M})$. We define the *consequence operator* as follows.

$\textsc{Cons}_{\Phi,W}(R) = R \cup \mathrm{C}^{\Phi}_{\leadsto}(\langle W, R\rangle)$

Now, the *closure operator* can be defined as the least fixed-point of *Cons*:

$\textsc{Closure}_{\Phi,W}(R) = \bigcup_{i>0} \textsc{Cons}^{i}_{\Phi,W}(R)$

Example 6.1. *Consider the tree $\langle W, R\rangle$ presented in Fig. 6.1 and $\Phi = \{xRz \wedge zRy \Rightarrow yRy, xRx \wedge xRy \wedge xRz \Rightarrow yRz\}$. Reflexive edges belong to $\textsc{Cons}_{\Phi,W}(R)$, dashed edges belong to $\textsc{Cons}^{2}_{\Phi,W}(R)$, and dotted edges belong to $\textsc{Cons}^{3}_{\Phi,W}(R)$. Quick check shows that $\textsc{Cons}^{3}_{\Phi,W}(R) = \textsc{Cons}^{4}_{\Phi,W}(R)$ and therefore $\textsc{Cons}^{3}_{\Phi,W}(R)$ is equal to $\textsc{Closure}_{\Phi,W}(R)$.*

For a tree $\mathcal{T} = \langle W, R\rangle$, we now define the *$\mathcal{T}$-based model of* Φ as $\mathfrak{C}_\Phi(\mathcal{T}) = \langle W, \textsc{Closure}_{\Phi,W}(R)\rangle$. We denote by Φ^+ the set of the clauses from Φ containing a positive literal, i.e. all clauses of Φ except those of the form $\Psi \Rightarrow \bot$. Note that $\mathfrak{C}_\Phi(\mathcal{T})$ is the smallest (w.r.t. inclusion of the set of edges) model of Φ^+ containing all edges from R. Of course, not all models can be obtained in this way. The following lemma shows, however, that we can restrict our attention to models that are *$\mathcal{T}$-based* for some tree $\mathcal{T}$ with bounded degree.

Lemma 6.2. *Let φ be a modal formula and let $\Phi \in$ UHF. If φ is $\mathcal{K}_\Phi$-satisfiable, then there exists a tree $\mathcal{T}$ with the degree bounded by $|\varphi|$ and a labeling $\pi_\mathcal{T}$, such that*

(i) $\langle \mathcal{T}, \pi_\mathcal{T}\rangle$ is a model of φ;

(ii) $\langle \mathfrak{C}_\Phi(\mathcal{T}), \pi_\mathcal{T}\rangle$ is a model of φ that satisfies Φ.

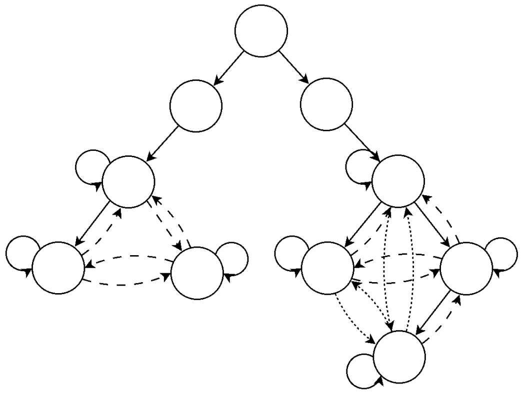

Figure 6.1.: A closure for $\Phi = \{xRz \wedge zRy \Rightarrow yRy, xRx \wedge xRy \wedge xRz \Rightarrow yRz\}$.

The result holds for the local satisfiability and for the global satisfiability.

The proof is in Section 6.6.

6.2. Definitions

We study several properties of models. The following frames will be useful.

Definition 6.3. *We define the linear structure $\mathcal{L}_\mathbb{Z}$ as $\langle\{\underline{i} : i \in \mathbb{Z}\}, \{(\underline{i}, \underline{i+1}) | i \in \mathbb{Z}\}\rangle$, and the infinite binary tree $\mathcal{T}_\infty$ as $\langle\{\underline{s} | s \in \{0,1\}^*\}, \{(\underline{s}, \underline{si}) | s \in \{0,1\}^* \wedge i \in \{0,1\}\}\rangle$. For each $s \in \mathbb{N}_\infty$, we define $\mathcal{I}_s = \mathcal{L}_{\mathbb{Z}\restriction W_s}$, where $W_s = \{\underline{i} | 0 \leq i < s\}$.*

The structures $\mathcal{L}_\mathbb{Z}$ and $\mathcal{T}_\infty$ play a crucial role in our proofs. We often reason in the following way. If for some $\mathcal{T}$ a property P is satisfied in a world of $\mathfrak{C}_\Phi(\mathcal{T})$, then we show that it is also satisfied in some world of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ or $\mathfrak{C}_\Phi(\mathcal{T}_\infty)$. Thanks to the uniformity of those structures, we show that the property P is satisfied in all $\mathfrak{g}(|\Phi|)$-proceeded worlds. Then we show that P has to be satisfied in all $\mathfrak{g}(|\Phi|)$-inner worlds of $\mathfrak{C}_\Phi(\mathcal{T})$.

Now we define our most important tool. We say that a function f from $\mathcal{M}_1$ into $\mathcal{M}_2$ is a *morphism* iff for all worlds w, w' if $\mathcal{M}_1 \models wRw'$, then $\mathcal{M}_2 \models f(w)Rf(w')$.

Observation 6.4. *Let $\mathcal{M}_1, \mathcal{M}_2$ be frames, let $\Phi \in$ UHF and let f be a function from $\mathcal{M}_1$ into $\mathcal{M}_2$. If f is a morphism from $\mathcal{M}_1$ into $\mathcal{M}_2$, then f is a morphism from $\mathfrak{C}_\Phi(\mathcal{M}_1)$ into $\mathfrak{C}_\Phi(\mathcal{M}_2)$.*

We use morphisms to transfer properties between $\mathfrak{C}_\Phi(\mathcal{T})$ and $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ or $\mathfrak{C}_\Phi(\mathcal{T}_\infty)$. One morphism, we often use, is $h_\mathcal{T} : \mathcal{T} \to \mathcal{L}_\mathbb{Z}$ defined as follows. For each v at the ith level of $\mathcal{T}$, $h_\mathcal{T}(v) = \underline{i}$. Now we define an important property that tells us whether a UHF formula enforces edges between different branches of a tree.

Definition 6.5. *We say that a formula $\Phi \in$ UHF* forks at the level i *if for all $\underline{s} \in \mathcal{T}_\infty$ with $|s| = i$ and $t, t' \in \{0,1\}^*$ there are no edges between $\underline{s0t}$ and $\underline{s1t'}$ in $\mathfrak{C}_\Phi(\mathcal{T}_\infty)$.*

We say that $\Phi \in \mathsf{UHF}$ has the tree-compatible model property *(TCMP) if for each i, Φ forks at the level i.*

It is not hard to see that if Φ has the tree-compatible model property, then in all tree-based models of Φ there are no edges among the worlds from disjoint subtrees. Indeed, if there is an edge between two different subtrees $\mathcal{S}_1, \mathcal{S}_2$ of a model $\mathcal{M}$, one can define a morphism from $\mathcal{M}$ to $\mathcal{T}_\infty$ which maps $\mathcal{S}_1$ and $\mathcal{S}_2$ into disjoint subtrees of $\mathcal{T}$. This implies that some world above $\mathcal{S}_1$ and $\mathcal{S}_2$ does not fork, and Φ does not have the tree-compatible model property.

In the next section, we study the linear structure $\mathcal{L}_\mathbb{Z}$, which turns out to be a good approximation of paths in trees. The formulas without the tree-compatible model property are discussed in Section 6.4.

6.3. The closures of linear structures

Now we study the possible shapes of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$. We say that the edge $(\underline{i}, \underline{j})$ is *forward* if $i < j$, *backward* if $i > j$, *short* if $|i - j| < 2$, and *long* if $|i - j| \geq 2$. We say that Φ *forces* long (resp. backward) edges if there is a long (resp. backward) edge in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and that Φ *forces only long forward edges* if it forces long edges but it does not force backward edges.

Definition 6.6. *We say that $\Phi \in \mathsf{UHF}$ satisfies*

S1 if Φ does not force long edges,

S2 if Φ forces only long forward edges and there exist $l, a_1, a_2, \ldots, a_l \in \mathbb{N}$ bounded by $\mathfrak{g}(|\Phi|)$ such that for all worlds $\underline{i}$, $\underline{i+b}$, there is an edge from $\underline{i}$ to $\underline{i+b}$ in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ if and only if $b \geq 0$ and $b - 1$ is in the additive closure of $\{a_1, a_2, \ldots, a_l\}$.

S3 if Φ forces long and backward edges and there exists m bounded by $\mathfrak{g}(|\Phi|)$ such that for all worlds $\underline{i}$, $\underline{i+b}$, there is an edge from $\underline{i}$ to $\underline{i+b}$ in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ if and only if m divides $|b - 1|$.

Properties S2 and S3 look complicated, so we present a few examples. Below we abbreviate $xRu_1 \wedge u_1Ru_2 \wedge \cdots \wedge u_{i-2}Ru_{i-1} \wedge u_{i-1}Ry$ by xR^iy.

Example 6.7. *Consider a formula $xR^2y \Rightarrow yRx$. Here, Property S3 is satisfied for $m = 3$. For example, $\underline{0}$ is connected to $\underline{1}$, $\underline{4}$, $\underline{7}$ and so on, while $\underline{2}$, $\underline{5}$, $\underline{8}$ and so on are connected to $\underline{0}$ (see Fig. 6.2a). In general, a formula $xR^iy \Rightarrow yRx$ satisfies Property S3 with $m = i + 1$.*

Example 6.8. *Consider a formula $\varphi_3 \wedge \varphi_4$, where $\varphi_i = xR^iy \Rightarrow xRy$. Here, Property S2 is satisfied for $l = 2$, $a_1 = 2$ and $a_2 = 3$. For example, $\underline{0}$ is connected to $\underline{1}$ (as in L_∞), $\underline{3}$ (because of φ_3), $\underline{4}$ (because of φ_4), $\underline{5}$ (because of φ_3, $\underline{0}R\underline{3}$, $\underline{3}R\underline{4}$, and $\underline{4}R\underline{5}$), and so on (see Fig. 6.2b). In general, for a formula of the form $\varphi_i \wedge \varphi_j$ Property S2 is satisfied with $l = 2$, $a_1 = i - 1$ and $a_2 = j - 1$.*

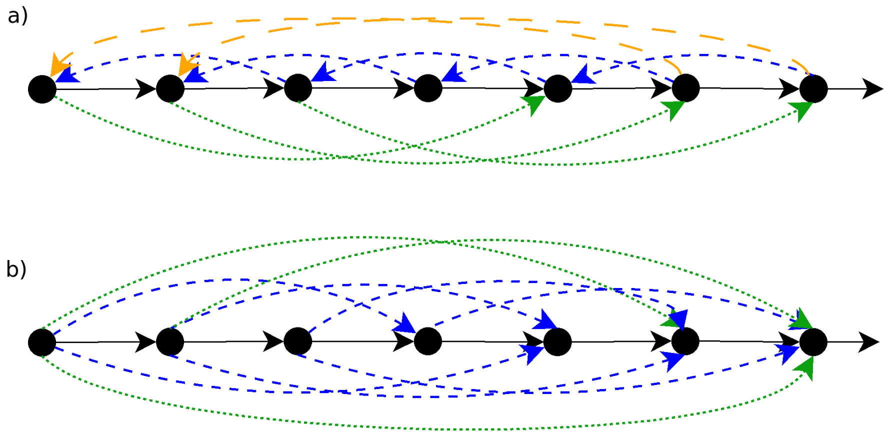

Figure 6.2.: Some closures for the linear structure.

It turns out that Properties S1, S2, and S3 cover all possible formulas.

Lemma 6.9. *Each* $\Phi \in \mathsf{UHF}$ *satisfies S1, S2, or S3.*

Now we show why these linear structures are important. In the tree-compatible case, along each path almost all worlds are connected as in the linear structure. The only exception is for the worlds that are close to the "end" of the model.

Lemma 6.10. *Let* $\Phi \in \mathsf{UHF}$, $\mathcal{T}$ *be a tree and* v_i, v_j *be* $\mathfrak{g}(|\Phi|)$*-inner worlds at the same path. Then there is an edge from* v_i *to* v_j *in* $\mathfrak{C}_\Phi(\mathcal{T})$ *if and only if there is an edge from* $h_\mathcal{T}(v_i)$ *to* $h_\mathcal{T}(v_j)$ *in* $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$*.*

6.4. Forks

In this section, we study models of the formulas without the tree-compatible model property.

Lemma 6.11. *If* $\Phi \in \mathsf{UHF}$ *forks at level* $\mathfrak{g}(|\Phi|)$*, then it has the tree-compatible model property.*

This lemma says that if there is a world that does not fork, then no world below some level forks.

We say that two worlds w, w' of a frame $\mathcal{M}$ are *equivalent* if for each world u we have uRw iff uRw'. Now we argue that if Φ does not fork at the level i, then in structures reachable from worlds at the level i such equivalence is very common:

Lemma 6.12. *Let* $\Phi \in \mathsf{UHF}$ *be a formula that does not fork,* $\mathcal{T}$ *be a tree with a bounded degree and* w *be a world at level* $k = (|\Phi| + 1)\mathfrak{g}(|\Phi|)$ *in* $\mathfrak{C}_\Phi(\mathcal{T})$*. Then for* $n = 2k + 1$ *and all* i*, all the* n*-followed descendants of* w *at level* $n + i$ *are equivalent in the frame* $\mathfrak{C}_\Phi(\mathcal{T})$*.*

Example 6.13. *Consider the formula $\Phi = \{\varphi_1, \varphi_2\}$, where $\varphi_1 = xRz \wedge zRy \Rightarrow yRy$ and $\varphi_2 = xRx \wedge xRy \wedge xRz \Rightarrow yRz$, and the tree in Fig. 6.1. The formula φ_1 enforces the following property: each world that has a predecessor that has a predecessor is reflexive. The formula φ_2 makes the relation R Euclidean except for the non-reflexive worlds. Formula Φ forks at the first two levels.*

6.5. Boundedness

The properties defined above are enough to prove the decidability, but not to obtain the optimal complexity.

We say that a formula Φ is *bounded* if $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ is not a model of Φ, and unbounded otherwise. If the formula is bounded, then there is a k such that the length of each path in each model of Φ is bounded by k, and the value of k depends only on Φ. Recall that in problems $\mathcal{K}_\Phi$-SAT and global $\mathcal{K}_\Phi$-SAT formula Φ is not a part of input. Hence the exact value of k is irrelevant, since it is regarded as a constant.

Example 6.14. *Consider any $n \in \mathbb{N}$ and the formula $\Phi_n = x_1Rx_2 \wedge x_2Rx_3 \wedge \cdots \wedge x_{n-1}Rx_n \Rightarrow \bot$. It is not hard to see that a given structure is a model of Φ_n if and only if it does not contain a path of the length n.*

Now we prove the polynomial model property for bounded formulas. The following argument works for local and global satisfiability.

Let Φ be a bounded formula and φ be a modal formula. Then for any model $\mathfrak{M} = \langle W, R, \pi \rangle$ of φ and Φ, we can find a $W' \subseteq W$ such that $\mathfrak{M}_{\restriction W'}$ is a model of φ and $|W'|$ is polynomial in $|\varphi|$. At first, we add an arbitrary world from $\mathfrak{M}$ that satisfies φ to W'. Then, recursively, for each world w in W' and each subformula $\Diamond\psi$ of φ, if w has a witness for $\Diamond\psi$ in W but not in W', then we add one such witness to W'. We proceed until a fixed-point is reached. Observe that since the length of each path is bounded by k, then this procedure takes at most k recursive steps, and in each of them, it adds at most $|\varphi|$ worlds for each element of W'. Therefore, at the end we have $|W'| = |\varphi|^k$ and $\mathfrak{M}_{\restriction W'}$ is a model of φ, so indeed we find a polynomial model of φ. Of course, since Φ is universal, $\langle W', R_{\restriction W'} \rangle$ satisfies Φ.

Clearly, the polynomial model property leads to a straightforward nondeterministic algorithm that guesses a model and verifies it.

Lemma 6.15. *If Φ is a bounded* UHF *formula, then $\mathcal{K}_\Phi$–SAT is in* NP.

6.6. Omitted proofs

6.6.1. Proof of Lemma 6.2

We start with an auxiliary lemma.

Lemma 6.16. *Let $\mathfrak{M}_1 = \langle W, R_1, \pi \rangle$ and $\mathfrak{M}_2 = \langle W, R_2, \pi \rangle$ be two Kripke structures sharing the universe and labeling, and such that $R_1 \subseteq R_2$. Let φ be a modal formula.*

If for each (u,v) in $R_2 \setminus R_1$ and each subformula of ϕ of the form $\Diamond\psi$ such that $\mathfrak{M}_2, v \models \psi$ there exists v' such that $(u,v') \in R_1$ and $\mathfrak{M}_2, v' \models \psi$, then for each $v \in W$ we have $tp_{\mathfrak{M}_1}(v) = tp_{\mathfrak{M}_2}(v)$.

Proof. We want to show that for each subformula ψ of φ and $v \in W$ we have $\mathfrak{M}_1, v \models \psi$ iff $\mathfrak{M}_2, v \models \psi$. The proof goes by structural induction on ψ. The cases where ψ is of the form P, $\psi_1 \vee \psi_2$ and $\neg\psi'$ are straightforward.

Suppose that $\psi = \Diamond\psi'$ and $v \in W$. If there exists a world w such that $(v,w) \in R_1$ and $\mathfrak{M}_1, w \models \psi'$, then by the inductive hypothesis we know that $\mathfrak{M}_2, w \models \psi'$ and therefore $\mathfrak{M}_2, v \models \psi$.

If there exists a world w such that $(v,w) \in R_2$ and $\mathfrak{M}_2, w \models \psi'$, then by the assumptions and the inductive hypothesis we know that there exists $w' \in W$ such that $(v,w') \in R_1$ and $\mathfrak{M}_1, w \models \psi'$, so $\mathfrak{M}_1, v \models \psi$.

If $\psi = \Box\psi'$ then, similarly, using the inductive hypothesis for all its successors we see that $\mathfrak{M}_1, v \models \psi$ iff $\mathfrak{M}_2, v \models \psi$. $\square$

Now we are ready to prove Lemma 6.2.

Proof. Assume that there exists $\mathfrak{M} = \langle W, R, \pi \rangle$ and $u_0 \in W$ such that $\mathfrak{M} \models \Phi$ and $\mathfrak{M}, u_0 \models \varphi$.

We construct $\mathfrak{M}_0 = \langle \mathcal{T}, \pi_{\mathcal{T}} \rangle$, where $\mathcal{T} = \langle W_0, R_0 \rangle$, by an unraveling of $\mathfrak{M}$ as follows. W_0 is a subset of the set of finite sequences of elements of W. We define W_0 and R_0 inductively. Initially, we put $(u_0) \in W_0$. Assume that $(u_0, \ldots, u_k) \in W_0$. Let $\Diamond\psi_1, \ldots, \Diamond\psi_s$ be all formulas of the form $\Diamond\psi$ from $tp_{\mathfrak{M}}(u_k)$. There exist $u_{k+1}^1, \ldots, u_{k+1}^s \in W$, such that for every $i \in \{1, \ldots, s\}$ we have $\mathfrak{M} \models u_k R u_{k+1}^i$ and $\psi_i \in tp_{\mathfrak{M}}(u_{k+1}^i)$. For each such i we put $(u_0, \ldots, u_k, u_{k+1}^i)$ into W_0 and add $((u_0, \ldots, u_k),\ (u_0, \ldots, u_k, u_{k+1}^i))$ to R_0. Define $\pi_{\mathcal{T}}$ as $\pi_{\mathcal{T}}((u_0, \ldots, u_k)) = \pi(u_k)$. Observe that $\mathcal{T}$ is a tree in which the degree of every node is bounded by $|\varphi|$.

Let $f : W_0 \to W$ be defined as $f((u_0, \ldots, u_k)) = u_k$. By a straightforward induction the reader may verify that, for every $\vec{u} \in W_0$ we have $tp_{\mathfrak{M}_0}(\vec{u}) = tp_{\mathfrak{M}}(f(\vec{u}))$. This implies that $\mathcal{T}, (u_0) \models \varphi$.

Now, in countably many stages we add to $\mathcal{T}$ the edges implied by Φ. We put $\mathcal{M}_0 = \mathcal{T}$ and we define the sequence of frames $(\mathcal{M}_i)_{i>0}$ and models $(\mathfrak{M}_i)_{i>0}$. The frames $(\mathcal{M}_i)_{i>0}$ share the universe W_0 and the structures $(\mathfrak{M}_i)_{i>0}$ share the universe W_0 and the mapping $\pi_{\mathcal{T}}$. For $K > 0$ let $\mathcal{M}_K = \langle W_0, \textsc{Cons}^K_{\Phi, W_0}(R_0) \rangle$, $\mathfrak{M}_K = \langle \mathcal{M}_K, \pi_{\mathcal{T}} \rangle$. Let $\mathfrak{M}_\infty$ be the natural limit $\mathfrak{M}_\infty = \langle \mathfrak{C}_\Phi(\mathcal{M}_0), \pi_{\mathcal{T}} \rangle$.

We show by induction on K, that for each $\vec{u}_1, \vec{u}_2 \in W_0$ if $\mathfrak{M}_K \models \vec{u}_1 R \vec{u}_2$, then $\mathfrak{M} \models f(\vec{u}_1) R f(\vec{u}_2)$. It follows that for each $\vec{u}_1, \vec{u}_2 \in W_0$ if $\mathfrak{M}_\infty \models \vec{u}_1 R \vec{u}_2$, then $\mathfrak{M} \models f(\vec{u}_1) R f(\vec{u}_2)$.
For $K = 0$ the conclusion is a straightforward consequence of the definition of $\mathfrak{M}_0$. Assume that $\mathfrak{M}_K$ satisfies the inductive hypothesis. For each $\vec{u}_1, \vec{u}_2 \in W_0$, if $\mathfrak{M}_{K+1} \models \vec{u}_1 R \vec{u}_2$, then either $\mathfrak{M}_K \models \vec{u}_1 R \vec{u}_2$ and by the inductive assumption $\mathfrak{M} \models f(\vec{u}_1) R f(\vec{u}_2)$, or for some $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k} \in W_0$ and $\Psi_1 \Rightarrow \Psi_2 \in \Phi$, we have $\mathfrak{M}_K \models \Psi_1(\vec{u_1}, \vec{u_2}, \vec{v_1}, \vec{v_2}, \ldots, \vec{v_k})$, and $\Psi_2(\vec{u_1}, \vec{u_2}, \vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}) = \vec{u_1} R \vec{u_2}$. In this case, $\mathfrak{M}_K \models \Psi_1(\vec{u_1}, \vec{u_2}, \vec{v_1}, \vec{v_2}, \ldots, \vec{v_k})$ implies by the inductive assumption that

$\mathfrak{M} \models \Psi_1(f(\vec{u_1}), f(\vec{u_2}), f(\vec{v_1}), f(\vec{v_2}), \ldots, f(\vec{v_k}))$. Since $\mathfrak{M} \models \Psi_1 \Rightarrow \Psi_2$, we have $\mathfrak{M} \models f(\vec{u}_1)Rf(\vec{u}_2)$.

Let $\mathfrak{M}_\infty = \langle W_0, R_\infty, \pi_\mathcal{T}\rangle$. The frames $\mathfrak{M}_0$ and $\mathfrak{M}_\infty$ have the same universe and $R_0 \subseteq R_\infty$. We show that for each $\vec{u} \in W_0$ we have $tp_{\mathfrak{M}_\infty}(\vec{u}) = tp_{\mathfrak{M}_0}(\vec{u})$. This implies that $\mathfrak{M}_\infty, (u_0) \models \varphi$. Since the labeling of the worlds is the same, it is enough to show that in $\mathfrak{M}_0$ and $\mathfrak{M}_\infty$ each world is connected with the worlds that satisfy the same subformulas. We show that by induction w.r.t. the size of subformula ψ of φ.

Clearly, for every edge $(\vec{u}, \vec{v})$ from $R_\infty \setminus R_0$ and a subformula $\Diamond\psi$ of φ, if a world $\vec{v}$ satisfies ψ in $\mathfrak{M}_\infty$, then by the inductive assumption we have $\psi \in tp_{\mathfrak{M}_0}(\vec{v}) = tp_{\mathfrak{M}}(f(\vec{v}))$, and since $\mathfrak{M} \models f(\vec{u})Rf(\vec{v})$ we have that $\Diamond\psi \in tp_{\mathfrak{M}}(f(\vec{u})) = tp_{\mathfrak{M}_0}(\vec{u})$ (by Lemma 6.16 applied to $\mathfrak{M}_1 = \mathfrak{M}_0$ and $\mathfrak{M}_2 = \mathfrak{M}_\infty$).

Finally, we have to prove that $\mathfrak{C}_\Phi(\mathcal{M}_0) \models \Phi$. By definition $\mathfrak{C}_\Phi(\mathcal{M}_0)$ satisfies every $\Psi_1 \Rightarrow \Psi_2 \in \Phi^+$. Suppose that $\mathfrak{C}_\Phi(\mathcal{M}_0)$ does not satisfy $\Psi \Rightarrow \perp \in \Phi$. For some $\vec{w_1}, \vec{w_2}, \vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ we have $\mathfrak{C}_\Phi(\mathcal{M}_0) \models \Psi(\vec{w_1}, \vec{w_2}, \vec{v_1}, \vec{v_2}, \ldots, \vec{v_k})$, but then $\mathfrak{M} \models \Psi(f(\vec{w_1}), f(\vec{w_2}), f(\vec{v_1}), f(\vec{v_2}), \ldots, f(\vec{v_k}))$. This contradicts the assumption that $\mathfrak{M} \models \Phi$. □

6.6.2. Proof of Lemma 6.9

Let us fix $\Phi \in \mathsf{UHF}$. The following definition will prove useful in the sequel. For $s \in \mathbb{Z}$ we define *the shift function* sh_s as $sh_s(\underline{i}) = \underline{i+s}$. Let $\mathcal{M}$ be a frame containing $\mathcal{L}_\mathbb{Z}$ over the same universe, i.e. $\{\underline{i} : i \in \mathbb{Z}\}$. We say that $\mathcal{M}$ is *uniform* if for every $s \in \mathbb{Z}$ the shift sh_s is an automorphism of $\mathcal{M}$. We say that $\mathcal{M}$ is *closed under composition* iff for every world $\underline{i}$, positive k and $a_1, a_2, \ldots, a_k \in \mathbb{Z}$:

1. if $\mathcal{M} \models \underline{i}R\underline{i+k}$ and $\mathcal{M} \models \underline{i}R\underline{i+a_1} \wedge \ldots \wedge \underline{i+a_{k-1}}R\underline{i+a_k}$ then $\mathcal{M} \models \underline{i}R\underline{i+a_k}$, and
2. if $\mathcal{M} \models \underline{i}R\underline{i-k}$ and $\mathcal{M} \models \underline{i+a_k}R\underline{i+a_{k-1}} \wedge \ldots \wedge \underline{i+a_1}R\underline{i}$ then $\mathcal{M} \models \underline{i}R\underline{i+a_k}$.

A frame $\mathcal{N}$ is the *composite closure* of $\mathcal{M}$ iff $\mathcal{N}$ is the least (w.r.t. the relation R) closed under composition frame containing $\mathcal{M}$.

Lemma 6.17. *Frame $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ is uniform and closed under composition.*

Proof. Clearly, for every $s \in \mathbb{Z}$, the shift sh_s is an automorphism of $\mathcal{L}_\mathbb{Z}$ onto itself, and due to Observation 6.4, sh_s is a morphism of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ onto itself and it is a bijection. Hence, sh_s is an automorphism of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ is uniform.

If $\mathcal{M} \models \underline{i}R\underline{i+k}$ and $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{i}R\underline{i+a_1} \wedge \underline{i+a_1}R\underline{i+a_2} \wedge \ldots \wedge \underline{i+a_{k-1}}R\underline{i+a_k}$, then the function g defined as

$$g(\underline{j}) = \begin{cases} \underline{j} & \text{if } j \leq i \\ \underline{i+a_{j-i}} & \text{if } i < j \leq i+k \\ \underline{j+a_k} & \text{otherwise} \end{cases}$$

is a morphism of $\mathcal{L}_\mathbb{Z}$ into $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$. Thus, by Observation 6.4, g is a morphism of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ into $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$. Since $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{i}R\underline{i+k}$, the morphism g implies that $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{i}R\underline{i+a_k}$.

Similarly, if $\mathcal{M} \models \underline{i}R\underline{i-k}$ and $\mathcal{M} \models \underline{i+a_k}R\underline{i+a_{k-1}} \wedge \ldots \wedge \underline{i+a_1}R\underline{i}$, then a function

$$g(\underline{j}) = \begin{cases} \underline{j} & \text{if } j < i + a_k \\ \underline{i + a_{k-(j-i)}} & \text{if } i + a_k \leq j < i + a_k + k \\ \underline{j - a_k - k} & \text{otherwise} \end{cases}$$

is again a morphism of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ into $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and since $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{i}R\underline{i-k}$, the morphism g implies that $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{i}R\underline{i+a_k}$. □

For any uniform and closed under composition $\mathcal{N}$ we define $X_\Phi^{\mathcal{N},i} = \{a : \mathcal{N} \models \underline{i}R\underline{i+a+1}\}$ and $X_\Phi^{\mathcal{N}} = X_\Phi^{\mathcal{N},0}$. Since the function sh_s is an automorphism of $\mathcal{N}$, for any i we have $X_\Phi^{\mathcal{N},i} = X_\Phi^{\mathcal{N}}$.

Lemma 6.18. *Let $\mathcal{N}$ be a uniform and closed under composition structure.*

(i) If $x, y \in X_\Phi^{\mathcal{N}}$ and $x \geq 0$, then $x + y \in X_\Phi^{\mathcal{N}}$.

(ii) If $x, y \in X_\Phi^{\mathcal{N}}$ and $s, x \geq 0$, then $x + sy \in X_\Phi^{\mathcal{N}}$.

(iii) For every $a > 2$, if $-a \in X_\Phi^{\mathcal{N}}$, then $a \in X_\Phi^{\mathcal{N}}$.

(iv) If $X_\Phi^{\mathcal{N}}$ contains any positive number, then for all $a \geq 0$, if $-a \in X_\Phi^{\mathcal{N}}$, then $a \in X_\Phi^{\mathcal{N}}$.

Proof. For the (i) part, observe that if $\mathcal{N} \models \underline{0}R\underline{x+1}$ and $\mathcal{N} \models \underline{0}R\underline{y+1}$, where $x > 0$, then $\mathcal{N} \models (\underline{0}R\underline{y+1}) \wedge (\underline{y+1}R\underline{y+2}) \wedge \ldots (\underline{i+y+x}R\underline{i+y+x+1})$. Hence, by composite closure $\mathcal{N} \models \underline{i}R\underline{i+x+y+1}$ and $x + y \in X_\Phi^{\mathcal{N}}$. Property (ii) follows from a straightforward induction based on (i).

For the (iii) part, note that if $\mathcal{N} \models \underline{0}R\underline{-(a-1)}$ where $a > 2$, then due to uniformity we have in $\mathcal{N}$ that $\underline{a+1}R\underline{2}$ and $\underline{2}R\underline{-a+3}$. Of course, $\underline{-a+3}R\underline{-a+2}$, $\underline{-a+2}R\underline{-a+1}$, …, $\underline{-1}R\underline{0}$. As $\mathcal{N}$ is closed under composition, it implies that $\mathcal{N} \models \underline{0}R\underline{a+1}$.

Finally, for the (iv) part, let $b \in X_\Phi^{\mathcal{N}}$ be a positive number and $a \geq 0$. If $a = 1$, then we by Property (ii) for $x = b$, $y = -1$ and $s = b - 1$ we have $a \in X_\Phi$. If $a = 2$, then consider two cases. If b is odd, we use Property (ii) for $x = b$, $y = a$ and $s = (b+1)/2$ to show that $1 \in X_\Phi^{\mathcal{N}}$ and, by (i), that $2 = 1 + 1 \in X_\Phi^{\mathcal{N}}$. If b is even, use Property (ii) for $x = b$, $y = a$ and $s = b/2 - 1$ to show that $2 \in X_\Phi^{\mathcal{N}}$. If $a > 2$, then the statement follows from Property (iii). □

Now we consider the case when $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ contains long backward edges.

Lemma 6.19. *Suppose that a uniform and closed under composition structure $\mathcal{N}$ contains a backward edge and a long edge. There exists a finite subset X' of $X_\Phi^{\mathcal{N}}$ such that $gcd(X') = gcd(X_\Phi^{\mathcal{N}})$ and, for every b, $\mathcal{N} \models \underline{i}R\underline{i+b+1}$ iff $gcd(X')$ divides b.*

Proof. Due to Lemma 6.18 (iv), for any positive b, if $-b \in X^{\mathcal{N}}_{\Phi}$, then $b \in X^{\mathcal{N}}_{\Phi}$. This implies that $gcd(X^{\mathcal{N}}_{\Phi}) = gcd(X^{+})$, where X^{+} contains all positive elements from $X^{\mathcal{N}}_{\Phi}$. Thanks to Fact 4.1, there exists finite X' such that for every $b > lcm(X')$ we have $\mathcal{N} \models \underline{i}R\underline{i+b+1}$ iff $gcd(X')$ divides b. We will show that for every $b \in \mathbb{Z}$, we have $b \in X^{\mathcal{N}}_{\Phi}$ iff $gcd(X')$ divides b.

Of course, all elements of $X^{\mathcal{N}}_{\Phi}$ are divisible by $gcd(X')$. Suppose that $b \in \mathbb{Z}$ is divisible by $gcd(X')$. Sine $\mathcal{N}$ contains backward edges and long edges, there exists $a_j \in X^{\mathcal{N}}_{\Phi}$ such that $a_j < -1$. From Lemma 6.18 we know that $-a_j \in X^{\mathcal{N}}_{\Phi}$ and, moreover, for any $s \geq 0$, $b + s \cdot (-a_j) \in X^{\mathcal{N}}_{\Phi}$. Let s be such that $b' = b + s \cdot (-a_j) > lcm(X')$. Since $gcd(X')$ divides b and $-a_j$, $b' \in X^{\mathcal{N}}_{\Phi}$. Due to additivity of $X^{\mathcal{N}}_{\Phi}$, we conclude that $b = (b' + s \cdot a_j) \in X^{\mathcal{N}}_{\phi}$. □

We say that a structure $\mathcal{M}$ is t, p-regular is for all $b > 0$

- if $\mathcal{M} \models \underline{i}R\underline{i+b+1}$ then p divides b;
- if $b > t$ and p divides b, then $\mathcal{M} \models \underline{i}R\underline{i+b+1}$.

We define the *period* of $\mathcal{M}$, denoted by $pi(\mathcal{M})$, in the following way. If there exists $t < \infty$ and p such that $\mathcal{M}$ is t, p regular, then we put $pi(\mathcal{M}) = p$ (note that such p is always unique). Otherwise, we put $p = 0$. Moreover, we define the *threshold* $tr(\mathcal{M}) \in \mathbb{N}_{\infty}$ as the least number t such that $\mathcal{M}$ is $t, pi(\mathcal{M})$-regular. Note that $pi(\mathcal{M}) = 0$ if and only if $tr(\mathcal{M}) = \infty$.

For a given two worlds $\underline{i}$, $\underline{j}$, where $i, j \in \mathbb{Z}$, we define a *distance* between $\underline{i}$ and $\underline{j}$, $dt(\underline{i}, wj)$, as $|i - j|$. The following lemma explains that the threshold is an important characteristic of a frame.

Lemma 6.20. *Let $\mathcal{M}$ be a closed under composition, uniform frame with $tr(\mathcal{M}) < \infty$ and $u_1, \ldots, u_k$ be a sequence of integers. There is a sequence $v_1, \ldots, v_k$ such that for every $i, j \in \{1, \ldots, k\}$*

1. *$\mathcal{M} \models \underline{u_i}R\underline{u_j}$ iff $\mathcal{M} \models \underline{v_i}R\underline{v_j}$,*
2. *$u_i \leq u_j$ iff $v_i \leq v_j$, and*
3. *$|v_i - v_j| \leq k \cdot tr(\mathcal{M}) + pi(\mathcal{M})$.*
4. *$|v_i - v_j| \equiv |u_i - u_j| \mod pi(\mathcal{M})$ and $|v_i - v_j| \geq tr(\mathcal{M})$ iff $|u_i - u_j| \geq tr(\mathcal{M})$.*

Moreover, for any $a, b \in \{1, \ldots, k\}$ there is a sequence $v'_1, \ldots, v'_k$ such that $v'_a = u_a$, $v'_b = u_b$ and for every $i, j \in \{1, \ldots, k\}$ $\mathcal{M} \models u_iRu_j$ iff $\mathcal{M} \models v'_iRv'_j$ and $min(dt(v'_i, v'_a), dt(v'_i, v'_b)) < (k-1) \cdot tr(\mathcal{M}) + pi(\mathcal{M})$.

Proof. Without loss of generality, assume that the sequence $u_1, \ldots, u_k$ is ascending. Let j be such $u_{j+1} - u_j > tr(\mathcal{M}) + pi(\mathcal{M})$. Define $u'_{j+i} = u_{j+i} - pi(\mathcal{M})$ for all $1 \leq i \leq k - j$. Since $\mathcal{M}$ is uniform, the relations among $\underline{u'_{j+1}}, \ldots, \underline{u'_n}$ are the same as the relations among $\underline{u_{j+1}}, \ldots, \underline{u_n}$.

Consider any $i \leq j$ and $i' > j$. Since $u'_{i'} - u_i - 1 \geq tr(\mathcal{M})$, there following statements are equivalent:

(i) there is an edge between $\underline{u'_{i'}}$ and $\underline{u_i}$;

(ii) $pi(\mathcal{M})$ divides $u'_{i'} - u_i - 1$;

(iii) $pi(\mathcal{M})$ divides $u'_{i'} - u_i - 1 + pi(\mathcal{M}) = u_{i'} - u_i - 1$;

(iv) there is an edge between $\underline{u_{i'}}$ and $\underline{u_i}$.

Clearly, $u_1, u_2, \ldots, u_j, u'_{j+1}, \ldots, u'_k$ satisfies 1 and 2, and $u'_{j+1} - u_j < u_{j+1} - u_j$. By iterating the operation defined above finitely many times we obtain sequences that satisfy the required properties. □

Let $W = \{\underline{i} : i \in \mathbb{Z}\}$ be a domain of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and R be any set of edges over W. We define R^s as a number such that $\langle \underline{0}, \underline{R^s} \rangle \in Cons_\Phi(R) \setminus R$ and for each b such that $\langle \underline{0}, \underline{b} \rangle \in Cons_\Phi(R) \setminus R$, $|b| \geq |R^s|$, and if $|b| = |R^s|$, then $b \leq R^s$. If $Cons_\Phi(R) \setminus R = \emptyset$, we put $R^s = 1$.

We define the *sequence of approximations* $\mathcal{N}_0, \mathcal{N}_1, \ldots$ of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ in the following way:

- $\mathcal{N}_0 = \mathcal{L}_\mathbb{Z} = \langle W, R_0 \rangle$ and
- $\mathcal{N}_{p+1} = \langle W, R_{p+1} \rangle$ is the compositely closure of $\langle W, R_p \cup \{(\underline{i}, \underline{i + R^s_p}) | i \in \mathbb{Z}\} \rangle$.

Clearly, the limit of the sequence $\mathcal{N}_0, \mathcal{N}_1, \ldots$ is $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and all approximations are closed under composition and uniform.

Lemma 6.21. *Let $\mathcal{N}_0, \mathcal{N}_1, \ldots$ be the sequence of approximations of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and n be the number of variables in Φ. For every i, if $\mathcal{N}_p^\Phi$ contains long edges, then $\mathcal{N}_p^\Phi$ has the threshold bounded by $n^{3(\log_2(n) - \log_2(pi(\mathcal{N}_p)) + 1)}$ and the period bounded by n. Moreover, the sequence stabilizes after some finite index.*

Proof. We prove the lemma by induction.

The induction base. If $\mathcal{N}_p$ is the first frame containing a long edge, then $\mathcal{N}_{p-1}$ is equal to $\mathcal{L}_\mathbb{Z}$ or a reflexive or symmetric closure (or both) of $\mathcal{L}_\mathbb{Z}$. A quick check shows that in this case both the threshold and period of $\mathcal{N}_p^\Phi$ are bounded by n.

The induction step. Suppose that $R^s_p > n \cdot tr(\mathcal{N}_p) + pi(\mathcal{N}_p)$. The edge $(\underline{0}, \underline{R^s_p})$ is implied by the formula Φ applied to some worlds $\underline{u_1}, \ldots, \underline{u_n}$ with $u_s = 0$ and $u_t = R^s_p$ for some s, t. Let $v_1, \ldots, v_n$ be a result of application of Lemma 6.20 to the sequence $u_1, \ldots, u_n$ and the frame $\mathcal{N}_p$. Since the connections among $\underline{v_1}, \ldots, \underline{v_n}$ are the same as among $\underline{u_1}, \ldots, \underline{u_n}$, the edge $(\underline{v_s}, \underline{v_t})$ is a consequence of Φ in $\mathcal{N}_p$. If $(\underline{v_s}, \underline{v_t})$ were an edge in $\mathcal{N}_p$, then also $(\underline{0}, \underline{R^s_p})$ would, because of Property 4 of Lemma 6.20. But it is not the case by the definition of R^s_p. The existence of the edge $(\underline{v_s}, \underline{v_t})$ contradicts the minimality of R^s_p. Therefore $R^s_p \leq n \cdot tr(\mathcal{N}_p) + pi(\mathcal{N}_p)$.

Assume that $R^s_p > 0$.

Let $j = R^s_p - 1$. If j is divisible by $pi(\mathcal{N}_i)$, then $\mathcal{N}_{p+1}$ is $tr(\mathcal{N}_p), pi(\mathcal{N}_p)$-regular, and therefore $tr(\mathcal{N}_{p+1}) \leq tr(\mathcal{N}_p)$ and $pi(\mathcal{N}_{p+1}) = pi(\mathcal{N}_p)$.

Assume otherwise, that j is not divisible by $pi(\mathcal{N}_p)$.

Let $P = gcd(pi(\mathcal{N}_p), j)$ and $T = tr(\mathcal{N}_p) + j \cdot pi(\mathcal{N}_p)$. Every number divisible by P from the interval $[0, pi(\mathcal{N}_p) - 1]$ is the remainder of some number form $\{j, 2 \cdot j, \ldots, pi(\mathcal{N}_p) \cdot j\}$ from division by $pi(\mathcal{N}_p)$. Hence, every number $b \geq 0$ divisible by P is equal to $\alpha j + \beta pi(\mathcal{N}_p)$ where $\alpha \in \{1, \ldots, pi(\mathcal{N}_p)\}$ and $\beta \in \mathbb{Z}$. If $b > T$, then $\beta pi(\mathcal{N}_p) > tr(\mathcal{N}_p)$ and $\beta pi(\mathcal{N}_p) \in X_\Phi^{\mathcal{N}_p}$. Thus, for every $b \geq T$, if P divides b, then $b \in X_\Phi^{\mathcal{N}_{p+1}}$. On the other hand, the frame $\mathcal{M} = \langle \{\underline{i} : i \in \mathbb{Z}\}, \{(\underline{i}, \underline{i + \alpha P}) : i \in \mathbb{Z}, \alpha \geq 0\} \rangle$ is closed under composition and it contains the frame $\mathcal{N}_p$ and the edges $\{(\underline{i}, \underline{i + j + 1}) : i \in \mathbb{Z}\}$. It implies that $\mathcal{N}_{p+1}$ is contained in $\mathcal{M}$ and if $b \in X_\Phi^{\mathcal{N}_{p+1}}$, then P divides b.

Hence, the threshold of $\mathcal{N}_{p+1}$ is bounded by T and the period is equal P.

Observe that $T = tr(\mathcal{N}_p) + j \cdot pi(\mathcal{N}_p) < (n^2 + 1) \cdot (tr(\mathcal{N}_p) + 1)$. By inductive assumption, it can be bounded by $(n^2 + 1) \cdot (n^{3(\log_2(n) - \log_2(pi(\mathcal{N}_p)) + 1)} + 1) < n^{3 + 3(\log_2(n) - \log_2(pi(\mathcal{N}_p)) + 1)}$. Since $gcd(pi(\mathcal{N})_p, j) \leq pi(\mathcal{N}_p)/2$, $\log_2(pi(\mathcal{N}_p)) - 1 \leq \log(pi(\mathcal{N}_{p+1}))$ and - $\log_2(pi(\mathcal{N}_p)) \geq -(\log(pi(\mathcal{N}_{p+1})) + 1)$, we can conclude that $T < n^{3 + 3(\log_2(n) - \log_2(pi(\mathcal{N}_{p+1})) - 3 + 1)} = n^{3(\log_2(n) - \log_2(pi(\mathcal{N}_{p+1})) + 1)}$.

If $R_p^s < 0$, then $pi(\mathcal{N}_{p+1})$ satisfies the condition of Lemma 6.19, i.e. there exists d such that $\underline{i} R_{p+1} \underline{j}$ iff $d | j - 1 - i$. Of course, d is the period of $\mathcal{N}_{p+1}$ and its threshold is equal 0. Moreover, since d divides all $b \in X_\Phi^{\mathcal{N}_p}$ and $X_\Phi^{\mathcal{N}_p}$ contains a number from the interval $[2, n]$ we have $d \leq n$.

Finally, at every step p, either the period decreases or a new edge $(\underline{0}, \underline{R_p^s})$ is added, where $|R_p^s| < n \cdot tr(\mathcal{N}_p) + pi(\mathcal{N}_p)$ and the period does not change. Hence, if between step p_1, p_2 the period does not decrease and the frames $\mathcal{N}_{p_1}, \mathcal{N}_{p_2}$ are different, then they have the same threshold and $|p_1 - p_2| \leq 2 \cdot (n \cdot tr(\mathcal{N}_{p_1}) + pi(\mathcal{N}_{p_1}))$. Hence, after at most $n \cdot 2 \cdot (n^{3(\log_2(n) - \log_2(pi(\mathcal{N}_p)) + 1)} + n)$ steps the sequence $\mathcal{N}_0, \mathcal{N}_1, \ldots$ stabilizes. $\square$

Now we are ready to prove Lemma 6.9. If $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ contains no long edges, then Φ satisfies S1 and we are done.

If $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ contains only forward edges, the let $\{a_1, \ldots, a_l\}$ be a minimal set such that X is the additive closure of $X_\Phi^{\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})}$. It is not hard to see that l and all $a_1, \ldots, a_l$ are smaller than the threshold of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$, and by Lemma 6.21 the threshold of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ is bounded by some $n^{O(\log_2(n))}$.

If $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ contains long and backward edges, then Lemma 6.19 implies that there is some m such that for all worlds $\underline{i}$, $\underline{i + b}$, there is an edge from $\underline{i}$ to $\underline{i + b}$ if and only if m divides $|b - 1|$. By Lemma 6.21 the value of m can be bounded by n.

6.6.3. Proof of Lemma 6.10

Let $\Phi \in \mathsf{UHF}$ and n be the number of variables in Φ. We start from two auxiliary lemmas.

Lemma 6.22. *Let $s \in \mathbb{N}_\infty$.*

(i) Let k be the maximal number such that v is k-proceeded (k-followed) in $\mathcal{I}_s$. If v is $k+1$-proceeded (resp. $k+1$-followed) in $\mathfrak{C}_\Phi(\mathcal{I}_s)$, then it is ∞-proceeded (resp. ∞-followed) in $\mathfrak{C}_\Phi(\mathcal{I}_s)$.

(ii) If $\underline{i}, \underline{i+j}$ are ∞-inner in $\mathfrak{C}_\Phi(\mathcal{I}_s)$, then there is a morphism f from $\mathcal{L}_\mathbb{Z}$ into $\mathfrak{C}_\Phi(\mathcal{I}_s)$ such that $f(\underline{i}) = \underline{i}$ and $f(\underline{i+j}) = \underline{i+j}$.

Proof. For part (i), let k be a maximal number such that v is k-inner in $\mathcal{I}_s$. Assume that k is also the maximal number such that v is k-proceeded in $\mathcal{I}_s$. Since v is $k+1$-proceeded in $\mathfrak{C}_\Phi(\mathcal{I}_s)$, there exists a path $v_1, v_2, \ldots, v_{k+1}, k$. If all worlds among $v_1, v_2, \ldots, v_{k+1}, v$ are different, then at least one of them is a descendant of v and therefore v belongs to a cycle in $\mathfrak{C}_\Phi(\mathcal{I}_s)$ and is ∞-inner. If $v_i = v_j$ for some $i < j$, then there is an infinite path $\ldots v_i, v_{i+1}, \ldots, v_j = v_i, v_{i+1}, \ldots v_j, v_{j+1}, \ldots, v_k + 1, k$ that proves that v is ∞-inner. The proof for the k-followed case is symmetric.

For part (ii), let $\ldots v_{-2}, v_{-1}, \underline{i}$ be a path that shows that $\underline{i}$ is ∞-proceeded and $\underline{i+j}, v_1, v_2, \ldots$ be a path that proves that $\underline{i+j}$ is ∞-followed. We define $f(\underline{l})$ equals v_{l-i} for $l < i$, $\underline{l}$ for $i \leq l \leq j$, and v_{l-j} otherwise. It is readily checkable that f is as required. □

Lemma 6.23. *There exists $k = O(\mathfrak{g}(|\Phi|))$ such that for every $s > 2 \cdot k$ if u, v are k-inner in $\mathfrak{C}_\Phi(\mathcal{I}_s)$, then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models uRv$ iff $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models uRv$.*

Proof. Clearly, for every s and all $u, v \in \mathcal{I}_s$, $\mathfrak{C}_\Phi(\mathcal{I}_s) \models uRv$ implies $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models uRv$. Let T and P denote the upper bounds on the threshold and the period of the structures among $\mathcal{N}_0, \mathcal{N}_1, \ldots$ with finite threshold, $b(p) = pn(T+n)$ and $\mathcal{N}_0, \mathcal{N}_1, \ldots$ be the sequence of approximations of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$. We show by induction w.r.t. p that if $\underline{i}, \underline{i+j+1}$ are $b(p)$-inner in $\mathcal{I}_s$ and $\mathcal{N}_p \models \underline{i}R\underline{i+j+1}$, then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+j+1}$.

The induction base. The frame $\mathcal{N}_0 = \mathcal{L}_\mathbb{Z}$ contains only short, forward edges that are also in $\mathcal{I}_s$ and therefore in $\mathfrak{C}_\Phi(\mathcal{I}_s)$.

The induction step. If $\mathcal{N}_{p+1} = \mathcal{N}_p$ then we are done. Otherwise, assume that for every all $s > 0, i \in \mathbb{Z}$ if $\underline{i}, \underline{i+j+1}$ are $b(p)$-inner in $\mathfrak{C}_\Phi(\mathcal{I}_s)$ and $\mathcal{N}_p \models \underline{i}R\underline{i+j+1}$, then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+j+1}$.

We first discuss the case when $\mathcal{N}_p$ contains only forward edges.

First, we show that for all $s > 0, i \in \mathbb{Z}$ if $\underline{i}, \underline{i+R^s_p}$ are $b(p+1)$-inner in $\mathfrak{C}_\Phi(\mathcal{I}_s)$, then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+R^s_p}$. Let $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_n}$, with $u_1 = i$ and $u_2 = i + R^s_p$, be the worlds that imply the edge $(\underline{i}, \underline{i+R^s_p})$ in $\mathcal{N}_{p+1}$. We consider two cases.

If $\mathcal{N}_p = \mathcal{L}_\mathbb{Z}$, then we may assume that $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_n}$ are contained in $\{\underline{i-n}, \underline{i-n+1}, \ldots, \underline{i+n}\}$, so if $\underline{i}$ is n-inner in $\mathcal{I}_s$, then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+R^s_p}$.

Otherwise, the frame $\mathcal{N}_p$ contains a long edge and, therefore, its threshold is finite. Clearly, $|R^s_p|$ is smaller than $tr(\mathcal{N}_{p+1}) + pi(\mathcal{N}_{p+1})$, in particular it is smaller than $T+n$. Let $v_1, v_2, \ldots$ be a result of application of Lemma 6.20 to $u_1, u_2, \ldots, a = 1$ and $b = 2$. If $\underline{v_1}, \underline{v_2}$ are $b(p+1)$-inner, then all $\underline{v_1}, \ldots, \underline{v_k}$ are $b(p)$-inner. By the inductive hypothesis for all $v, v' \in \{v_1, \ldots, v_n\}$, if $\mathcal{N}_p \models \underline{v}Rw_{v'}$, then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{v}Rw_{v'}$. Hence, $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+R^s_p}$.

If $R^s_p \leq 0$, then the above fact implies that for any $s > 0$ if some worlds $\underline{i}, \underline{i+R^s_p}$ are $b(p+1)$-inner in $\mathcal{I}_s$, then $\underline{i}, \underline{i+R^s_p}, \underline{i+R^s_p+1}, \ldots, \underline{i}$ is a cycle, and therefore $\underline{i}$ and $\underline{i+R^s_p}$ are ∞-inner in $\mathfrak{C}_\Phi(\mathcal{I}_s)$. If some world $\underline{i}$ is not $b(p+1)$-proceeded in $\mathcal{I}_s$ but it is $b(p+1)$-proceeded in $\mathfrak{C}_\Phi(\mathcal{I}_s)$, then by Lemma 6.22 (i) it is ∞-proceeded and, since it has a ∞-inner descendant, it is ∞-proceeded. Therefore all $b(p+1)$-inner worlds in $\mathfrak{C}_\Phi(\mathcal{I}_s)$ are ∞-inner. For any ∞-inner worlds $\underline{i}, \underline{i+j}$ Lemma 6.22 (ii) shows that if $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{i}R\underline{i+j}$ then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+j}$. This proves the lemma and therefore we can stop the induction here.

It remains to consider the case when $R^s_p > 1$. Let $r = R^s_p - 1$. We observe that $X_\Phi^{\mathcal{N}_{p+1}}$ is the additive closure of $X_\Phi^{\mathcal{N}_p} \cup \{r\}$. Indeed, the set $X_\Phi^{\mathcal{N}_{p+1}}$ contains $X_\Phi^{\mathcal{N}_p} \cup \{r\}$ and Lemma 6.18 implies that $X_\Phi^{\mathcal{N}_{p+1}}$ is additively closed. On the other hand, for every additively closed set Y the frame $\langle \{\underline{i} : i \in Z\}, \{(\underline{i}, \underline{i+j+1}) : i \in \mathbb{Z}, j \in Y\}\rangle$ is closed under composition. Hence, $X_\Phi^{\mathcal{N}_{p+1}}$ is equal to the additive closure of $X_\Phi^{\mathcal{N}_p} \cup \{r\}$. Since $X_\Phi^{\mathcal{N}_p}$ is additively closed, the set $X_\Phi^{\mathcal{N}_{p+1}}$ is equal to $\{j + \alpha r : \alpha \geq 0, j \in X_\Phi^{\mathcal{N}_p}\}$.

Thus, we have to show that for every $s, \alpha, i \in \mathbb{N}$, if $\underline{i}, \underline{i+j+\alpha r+1}$ are $b(p+1)$-inner in $\mathcal{I}_s$ and $j \in X_\Phi^{\mathcal{N}_{p+1}}$, then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+j+\alpha r+1}$. We will show that by induction w.r.t. α.

The base case, $\alpha = 0$, follows from the inductive hypothesis for p, that is if $\underline{i}, \underline{i+j+1}$ are $b(p)$-inner in $\mathfrak{C}_\Phi(\mathcal{I}_s)$ and $\mathcal{N}_p \models \underline{i}R\underline{i+j+1}$ (which means that $j \in X_\Phi^{\mathcal{N}_p}$), then $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+j+1}$.

The induction step. Let $t = s - (j + r(\alpha - 1))$. Assume that $\underline{i}, \underline{i+j+\alpha r+1}$ are $b(p+1)$-inner in $\mathcal{I}_s$. Then $\underline{i+r}$ is $b(p+1)$-inner in $\mathcal{I}_t$ and $\mathcal{I}_s$. Hence, $\mathfrak{C}_\Phi(\mathcal{I}_t) \models \underline{i}R\underline{i+r-1}$ and by the induction assumption $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i+r}R\underline{i+j+\alpha r+1}$.

Let us consider a function f defined as follows:

$$f(\underline{k}) = \begin{cases} \underline{k} & \text{if } k \leq i + r \\ \underline{k + j + (\alpha - 1) \cdot r} & \text{if } k > i + r \end{cases}$$

We have that $f(\underline{i+r}) = \underline{i+r}$ and $f(\underline{i+r+1}) = \underline{i+j+\alpha r+1}$, therefore $\mathfrak{C}_\Phi(\mathcal{I}_s)$ contains edge $(f(\underline{i+r}), f(\underline{i+r+1}))$. For $\underline{k}, \underline{k+1} \neq \underline{i+r}$ such that $\underline{k}, \underline{k+1} \in \mathcal{I}_t$ we have $f(\underline{k}), f(\underline{k}+1) \in \mathcal{I}_s$ and $\mathcal{I}_s \models f(\underline{k})Rf(\underline{k}+1)$. Hence, f is a morphism from $\mathcal{I}_t$ into $\mathfrak{C}_\Phi(\mathcal{I}_s)$. By Observation 6.4 f is a morphism from $\mathfrak{C}_\Phi(\mathcal{I}_t)$ into $\mathfrak{C}_\Phi(\mathcal{I}_s)$ and since $\mathfrak{C}_\Phi(\mathcal{I}_t) \models \underline{i}R\underline{i+r+1}$, we have $\mathfrak{C}_\Phi(\mathcal{I}_s) \models f(\underline{i})Rf(\underline{i+r+1})$. That is $\mathfrak{C}_\Phi(\mathcal{I}_s) \models \underline{i}R\underline{i+j+\alpha r+1}$.

□

For the proof of Lemma 6.10 is a consequence of Lemma 6.23. The proof of "⇒" is a simple application of Observation 6.4 to the morphism $h_\mathcal{T}$. For the "⇐", let $v_0, v_1, \ldots$ be a path containing v_i and v_j such that v_i, v_j are $\mathfrak{g}(|\Phi|)$-inner in this path and let $s \in \mathbb{N}_\infty$ be the length of this path. Then this path is isomorphic with $\mathcal{I}_s$. Due to Lemma 6.23 there is an edge from $h_\mathcal{T}(v_i)$ to $h_\mathcal{T}(v_j)$ in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ if and only if there is such an edge in $\mathfrak{C}_\Phi(\mathcal{I}_s)$, and therefore there is such edge in $\mathfrak{C}_\Phi(\mathcal{T})$.

6.6.4. Proof of Lemma 6.11

Let $\Phi \in \mathsf{UHF}$. For a tree $\langle W, R\rangle$, we define a *partial closures* $\mathcal{M}_0$, $\mathcal{M}_1$, …such that $\mathcal{M}_0 = \mathcal{T}$ and $\mathcal{M}_{i+1} = \langle W, R_{i+1}\rangle = \langle W, \textsc{Cons}_{\Phi,W}(R_i)\rangle$. Evidently, $\mathfrak{C}_\Phi(\mathcal{T})$ is a natural limit of the sequence of partial closures. We say that a partial closure $\mathcal{M}_i$

is tree-compatible if for all worlds w, v, if there is an edge between w and v in $\mathcal{M}_i$, then there is a path between w and v in $\mathcal{T}$. Clearly, if Φ has the tree-compatible model property, then all partial closures of $\mathcal{T}_\infty$ are tree-compatible.

Recall that $|s|$ is the length of s. For a given word s, by $s_{|k}$ we denote the prefix of $s0^\omega$ with length k. In other words, $s_{|k}$ contains first k letters of s, and if s is shorter than k, then last $k-|s|$ letters are 0.

Proof. Let k be a number from Lemma 6.23 and k' be a number from Lemma 6.21. We show that if a formula $\Phi \in \mathsf{UHF}$ does not fork on some level, then it does not fork at a level $k+nk'+n$, where n is a maximal number of variables in clauses from Φ.

Let $\mathcal{M}_0, \mathcal{M}_1, \ldots$ be partial closures of $\mathcal{T}_\infty$ and $p+1$ be the first index, such that $\mathcal{M}_p$ is tree-compatible and $\mathcal{M}_{p+1}$ is not. In this case, there is a clause Ψ of Φ such that for some worlds $\mathcal{M}_k \models \Psi(u_1, \ldots, u_l)$ for $u_1 = \underline{sit}$, $u_2 = \underline{si't'}$ for some $s, t, t' \in \{0,1\}^*$ and $i \neq i'$. Since $\mathcal{T}_\infty$ is isomorphic with all of its subtrees, we may assume the all of those worlds are at levels greater that k.

If $|s| < k + nk' + n$, then we are done. Otherwise, for each $j \in \{1, \ldots, l\}$, we define u'_j in the following way. If u_j is at the level $m \leq |s|$ in $\mathcal{T}_\infty$, we put $u'_j = \underline{s_{|m}}$. If $u_j = \underline{s''i''t''}$, where $|s''| = |s|$, $i \in \{0,1\}$ and $t'' \in \{0,1\}^*$, we set $u'_j = \underline{si''\bar{t}}$ where $\bar{t} = t_{||t''|}$ if $i = i''$ and $\bar{t} = t'_{||t''|}$ otherwise. A quick check shows that $\mathfrak{C}_\Phi(\mathcal{T}_\infty) \models \Psi(u'_1, \ldots, u'_l)$.

Let $v_s, v_1, \ldots, v_l$ be a result of application of Lemma 6.20 to $\pi_{\mathcal{T}_\infty}(\underline{s})$, $\pi_{\mathcal{T}_\infty}(u'_1)$, $\ldots$, $\pi_{\mathcal{T}_\infty}(u'_l)$. Thanks to automorphisms of $\mathcal{L}_\mathbb{Z}$ we may assume that the world with the lowest number among $v_s, v_1, \ldots, v_l$ is $\underline{k}$. Lemma 6.21 guarantees then that the highest number is bounded by $k + lk' + n$. It is not hard to see that if there is an edge between u'_j and $u'_{j'}$ in $\mathcal{M}_p$, then there is also such an edge in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ between $\pi_{\mathcal{T}_\infty}(u'_j)$ and $\pi_{\mathcal{T}_\infty}(u'_{j'})$, and thus there is an edge between v_j, $v_{j'}$.

Write $v_s = \underline{c}$. Let $s' = s_{|c}$. Now, for each v_j of the form $\underline{d}$, we define a world v'_j of $\mathcal{T}_\infty$ in the following way. If u'_j is of the form $\underline{s\bar{i}\bar{t}}$ for some $\bar{i}, \bar{t}$, then $v'_j = \underline{s'\bar{i}\bar{t}'}$, where $\bar{t}' = \bar{t}_{|c-|s'|-1}$ (note that Lemma 6.20 preserves the order and therefore $c-|s'|-1 \geq 0$). Otherwise, we put $v'_j = \underline{s'_{|d}}$. Note that each v'_j is at the level v_j.

Lemma 6.23 guarantees that for each v'_j and $v'_{j'}$, if these worlds are at the same path, then there are connected if and only if v_j, $v_{j'}$ are. It is not hard to see that v'_j and $v'_{j'}$ are on the same path if and only if u'_j and $u'_{j'}$ are. It means that $\Psi(v'_1, \ldots, v'_l)$ holds and therefore $\underline{s'}$ is not forking and $|s'| < k + k'n + n$. $\square$

6.6.5. Proof of Lemma 6.12

For $k_1 \in \mathbb{N}$ and $k_2 \in \mathbb{N}_\infty$, we define the frame $\mathcal{Y}(k_1, k_2)$ as $\mathcal{T}_{\infty\restriction W_Y}$, where $W_Y = \{\underline{s} : s \sqsubseteq 0^{k_1+k_2} \vee s \sqsubseteq 0^{k_1}1^{k_2}\}$ ($\sqsubseteq$ denotes the prefix relation).

Let Φ be a UHF formula that does not fork, n be the number of variables in Φ and $m_i = (n+i)\mathfrak{g}(|\Phi|)$. We start with auxiliary lemmas.

Lemma 6.24. *Let $m = m_1 - 1$. There exist $x, y \in [1, m_0]$ such that $|x - y| < m_0$ and $\mathfrak{C}_\Phi(\mathcal{Y}(m, m_2)) \models \underline{0^x} R \underline{0^m 1^{y-m}}$.*

Proof. Let $\mathcal{M}_0 = \mathcal{T}_\infty$ and for every $i \geq 0$ let $\mathcal{M}_{i+1} = \langle W, R_{i+1}\rangle = \langle W, \textsc{Cons}_{\Phi,W}(R_i)\rangle$. The frame $\mathfrak{C}_\Phi(\mathcal{T}_\infty)$ is the limit of $\mathcal{M}_0, \mathcal{M}_1, \ldots$. Let i, p, p_0, p_1 be such that $\mathcal{M}_{i+1} \models \underline{p0p_0} R \underline{p1p_1}$, but for all v, v', $\mathcal{M}_i \models vRv'$ implies that v, v' are along the same path in $\mathcal{T}_\infty$. In other words the edge $(\underline{p0p_0}, \underline{p1p_1})$ violates tree-compatibility.

Let $v_1 = \underline{p0p_0}$, $v_2 = \underline{p1p_1}$ and $v_3, \ldots, v_n$ be worlds such that Φ applied to $v_1, v_2, \ldots, v_n$ in $\mathcal{M}_i$ implies the edge $(\underline{p0p_0}, \underline{p1p_1})$ in $\mathcal{M}_{i+1}$. Since the function $f : \mathcal{T}_\infty \to \mathcal{T}_\infty$ defined as $f(\underline{u}) = \underline{0^{m_2} u}$ is a morphism of $\mathcal{M}_i$ into $\mathcal{M}_i$, we may assume that $|p| \geq m_2$ and all worlds $v_1, \ldots, v_n$ are below the level m_2 in $\mathcal{T}_\infty$.

Let $\mathcal{N}$ be the result of removing, from $\mathfrak{C}_\Phi(\mathcal{Y}(|p|, \infty))$, all edges $(\underline{p}, \underline{p'})$ such that $\underline{p}$ and $\underline{p'}$ are on different paths in $\mathcal{Y}(|p|, \infty)$, i.e. $p \not\sqsubseteq p' \wedge p' \not\sqsubseteq p$.

Let g be a function from $\mathcal{M}_i$ (containing $\mathcal{T}_\infty$) into $\mathcal{N}$ (containing $\mathcal{Y}(|p|, \infty)$) defined as follows:

$$g(\underline{s}) = \begin{cases} \underline{0^{|s|}} & \text{if } |s| \leq |p_1| \\ \underline{0^{|p_1|} z^{|s| - |p_1|}} & \text{if } |s| > |p_1| \text{ and } z \text{ is the } (|p_1| + 1)\text{th letter in } s \end{cases}$$

The function g is a morphism from $\mathcal{M}_i$ into $\mathcal{N}$ such that $g(\underline{p}) = \underline{0^{|p|}}$, $g(\underline{p0p_0}) = \underline{0^{|p|} 0^{|p_0|+1}}$ and $g(\underline{p1p_1}) = \underline{0^{|p|} 1^{|p_1|+1}}$.

Let $l_1, \ldots, l_n$ be levels of $g(v_1), \ldots, g(v_n)$. The numbers $l_1, \ldots, l_n$ not smaller than m_2. Let $l_1^a, \ldots, l_n^a$ be the result of applying Lemma 6.20 to $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and the sequence $l_1, \ldots, l_n$. For $i, j \in \{1, \ldots, n\}$ we have $|l_i^a - l_j^a| < m_0$ and the realtions in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ are the same among $\underline{l_1^a}, \ldots, \underline{l_n^a}$ as among $\underline{l_1}, \ldots, \underline{l_n}$.

Let m' be a maximal number such that $l_i^a > m'$ iff $l_i > |p|$ and $m = m_1 - 1$. For each i, we define $l_i' = l_i^a - m' + m$ Since the shift $sh_{(m-m')}$ is automorphism of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$, the realtions in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ are the same among $\underline{l_1'}, \ldots, \underline{l_n'}$ as among $\underline{l_1}, \ldots, \underline{l_n}$. Moreover, by the maximality of m, there is j such that $l_j' = m + 1$, and therefore for $i \in \{1, \ldots, n\}$ we have $l_i' \in [\mathfrak{g}(|\Phi|), m_2 - 1]$. Since $l_i \geq m_2$ we have $l_i' < l_i$.

We define worlds $v_1', \ldots, v_n'$ in $\mathcal{Y}(m, \infty)$ as follows: for $i \in \{1, \ldots, n\}$,

$$v_i' = \begin{cases} \underline{0^m 1^{l_i'' - m}} & \text{if } v_i = \underline{0^{|p|} 1^{l_i - |p|}} \\ \underline{0^{l_i''}} & \text{otherwise (when } v_i = \underline{0^{l_i}}) \end{cases}$$

By the definition of m, for $i, j \in \{1, \ldots, n\}$ worlds v_i, v_j are along the same path in $\mathcal{Y}(|p|, \infty)$ if and only if v_i', v_j' are along the same path in $\mathcal{Y}(m, \infty)$. In particular v_1', v_2' are on different paths.

Let $\mathcal{N}'$ be the result of removing from $\mathfrak{C}_\Phi(\mathcal{Y}(m, \infty))$ all edges $(\underline{p}, \underline{p'})$ such that $\underline{p}$ and $\underline{p'}$ are on different paths in $\mathcal{Y}(m, \infty)$. Since for all $i, j \in \{1, \ldots, n\}$ the worlds v_i, v_j are $\mathfrak{g}(|\Phi|)$-inner in $\mathcal{N}$, we have $\mathcal{N} \models v_i R v_j$ iff they are along the same path in $\mathcal{Y}(|p_0|, \infty)$ and $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{l_i} R \underline{l_j}$. The same equivaence holds for $\mathcal{N}'$. Then, for all $i, j \in \{1, \ldots, n\}$, we have $\mathcal{N} \models v_i R v_j$ iff $\mathcal{N}' \models v_i' R v_j'$. This implies that Φ applied to $v_1', \ldots, v_n'$ in $\mathcal{N}'$ implies the edge (v_1', v_2').

Finally, let $\mathcal{N}^f$ be the frame resulting from closure of $\mathcal{Y}(m, m_2)$ w.r.t. Φ along paths. The frame $\mathcal{N}^f$ is a finite subframe of $\mathcal{N}'$. The worlds $v'_1, \ldots, v'_n$ belong to $\mathcal{N}^f$ and they are $\mathfrak{g}(|\Phi|)$-inner in $\mathcal{N}^f$. Hence, Lemma 6.23 implies that for all $i, j \in \{1, \ldots, n\}$, we have $\mathcal{N}^f \models v'_i R v'_j$ iff $\mathcal{N}' \models v'_i R v'_j$. Thus, $\mathfrak{C}_\Phi(\mathcal{N}^f) \models v'_1 R v'_2$. Since $v'_1 = \underline{0^{l'_1}}$, $v'_2 = \underline{0^m 1^{l'_2 - m}}$ and $\mathfrak{C}_\Phi(\mathcal{N}^f) = \mathfrak{C}_\Phi(\mathcal{Y}(m, m_2))$, the result follows with $x = l'_1 - m < m_0$ and $y = l'_2 - m < m_0$. □

Lemma 6.25. *Let $m, l, x, y \in \mathbb{N}$ be such that $\mathfrak{C}_\Phi(\mathcal{Y}(m, l)) \models \underline{0^{m+x}} R \underline{0^m 1^y}$. Then, $\mathfrak{C}_\Phi(\mathcal{Y}(m+1, l+|x-y|)) \models \underline{0^{m+k}} R \underline{0^{m+1} 1^k}$ where $k = max(x, y)$.*

Proof. Since rotation of branches, i.e. the function $r : \mathfrak{C}_\Phi(\mathcal{Y}(m, l)) \to \mathfrak{C}_\Phi(\mathcal{Y}(m, l))$ defined as $f(\underline{0^z}) = \underline{0^z}$ for $z \leq m$, $f(\underline{0^{m+z}}) = \underline{0^m 1^z}$ for $z \leq l$ and $f(\underline{0^m 1^z}) = \underline{0^{m+z}}$ for $z \leq l$, is an automorphism of $\mathfrak{C}_\Phi(\mathcal{Y}(m, l))$, we can assume that $x \leq y$.

The morphism $f : \mathcal{Y}(m, l) \to \mathcal{Y}(m, l)$ defined as $f(\underline{s}) = \underline{0^{|s|}}$ implies that $\mathfrak{C}_\Phi(\mathcal{Y}(m, l))$ contains edge $(\underline{0^{m+x}}, \underline{0^{m+y}})$.

Finally, we define a function $g : \mathcal{Y}(m, l) \to \mathcal{Y}(m+1, l+|x-y|)$ as follows.

$$
g(\underline{p}) = \begin{cases} \underline{0^{|p|+1}} & \text{if } p \text{ ends with 0 and } |p| < m + x \\ \underline{0^{|p|+y-x}} & \text{if } p \text{ ends with 0 and } |p| \geq m + x \\ \underline{0p} & \text{if } p \text{ ends with 1} \end{cases}
$$

Since $\mathfrak{C}_\Phi(\mathcal{Y}(m+1, l+|x-y|)) \models \underline{0^{m+x}} R \underline{0^{m+y}}$, the function g is a morphism from $\mathcal{Y}(m, l)$ to $\mathfrak{C}_\Phi(\mathcal{Y}(m+1, l+|x-y|))$ and, by Observation 6.4, it is a morphism from $\mathfrak{C}_\Phi(\mathcal{Y}(m, l))$ to $\mathfrak{C}_\Phi(\mathcal{Y}(m+1, l+|x-y|))$. Hence, $\mathfrak{C}_\Phi(\mathcal{Y}(m+1, l+|x-y|)) \models g(\underline{0^{m+x}}) R g(\underline{0^m 1^y})$. Since $g(\underline{0^{m+x}}) = \underline{0^{m+x+(y-x)}}$, $g(\underline{0^m 1^y}) = \underline{0^{m+1} 1^y}$, the result follows.

□

Lemma 6.26. *There is $x \in [1, m_0]$ such that $\mathfrak{C}_\Phi(\mathcal{Y}(m_1, m_{n+2})) \models \underline{0^{m_1+x}} R \underline{0^{m_1} 1^{x+1}}$.*

This lemma is a straightforward consequence of Lemmas 6.24 and 6.25.

Lemma 6.27. *Let $\mathcal{T}$ be a tree of bounded degree and w be a world at level m_1 in $\mathcal{T}$. There exists $x \in [1, m_0]$ such that for every $i \geq 0$, if u_1, u_2 are m_{n+2}-followed descendants of w in $\mathfrak{C}_\Phi(\mathcal{T})$ at levels m_1+x+i and $m_1+x+i+1$, then $\mathfrak{C}_\Phi(\mathcal{T}) \models u_1 R u_2$.*

We remark that in contrary to Lemmas 6.24 and 6.25 in Lemma 6.27 the worlds need to be m_{n+2}-followed only in $\mathfrak{C}_\Phi(\mathcal{T})$. Some worlds that are not m_{n+2}-followed in $\mathcal{T}$ may become m_{n+2}-followed in $\mathfrak{C}_\Phi(\mathcal{T})$.

Proof. Let $m = m_1$, $x \in [1, m_0]$ be such that for $s_1 = 0^{m+x}$ and $s_2 = 0^m 1^{x+1}$ we have $\mathfrak{C}_\Phi(\mathcal{Y}(m, m_{n+2})) \models \underline{s_1} R \underline{s_2}$. Such x exists due to Lemma 6.26.

The proof is by induction w.r.t. i.

The base case, $i = 0$. Let u_1, u_2 be m_{n+2}-followed descendants of w in $\mathfrak{C}_\Phi(\mathcal{T})$ at levels $m + x$ and $m + x + 1$. There is a morphism f of $\mathcal{Y}(m, m_{n+2})$ into $\mathfrak{C}_\Phi(\mathcal{T})$

such that $f(\underline{0^m}) = w$, $f(\underline{s_1}) = u_1$ and $f(\underline{s_2}) = u_2$ for $s_1 = 0^{m+x}$ and $s_2 = 0^m1^{l+1}$. Therefore $\mathfrak{C}_\Phi(\mathcal{T}) \models u_1 R u_2$.

The inductive step. Let u_1, u_2 be m_{n+2}-followed descendants of w in $\mathfrak{C}_\Phi(\mathcal{T})$ at levels $m+x+(i+1)$ and $m+x+(i+1)+1$ and let v_1, v_2 be predecessors of u_1, u_2 in $\mathfrak{C}_\Phi(\mathcal{T})$ at levels $m+x+i, m+x+i+1$. Clearly, v_1, v_2 are m_{n+2}-followed descendants of w in $\mathfrak{C}_\Phi(\mathcal{T})$. By the induction assumption, $\mathfrak{C}_\Phi(\mathcal{T}) \models (v_1 R v_2)$.

Notice that there is a morphism g from $\mathcal{Y}(m, m_{n+2})$ into $\mathcal{Y}(m+x, m_{n+2}-x)$ such that $g(0^{m+x}) = 0^{m+x}$ and $g(0^m1^{x+1}) = 0^{m+x-1}11$. Since $\mathfrak{C}_\Phi(\mathcal{Y}(m, m_{n+2})) \models \underline{s_1} R \underline{s_2}$, the morphism g, extended to a morphism from $\mathcal{Y}(m, m_{n+2})$ to $\mathcal{Y}(m+x, m_{n+2}-x)$, implies that $\mathfrak{C}_\Phi(\mathcal{Y}(m+x, m_{n+2}-x)) \models \underline{0^{m+x-1}0} R \underline{0^{m+x-1}11}$.

There is a morphism h of $\mathcal{Y}(m+x, m_{n+2}-x)$ into $\mathfrak{C}_\Phi(\mathcal{T})$ such that $h(\underline{0^{m+x-1}}) = v_2$, $h(\underline{0^{m+x-1}0}) = u_1$ and $h(\underline{0^{m+x-1}11}) = u_2$. By Observation 6.4, h is a morphism from $\mathfrak{C}_\Phi(\mathcal{Y}(m+x, m_{n+2}-x))$ to $\mathfrak{C}_\Phi(\mathcal{T})$. Since $\mathfrak{C}_\Phi(\mathcal{Y}(m+x, m_{n+2}-x)) \models \underline{0^{m+x-1}0} R \underline{0^{m+x-1}11}$, we have $\mathfrak{C}_\Phi(\mathcal{T}) \models u_1 R u_2$.

□

Let Φ be a formula that does not fork at the level k, $\mathcal{T}$ be a tree of bounded degree, w be a world at level $m = m_1$ and $i > 0$. Now we prove that all $(m_{n+2}+1)$-followed descendants of w in $\mathfrak{C}_\Phi(\mathcal{T}^+)$ at level $m + m_0 + i$ are equivalent.

Let $i > 0$ and a, b be $(m_{n+2}+1)$-followed (in $\mathfrak{C}_\Phi(\mathcal{T}^+)$) descendants of w at level $L = m + m_0 + i$. It is sufficient to show that all predecessors of a are predecessors of b. Let c be a predecessor of a in $\mathfrak{C}_\Phi(\mathcal{T})$. Clearly, c is $(m_{n+2}+2)$-followed in $\mathfrak{C}_\Phi(\mathcal{T})$.

Let p_a (p_b) be the predecessor of a (b resp.). Let $S_{a,0}, S_{a,1}$ ($S_{b,0}, S_{b,1}$) be the partition of the successors of a (b resp.) such that worlds $S_{a,1}$ are (m_{n+2})-followed in $\mathfrak{C}_\Phi(\mathcal{T})$ and worlds $S_{a,1}$ are not. The sets $S_{a,1}, S_{b,1}$ are nonempty, since a, b are $(m_{n+2}+1)$-followed. We define $\mathcal{T}^+$ as the frame, based on $\mathcal{T}$, resulting from removing from $\mathcal{T}$ worlds $S_{a,0}, S_{b,0}$ and their descendants in $\mathcal{T}$, i.e. the successors of a or b which are not (m_{n+2})-followed in $\mathfrak{C}_\Phi(\mathcal{T})$ and their descendants (in $\mathcal{T}$). Additionally, $\mathcal{T}^+$ has the edges $(p_a, b), (p_b, a)$ and edges (x, y) for $x \in \{a, b\}$ and $y \in S_{a,1} \cup S_{b,1}$. Notice that a, b have the same predecessors and successors in $\mathcal{T}^+$. Since c is $(m_{n+2}+2)$-followed in $\mathfrak{C}_\Phi(\mathcal{T})$, it belongs to $\mathcal{T}^+$.

In the following we show three facts:

(i) $\mathfrak{C}_\Phi(\mathcal{T}^+) \models cRa$ $\quad (ii)$ $\mathfrak{C}_\Phi(\mathcal{T}^+) \models cRb$ $\quad (iii)$ $\mathfrak{C}_\Phi(\mathcal{T}) \models cRb$

In order to show (i), we consider a morphism f from $\mathcal{T}$ into itself, which is identity on $\mathcal{T}^+$ and maps worlds from $S_{a,0}$ and their descendants (in $\mathcal{T}$) into a path of maximal length beginning in a. Similarly, worlds from $S_{b,0}$ and their descendants are mapped into a path of maximal length beginning in b. We notice that the range of f is contained in $\mathcal{T}^+$, since any maximal path beginning in a has to contain a world from $S_{a,1}$. Otherwise, when a maximal path π beginning in a contains a world s_0 from $S_{a,0}$, there is a morphism which maps all the descendants of a on π, and this implies that s_0 is m_{n+2}-followed in $\mathfrak{C}_\Phi(\mathcal{T})$.

Hence, the morphism f is really a morphism from $\mathcal{T}$ into $\mathcal{T}^+$, $f(a) = a$ and $f(c) = c$. This implies that $\mathfrak{C}_\Phi(\mathcal{T}^+) \models cRa$.

For (ii), notice that since a and b have the same predecessors and the same successors in $\mathcal{T}^+$, the function g which is identity on $\mathcal{T}^+$ except that it swaps a with b is an automorphism of $\mathcal{T}^+$. Hence, by Observation 6.4 g is an automorphism of $\mathfrak{C}_\Phi(\mathcal{T}^+)$. This implies that if $a \neq c$, then $\mathfrak{C}_\Phi(\mathcal{T}^+) \models cRb$.

If $a = c$, then a is reflexive in $\mathfrak{C}_\Phi(\mathcal{T}^+)$ and g implies that b is also reflexive in $\mathfrak{C}_\Phi(\mathcal{T}^+)$. Then, the function h from $\mathcal{T}^+$ into $\mathcal{T}^+$ which maps all the descendants of b on b is a morphism from $\mathcal{T}^+$ into $\mathfrak{C}_\Phi(\mathcal{T}^+)$. The composition of h with f is a morphism from $\mathcal{T}$ into $\mathfrak{C}_\Phi(\mathcal{T}^+)$. The function $h \circ f$ is also a morphism from $\mathfrak{C}_\Phi(\mathcal{T})$ into $\mathfrak{C}_\Phi(\mathcal{T}^+)$ which maps a to a and the descendants of b into b. Let s be m_{n+2}-followed descendant of b. Due to Lemma 6.27, $\mathfrak{C}_\Phi(\mathcal{T}) \models aRs$. Hence, $\mathfrak{C}_\Phi(\mathcal{T}^+) \models h \circ f(a)Rh \circ f(s)$ where $h \circ f(a) = a$ and $h \circ f(s) = b$.

For (iii), Lemma 6.27 implies that the identity is a morphism from $\mathcal{T}^+$ into $\mathfrak{C}_\Phi(\mathcal{T})$. Hence, the identity a morphism from $\mathfrak{C}_\Phi(\mathcal{T}^+)$ into $\mathfrak{C}_\Phi(\mathcal{T})$ and $\mathfrak{C}_\Phi(\mathcal{T}^+) \models cRb$ implies $\mathfrak{C}_\Phi(\mathcal{T}) \models cRb$.

We showed that all predecessors of a are predecessors of b and, by symmetry, all predecessors of b are predecessor of a, and therefore a and b are equivalent.

7. The decidability

A well-known result shows that every satisfiable modal formula is satisfied in a finite tree. This *tree-model property* is crucial for the robust decidability of modal logics. Standard restrictions of classes of frames lead to similar results, stating that some "nice" models exists for all satisfiable formulas. Here we generalize those results for the classes of models that are definable by the Horn formulas.

Recall that the UHF formula Φ is not a part of an instance. To prove the decidability, it is enough to show that for every Φ there is an algorithm solving $\mathcal{K}_\Phi$-SAT. We are not going to present one *uniform algorithm* solving the satisfiability problem for all UHF formulas, because the complexity of such algorithm would be high.

If Φ is an inconsistent Horn formula, the satisfiability problem is in P (the answer is always "no"). For the case of consistent bounded formulas, we already proved that the satisfiability is NP-complete, (see Section 6.5).

Below we study consistent and unbounded formulas. We show different algorithms for formulas that fork at all levels and for formulas that do not fork at some level.

7.1. Tree-compatible case

In the following subsections we study formulas that fork at all levels. We show algorithms for formulas satisfying S1 and S2, and prove that it is impossible that a formula, which forks at all levels, satisfies S3.

7.1.1. Formulas that do not force long edges

Assume that all edges in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ are short. Here we can use standard approaches to satisfiability of modal logic over the class of all models. For local satisfiability we can bound the depth of tree-models and the degree of their worlds linearly in φ and then check the existence of such models in a depth-first search manner in PSpace (see e.g. [23]; please note that while the cited result does not consider reflexivity and symmetry, there are only some minor changes needed to cover these cases).

For the global satisfiability, we can enforce models of depth exponential with respect to the length of the modal formula φ. The existence of models can be checked by an alternating procedure, which first guesses the type of the root and then guesses types of its children and universally repeats the procedure for the children. This algorithm works in alternating polynomial space, and thus the problem is in ExpTime. The corresponding lower bound can be obtained by encoding the halting problem for alternating Turing machine with polynomial space.

7.1.2. Formulas that force only long forward edges

Assume that the condition S2 holds for some $l, a_1, \ldots, a_l$. This case can be treated similarly to the case of satisfiability over the class of transitive models, i.e. the case of logic K4 (see [23] or Section 6.7 in [4]). Let A be the additive closure of $\{a_1, \ldots, a_l\}$ and c be the product of all positive a_i ($c = \prod_{1 \leq i \leq l, a_i > 0} a_i$).

Let $P_{\mathcal{M}}(v)$ be a set of proper k-inner predecessors of v in M and $W_i = \{\underline{j} | j \leq i\}$. We have the following properties.

$$\text{For all } a \in A \text{ and } i, P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i}) \subseteq P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i+a}) \tag{7.1}$$

$$\text{For all } a \in A \text{ and } i, P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i+a}) \cap W_{i-c} \subseteq P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i}) \tag{7.2}$$

For the (7.1) note that for any $\underline{j} \in P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i})$ we have $i - j - 1 \in A$ and $a \in A$, and therefore $j + a - i - 1 \in A$ simply because A is closed under addition. Property (7.2) follows from property S2 and fact that for each $a \in A$ there exists $a' \in A$ such that $a = a' \mod c$ and $a' < c$ (which follows from Chinese remainder theorem).

$$\text{For } i \geq k, P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i}) = \bigcup_{a \in A, a < 2c} P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i-a}) \cup \{\underline{i-1}\} \tag{7.3}$$

The inclusion "$\subseteq$" comes from property (7.1). For the "$\supseteq$" case, consider any k-inner predecessor $\underline{j}$ of $\underline{i}$. If $i - j > 2c$, then Property (7.2) for $a = c$ guarantees that $\underline{j} \in P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i-j})$ only if $\underline{j} \in P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i})$. If $1 < i - j \leq 2c$, then $i - j - 1 \in A$. Since $\underline{j}$ is a predecessor of $\underline{j+1}$, $\underline{j} \in P_{\mathcal{L}_{\mathbb{Z}}}(\underline{j+1}) = P_{\mathcal{L}_{\mathbb{Z}}}(\underline{i - (i-j-1)})$ and $i - j - 1 < 2c$. Case when $i - j = 1$ is trivial.

For a given world w with a type t, we define *universal requirements* of w, denoted by $UR(w)$, as the subset of t that consists of formulas of the form $\Box\varphi$. Moreover, we define *predecessors requirements* of w, denoted by $PR(w)$, as the set of the universal requirements of the predecessors of w, i.e., $\bigcup\{UR(v) | v \text{ is a predecessor of } w\}$.

Clearly, property (7.3) implies that for all $i \geq k$

$$PR(\underline{i}) = \bigcup_{a \in A, a < 2c} PR(\underline{i-a}) \cup UR(\underline{i-1}) \cup PR_{ni}(\underline{i}) \tag{7.4}$$

where $PR_{ni}(wi)$ is a sum of requirements given by those predecessors of $\underline{i}$ that are not k-inner.

Now, we are ready to design an alternating algorithm that guesses a tree-based structure in top-down manner. For input φ, it starts from guessing and verifying first k levels. Then, the algorithm recursively calls procedure $verify(head, URs, PR, \diamond\psi)$ where

- $head$ contains information about the first k levels of structure;
- PRs is a list of predecessors requirements of previous $2c$ k-inner worlds;
- CR is a set of predecessors requirements for the current world;
- $\Diamond\psi$ is a subformula of φ.

The procedure guesses a type t that satisfies ψ and all requirements. Then it guesses a subset of subformulas of φ in order to provide all witnesses for the current world, and for each of them guesses whether they are k-inner. For each witness that is k-inner it simply guesses and verifies the remaining levels (at most $k-1$). For all others witnesses, it universally calls itself for this subformula with PRs and CR updated using Equation (7.4).

The algorithm described above verifies if φ has a model, but it may run forever. Therefore we add one more parameter to procedure $verify$: a list of visited configurations (i.e. triples $(PRs, CR, \Diamond\psi)$), and additional condition: return "Yes" if the same configuration is visited second time.

It is not hard to see that if this algorithm returns "Yes", then it is possible to build a model. Also, thanks to the property (i) of Lemma 6.2, if φ has a model, then it has a tree-based model such that all witnesses for the world at the level k are realized at the level $k+1$. In such tree-based model, worlds are connected only if they are on the same path in tree and, moreover, k-inner worlds v, w are connected if and only if $h_{\mathcal{T}}(v)$ and $h_{\mathcal{T}}(w)$ are. Such a canonical model can be guessed and verified by the algorithm. What remain to be explained is that this algorithm works in polynomial time.

The key observation here is that predecessors requirements cannot shrink, i.e., if we have two configurations $(PRs_1, CR_1, \diamond\psi_1)$ and $(PRs_2, CR_2, \diamond\psi_2)$ such that the algorithm visits the second one after the first one, then for each $r \in PRs_1 \cup \{CR_1\}$ (we abuse a notation here since no confusion will result) there is $r' \in PRs_2 \cup \{CR_2\}$ such that $r \subseteq r'$. It means that the number of possible PRs lists can be bounded by $|\varphi|^{2c} \cdot (2c)!$, and the number of all configurations can be bounded by $|\varphi|^{2c+1} \cdot (2c)! \cdot |\varphi|$, which is clearly polynomial in $|\varphi|$. Therefore, after a polynomial number of steps some configuration must occur twice. Since APTime = PSpace, it leads to the membership is PSpace in both global and local case.

7.1.3. Formulas that force long and backward edges

We prove that this case is not possible — S3 is inconsistent with the tree-compatible model property.

Let Φ satisfy S3 for some $m > 0$. Let $k = \mathfrak{g}(|\Phi|)$ and $w = 0^k$. By Lemma 6.10 we see that there are edges from $0^{k+(i+1)(m-1)}$ to $0^{k+i(m-1)}$ in $\mathfrak{C}(\mathcal{T}_\infty)$ for any $i \geq 0$. Define $h : \mathcal{L}_\mathbb{Z} \to \mathfrak{C}(\mathcal{T}_\infty)$ as $h(\underline{x}) = 0^{k-x(m-1)}$ for $x < 0$ and $h(\underline{x}) = 0^k 1^x$ otherwise. Clearly h is a morphism, and by Observation 6.4 it is also a morphism from $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ to $\mathfrak{C}(\mathcal{T}_\infty)$. Since in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ there is an edge from $\underline{1}$ to $\underline{1-3m+1}$, there is also an edge from $0^k 1$ to $0^{k+1-3m+1}$ and therefore w is not forking.

7.2. The tree-incompatibility

Let Φ be a formula without the tree-compatible model property. Recall that two worlds w, w' of a frame $\mathcal{M}$ are *equivalent* if for each world u we have uRw iff uRw'. We are going to exploit the property guaranteed by Lemma 6.12. We start with the observation that says that if we have two equivalent worlds v, w with the same types, then we can remove one of them.

Observation 7.1. *Let $\mathfrak{M} = \langle W, R, \pi \rangle$ be a structure such that $\langle W, R \rangle \models \Phi$, and φ be a modal formula such that $\mathfrak{M} \models \varphi$. If for all subformulas ψ of φ satisfied by w there is a world w' of $\mathfrak{M}$ such that $w \neq w'$, w, w' are equivalent and w' satisfies ψ, then for $W' = W \setminus \{w\}$ we have $\langle W', R_{\restriction W'} \rangle \models \Phi$ and $\mathfrak{M}_{\restriction W'} \models \varphi$.*

The proof is straightforward — the types of remaining worlds do not change.

Let $\mathfrak{M}$ be a tree-based model based on the frame $\mathfrak{C}(\mathcal{T})$. We denote by *level i of* $\mathfrak{M}$ the set of worlds from $\mathfrak{M}$ such that the length of the path from root to w in $\mathcal{T}$ (notice that $\mathcal{T}$ is a tree) is equal i.

Observation 7.2. *Let φ be a formula and $\mathfrak{M}$ be a tree-based model of Φ and φ. Then there is a model of Φ and φ such that the size of each level of $\mathfrak{M}$ is bounded polynomially in $|\varphi|$.*

First, observe that the number of worlds at level $i \leq 2\mathfrak{g}(|\Phi|)$ can be bounded by $|\varphi|^i$ because Lemma 6.2 guarantees that the degree of the tree is bounded by $|\varphi|$. Thanks to Lemma 6.11, Φ does not fork at level $\mathfrak{g}(|\Phi|)$. It follows from Lemma 6.12 that for all worlds w at the level $\mathfrak{g}(|\Phi|)$ and all $i \geq 2\mathfrak{g}(|\Phi|)$, all descendants of w at the level i are equivalent. Therefore, using Lemma 6.16, we can remove all but $|\varphi|$ of them. Since the number of worlds at the level $\mathfrak{g}(|\Phi|)$ can be bounded by $|\varphi|^{\mathfrak{g}(|\Phi|)}$, the number of worlds at the level $i > 2\mathfrak{g}(|\Phi|)$ can be bounded by $|\varphi|^{\mathfrak{g}(|\Phi|)} \cdot |\varphi|$, so polynomially in $|\varphi|$.

Observation 7.2 says that we can reduce the number of worlds needed at each level by some polynomial of $|\varphi|$. The existence of such models can be verified by a non-deterministic machine working in polynomial space that first guesses first $2\mathfrak{g}(|\Phi|)$ levels, and then recursively guesses and verifies the consecutive levels, similarly to the tree-compatible case. Since the number of worlds needed at each level can be bounded polynomially in $|\varphi|$, such an algorithm would work in NPSpace=PSpace [39]. We can conclude that here the satisfiability problem is in PSpace. This ends the proof of Theorem 3.1. However, it does not lead to the optimal complexity.

7.3. Sharpening the complexity

In this section, we study the satisfiability problems more carefully to obtain the precise complexity. The complexity results are summarized in Table 7.1.

7.3.1. Formulas with TCMP

Proposition 7.3. *For a given* UHF *formula Φ, if Φ has the tree-compatible model property and satisfies S2, then global $\mathcal{K}_\Phi$-SAT is in* NP.

Properties of Φ	global-$\mathcal{K}_\Phi$-SAT	$\mathcal{K}_\Phi$-SAT
Φ is inconsistent	P	P
Φ is consistent and bounded	NP-c	NP-c
Φ is consistent, unbounded, ...		
... has the TCMP and satisfies S1	ExpTime-c	PSpace-c
... has the TCMP and satisfies S2	NP-c	PSpace-c
... has the TCMP and satisfies S3	impossible	
... and does not have the TCMP and satisfies S1	PSpace-c	NP-c
... and does not have the TCMP and satisfies S2	NP-c	NP-c
... and does not have the TCMP and satisfies S3	NP-c	NP-c

Table 7.1.: A summary of a complexity of a satisfiability problem for modal logic defined by Horn formulas.

Proof. Let Φ satisfy S2 for some $l, a_1, \ldots, a_l$ bounded by $\mathfrak{g}(|\Phi|)$ and let c be the product of all a_i and $\mathfrak{M}$ be a T-based model of φ from $\mathcal{K}_\Phi$. We prove that φ has a $\mathcal{K}_\Phi$-based model with the number of types bounded by $|\varphi| \cdot c$.

We say that a world w at the level i (of $\mathcal{T}$) is *saturated* if for all k and every successors w' of w at levels $i + kc$, $PR(w) = PR(w')$.

Observe that in $\mathfrak{M}$ there is a world w such that the subtree rooted in w contains only saturated worlds. Let $\mathfrak{M}'$ be this subtree. Of course, $\mathfrak{M}'$ is a $\mathcal{K}_\Phi$-model of φ. For each subformula $\Diamond\psi$ of φ and each $i < c$, if there is a world in $\mathfrak{M}'$ at level $jc + i$ for some j that satisfies ψ, then we take a 1-type of one such world and call it $t_{\psi,j}$. It is not hard to see that there exists a model $\mathfrak{M}''$ that contains only w and worlds of these types — we can construct such a model starting from w, and then recursively constructing new levels that contain all needed witnesses for the previous level.

The non-deterministic algorithm proceeds as follows. First, it guesses sets of requirements PR_0, PR_1, ..., PR_{c-1}, and a subset of types of the form $t_{\psi,j}$. If this types are consistent with requirements and for each $t_{\psi,i}$ we can find $t_{\psi_1,i+1 \mod c}, \ldots,$ $t_{\psi_s,i+1 \mod c}$ such that these types provide all needed witnesses for a world of type $t_{\psi,i}$, then it returns "Yes", otherwise it returns "No". Clearly, it works in polynomial time and solves global $\mathcal{K}_\Phi$-SAT. □

Proposition 7.4. *For a given* UHF *formula Φ, if Φ has the tree-compatible model property, then $\mathcal{K}_\Phi$-SAT is* PSpace-*hard.*

Proof. We encode the QBF problem, adjusting the usual technique (see e.g. [23]). Let $P = \vartheta_1 p_1 \vartheta_2 p_2 \ldots \vartheta_n p_n.\rho$ be an instance of QBF problem, where $\vartheta_i \in \{\forall, \exists\}$ and ρ is quantifier-free. We define a modal formula φ such that P is true is and only if φ has a $\mathcal{K}_\Phi$-based model.

We define an operator $\Box_i\psi = \psi \wedge \Box\psi \wedge \Box\Box\psi \wedge \cdots \wedge \Box^i\psi$. Formula φ contains the variables $l_0, l_1, \ldots, l_n$ and $p_1, \ldots, p_n$ and is a conjunction of the following formulas.

1. l_0

2. $\Box_n \bigvee_{0\leq i\leq n} l_i \wedge \Box_n \bigwedge_{j\neq i} \neg(l_i \wedge l_j)$
3. $\Box_n(l_i \Rightarrow \Diamond l_{i+1})$ for each $i < n$ such that $\vartheta_{i+1} = \exists$
4. $\Box_n(l_i \Rightarrow \Diamond(l_{i+1} \wedge p_{i+1}) \wedge \Diamond(l_{i+1} \wedge \neg p_{i+1}))$ for each $i < n$ such that $\vartheta_{i+1} = \forall$
5. $\Box_n((l_i \wedge p_i \Rightarrow \Box_{n-i} p_i) \wedge (l_i \wedge \neg p_i \Rightarrow \Box_{n-i} \neg p_i))$ for each $i < n$
6. $\Box_n(l_n \Rightarrow \psi)$

Consider a tree $\mathcal{T}$ that consists of $n+1$ levels, and each world at ith level has one successor if $\vartheta_i = \exists$ and two successors otherwise.

Assume that P is true. We define a labeling π of $\mathcal{T}$ inductively, starting from the root. Let root satisfy only l_0. Let w be at level i. Define $t_{i+1} = \pi(w) \setminus \{l_i\} \cup \{l_{i+1}\}$. If $\vartheta_i = \forall$, then w has two successors and we set their labellings to be t_{i+1} and $t_{i+1} \cup \{p_{i+1}\}$. Otherwise, if set the labeling of the successor of w to t_{i+1} if formula $\vartheta_{i+2} p_{i+2} \ldots \vartheta n p_n.\rho$ is satisfied for a valuation that makes true precisely the variables from t_{i+1} and $t_{i+1} \cup \{p_{i+1}\}$ otherwise. Then, $\langle \mathfrak{C}_\Phi(\mathcal{T}), \pi\rangle$ is a $\mathcal{K}_\Phi$-based model of φ.

On the other hand, if φ has a model, then we can show that $\mathcal{T}$ can be homomorphically embedded in this model and the image of this embedding is a justification that P is true. $\Box$

Proposition 7.5. *For a given* UHF *formula Φ, if Φ has the tree-compatible model property and satisfies S1, then global $\mathcal{K}_\Phi$-SAT is* PSPACE*-hard.*

The proof is almost the same as the previous one, except that we replace the conjunct 1 by $l_n \Rightarrow \Diamond l_0$.

7.3.2. Formulas without TCMP that do not force long edges

Proposition 7.6. *For a given* UHF *formula Φ, if Φ does not have the tree-compatible model property and satisfies S1, then it has a polynomial model property for the local satisfiability problem.*

Proposition 7.6 follows from the fact that in the local satisfiability case for any tree-based model $\mathfrak{M}$ based of $\mathfrak{C}(\mathcal{T})$ such that of $\mathfrak{M}_0, \underline{0} \models \varphi$, we can simply remove all worlds w that are at the levels greater than d, the quantifier depth of φ. Indeed, S1 says that there are only short edges in closures and therefore the removed worlds were not reachable by φ. The resulting model contains at most $|\varphi|^{2\mathfrak{g}(|\Phi|)+1}$ worlds at first $2\mathfrak{g}(|\Phi|)$ levels and then at most $|\varphi|$ worlds at each of remaining $d - 2\mathfrak{g}(|\Phi|)$ levels, so clearly a polynomial number of worlds.

Proposition 7.7. *For a given* UHF *formula Φ, if Φ does not have the tree-compatible model property and satisfies S1, then global $\mathcal{K}_\Phi$-SAT is* PSPACE*-hard.*

Proof. To make the proof more readable, we consider only the formula $\Phi = \{sRt \wedge tRy \wedge sRx \Rightarrow xRy\}$. Proofs for other cases are similar.

Let $\langle \mathcal{D}, V_\mathcal{D}, H_\mathcal{D}\rangle$ be an instance of the bounded-space domino problem and $n = O(|\mathcal{D}|)$. We define a formula $\varphi = \psi_c \wedge \psi_v \wedge \psi_h \wedge \psi_e$ over variables $\{t_0, \ldots, t_{n-1}\} \cup \mathcal{D}$ where:

- $\psi_c = \bigvee_{d \in \mathcal{D}} d \wedge \bigwedge_{d,d' \in \mathcal{D}, d \neq d'} (\neg d \vee \neg d')$;
- $\psi_e = \bigwedge_{i<n} \Diamond t_i$;
- $\psi_v = \bigwedge_{i<n} \bigwedge_{d \in \mathcal{D}} (t_i \wedge d \Rightarrow (\bigvee_{(d,d') \in V_{\mathcal{D}}} \Box(t_i \Rightarrow d')))$;
- $\psi_h = \bigwedge_{i<n-1} \bigwedge_{d \in \mathcal{D}} (\Box(t_i \wedge d) \Rightarrow (\bigvee_{(d,d') \in H_{\mathcal{D}}} \Box(t_{i+1} \Rightarrow d')))$.

Clearly, the reduction is polynomial. Suppose that $\mathfrak{M}$ is a model of Φ and φ and v_0 is any world of $\mathfrak{M}$. We define the tiling t by repeating the following procedure. For a given i, we define $v_{j,i}$ as a successor of v_i that satisfies t_j and we put $t(j,i) = d$, where d is satisfied in $v_{j,i}$. Note that ψ_e guarantees that such a successor exists, ψ_v guarantees that if there is more than one such successor, then all of them satisfy the same d, and ψ_c guarantees that all worlds satisfy precisely one d. Finally, we set v_{i+1} equal to any successor of v_i that satisfies t_0.

It is not hard to see that for all $k < n-1$ and $l \in \mathbb{N}$ property $(t(k,l), t(k+1,l)) \in H_{\mathcal{D}}$ is guaranteed by ψ_h since both $v_{k,l}$ and $v_{k+1,l}$ are successors of v_l. To check the other property, consider any $l \in \mathbb{N}$ and $k < n$. Since $v_l R v_{l+1}$, $v_{l+1} R v_{k,l+1}$, and $v_l R v_{k,l}$, Φ guarantees that we have $v_{k,l} R v_{k,l+1}$ and therefore ψ_v guarantees that $(t(k,l), t(k,l+1)) \in V_{\mathcal{D}}$.

We showed that if φ has a model that satisfies Φ, then the domino problem has a solution. It should be now easy to see that the converse is also true. $\Box$

7.3.3. Formulas without TCMP that force only long forward edges

Proposition 7.8. *For a given* UHF *formula Φ, if Φ does not have the tree-compatible model property and satisfies S2, then global and local $\mathcal{K}_\Phi$-SAT are* NP*-complete.*

Proof. Let Φ be a UHF formula that does not have the tree-compatible model property and satisfies S2 for some $l, a_1, \ldots, a_l$, φ be a modal formula and $\mathfrak{M}$ be a tree-based model of Φ and φ. Let $c = a_1 \cdot \cdots \cdot a_l$ and for a world w at level $\mathfrak{g}(|\Phi|)$ and $i > \mathfrak{g}(|\Phi|)$, set C_i^w be the set of all descendants of w at level i. According to previous observations, we may assume that the size of each such set is polynomial in $|\varphi|$. Our goal is to show that for any w, it is enough to consider only polynomially many non-isomorphic sets C_i^w. Clearly, it will make the algorithm described above run in the polynomial time.

In Section 7.1 we showed similar property, but the technique used there is not sufficient for this case — now, it is not enough just to satisfy one formula of the form $\Diamond\psi$ at each level. We solve this problem in the following way: in each C_i^w, we put as many witnesses as possible. We extend the notation from Section 7.1 defining $PR(X) = \bigcup_{w \in X} PR(w)$. Note that since all worlds in C_i^w are equivalent, for any $v \in C_i^w$ we have $PR(v) = PR(C_i^w)$. Moreover, Properties (7.1) and (7.2) also holds in this case.

Observation 7.9. *Let w, v be worlds such that $v \in C_j^w$ for some j and let i be such that $\mathfrak{g}(|\Phi|) < i < j$ and c divides $j-i$. If $UR(v) \subseteq PR(C_{i+1}^w)$ and $PR(v) = PR(C_i^w)$, then model obtained by adding a copy v' of v to C_i^w satisfies both Φ and φ.*

Note that the set of successors of v is a subset of the set of successors of v', and therefore v has all the needed witnesses. Moreover, the set of predecessors of v' is a subset of the set of predecessors of v, so v' does not violate any predecessor requirements. Finally, since v' does not add any new requirements, it should be clear that new model satisfies φ. Therefore the new model satisfies φ and, in an obvious way, Φ.

Observation 7.10. *Let w be a world at level $\mathfrak{g}(|\Phi|)$, let $i > \mathfrak{g}(|\Phi|)$, and let $A = \{0, 1, \dots\}$ be a (possibly finite) set of consecutive numbers. Let $\mathcal{C} = \bigcup\{C^w_{i+ac} | a \in A\}$ be such that for all $j, j' \in A$, $PR(C^w_j) = PR(C^w_{j'})$ and $PR(C^w_{j+1}) = PR(C^w_{j'+1})$. Then, we can define a set C' with $|C'| \leq |\varphi|$ such that each element of $\bigcup \mathcal{C}$ can be replaced by a copy of an element from C' in a way such that the obtained model is still a model of φ and Φ.*

Let $C = \bigcup \mathcal{C}$. We define a $C' \subseteq C$ in the following way. For every subformula of φ of the form $\Diamond\psi$, if there is a type t satisfying ψ such that t is realized in infinitely many elements of $\mathcal{C}$, then we take one world of this type and add it to C'. If there is no such type, but there is a world in C that satisfies ψ, then we find a maximal $j \in A$ such that there is such a world $v \in C^w_{i+jc}$ and we add v to C'. Clearly, $|C'| \leq |\varphi|$. Then, we define $C'\underline{^{i+jc}} = C' \cap \bigcup_{a \in A, a \geq j} C^w_{i+ac}$ and replace each C^w_{i+jc} by $C'\underline{^{i+jc}}$. Note that such a model can be obtained by applying Observation 7.9 first, and then Lemma 6.16, and therefore it satisfies both φ and ϕ.

Let w be a world at level $\mathfrak{g}(|\Phi|)$ and i be such that $\mathfrak{g}(|\Phi|) \leq i < \mathfrak{g}(|\Phi|) + c$. Property (7.1) still holds and shows that the sequence $PR(C^w_i), PR(C^w_{i+c}), PR(C^w_{i+2c}) \dots$ never shrinks, and the same holds for $PR(C^w_{i+1})$, $PR(C^w_{i+c+1})$, $PR(C^w_{i+2c+1})$, $\dots$. Therefore, the sequence $C^w_i, C^w_{i+c}, C^w_{i+2c}$ can be split into at most $|\varphi|^2$ subsequences that satisfy the requirements of Observation 7.10, so the number of different sets of the form C^w_i can be bounded by $|\varphi|^3$. Taking into account all possible w and i, we can bound the number of possible sets C^w_i by $|\varphi|^{\mathfrak{g}(|\Phi|)} \cdot c \cdot |\varphi|^3$, which is clearly polynomial in φ. $\square$

7.3.4. Formulas without TCMP that force long and backward edges

Proposition 7.11. *For a given* UHF *formula Φ, if Φ does not have the tree-compatible model property and satisfies S3, then it has the polynomial model property.*

Proof. Suppose that Φ does not have the tree-compatible model property and satisfies S3 for some k, m. Observe that in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ for all $i > k$ and $l \geq 0$, worlds $\underline{i}$ and $\underline{i + lm}$ are equivalent. Let $\mathfrak{M}$ be a model of φ. It follows from Lemmas 6.11 and 6.12 that for all w at the level $\mathfrak{g}(|\Phi|)$ and all i, all descendants of w at levels $2\mathfrak{g}(|\Phi|) + i$, $2\mathfrak{g}(|\Phi|) + i + m$, $2\mathfrak{g}(|\Phi|) + i + 2m$, $\dots$ are equivalent. Thanks to Lemma 6.16 we can remove all but polynomially many of them and obtain a smaller model that still satisfies φ. We may repeat this procedure for all such w, finally obtaining model of polynomial size in $|\varphi|$. $\square$

7.4. Horn formulas and equality

In this section, we prove Theorem 3.2. It is not hard to see that each negative occurrence of equality may be eliminated by simply identifying variables. Thus, in the rest of this section we focus on formulas without negative occurrences of equality. For a given $\Phi \in \mathsf{UHF}^=$, let $\Phi^\#$ contain all the clauses of Φ except for those with the positive occurrence of equality.

We say that an $\mathsf{UHF}^=$ formula Φ *essentially uses* equality if there is a modal formula φ such that φ has a $\Phi^\#$–based model but it does not have a Φ–based model. Clearly, if Φ does not essentially use equality, then $\mathcal{K}_\Phi$–SAT is equal to $\mathcal{K}_{\Phi^\#}$–SAT.

Proposition 7.12. *For any* $\mathsf{UHF}^=$ *formula* Φ *that essentially uses equality,* $\mathcal{K}_\Phi$*–SAT and global* $\mathcal{K}_\Phi$*–SAT are in* NP.

Proof (sketch). Let Φ be an $\mathsf{UHF}^=$ formula that essentially uses equality.

Clearly, there is a tree $\mathcal{T}$ such that $\mathfrak{C}_{\Phi^\#}(\mathcal{T})$ is a model of $\Phi^\#$ but not a model of Φ — otherwise, every tree-based model over $\mathcal{K}_\#$ would be a model over $\mathcal{K}_\Phi$, contradicting the fact Φ essentially uses equality. We consider two cases.

Suppose that there are two worlds w, v at different levels in $\mathcal{T}$ such that for some clause $\Psi \Rightarrow x = y$ of Φ in $\mathfrak{C}_{\Phi^\#}(\mathcal{T})$ the formula Ψ is satisfied for some instantiation that substitutes x by w and y by v. Consider the morphism $h_\mathcal{T}$ (recall that $h_\mathcal{T}(w) = \underline{i}$ if w is at the level i) and the worlds $h_\mathcal{T}(w)$ and $h_\mathcal{T}(v)$. Clearly, in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ we have $\Psi(h_\mathcal{T}(w), h_\mathcal{T}(v))$. By the definition of $h_\mathcal{T}$ we know that $h_\mathcal{T}(w) \neq h_\mathcal{T}(v)$, and therefore $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ is not a model of Φ. It implies that Φ is bounded and therefore it has the polynomial model property.

The other case is more interesting. In this case, there are worlds w, v at the same level in $\mathcal{T}$ such that some clause of Φ of the form $\Psi \Rightarrow x = y$ is not satisfied in $\mathfrak{C}_{\Phi^\#}(\mathcal{T})$ for some instantiation that substitutes x with w and y with v. An example of formula with this property if $vRx \wedge vRy \Rightarrow x = y$. It is not hard to see that there are such worlds also in $\mathfrak{C}_{\Phi^\#}(\mathcal{T}_\infty)$.

We say that (possibly infinite) directed acyclic graph (DAG) is proper if it has a root r such that all vertices of this DAG are reachable from r, and for all elements v, v' all paths from v to v' have the same length. Note that trees are special cases of proper DAGs.

We adjust Lemma 6.2: if φ is $\mathcal{K}_\Phi$-satisfiable, then there exists a proper directed acyclic graph $\mathcal{T}$ with the degree bounded by $|\varphi|$ and a labeling $\pi_\mathcal{T}$, such that $\langle \mathcal{T}, \pi_\mathcal{T} \rangle$ is a model of φ and $\langle \mathfrak{C}_\Phi(\mathcal{T}), \pi_\mathcal{T} \rangle$ is a model of φ that satisfies Φ.

As for trees, we define the level of v in DAG as the length of path from the root to v. Therefore, morphism $\pi_\mathcal{T}$ is well-defined also for DAGs. Then, we adjust Lemma 6.12: if, for any proper DAG $\mathcal{T}$, w is a world at level $\mathfrak{g}(|\Phi|)$ in $\mathfrak{C}_\Phi(\mathcal{T})$, then for all i, there is at most one $\mathfrak{g}(|\Phi|)$-followed descendant of w at level $2 \cdot \mathfrak{g}(|\Phi|) + i$ in the frame $\mathcal{T}$. It means that the number of worlds at each level can be bounded by $|\varphi|^{\mathfrak{g}(|\Phi|)}$. Then we show, as in the case of tree-incompatibility, that it is enough to consider polynomial number of different types of levels, and therefore that both the global and the local satisfiability problems are NP-complete. $\square$

8. Finite satisfiability

In this section, we show that for any $\Phi \in \mathsf{UHF}$, finite $\mathcal{K}_\Phi$–SAT and finite global $\mathcal{K}_\Phi$–SAT are decidable. In case of some UHF formulas, decidability of corresponding modal logics is shown in previous section by proving that every modal formula satisfiable over $\mathcal{K}_\Phi$ has also a finite model in $\mathcal{K}_\Phi$ — for such a formulas, the finite local (global) satisfiability problem is equal to the local (resp. global) satisfiability problem and, therefore, has the same complexity. However, some logics lack the finite model property (with respect to the given class of frames). Below we study the satisfiability problem for all possible UHF formulas.

8.1. Formulas that do not force long edges

In this section, we consider formulas $\Phi \in \mathsf{UHF}$ that satisfy S1. We have already seen that often the logics defined by UHF formulas have the finite model property, e.g., the logics defined by the bounded formulas. In such cases, the question about the existence of a finite model is equivalent to the question about the existence of any model, and therefore these problems have the same complexity. Below we study unbounded (and therefore consistent) formulas.

8.1.1. Local satisfiability

Proposition 8.1. *Each unbounded* UHF *formula* Φ *that does not force long edges has the finite model property in the local satisfiability case.*

Proof. Let $\mathfrak{M}$ be a $\mathcal{T}$-based model of a modal formula φ for some $\mathcal{T}$ such that $\mathfrak{C}_\Phi(\mathcal{T}) \in \mathcal{K}_\Phi$. Such a model exists due to Lemma 6.2. Since Φ does not force long edges, under the morphism $h_\mathcal{T}$ it follows that $\mathfrak{C}_\Phi(\mathcal{T})$ can only contain edges between states on identical or consecutive levels.

In order to obtain a finite model, we simply remove from $\mathfrak{M}$ all worlds from levels greater than $|\varphi|$. Since the truth of φ depends only on the worlds that are reachable from a root by a path with the length bounded by $|\varphi|$ (in fact, by a modal depth of φ), the resulting model is a finite model of φ and, of course, it satisfies Φ since Φ is universal. $\square$

We showed that φ has a $\mathcal{K}_\Phi$-based model if and only if it has a finite $\mathcal{K}_\Phi$-based model, so $\mathcal{K}_\Phi$-SAT is equal to finite $\mathcal{K}_\Phi$-SAT.

Properties of Φ	finite global-$\mathcal{K}_\Phi$-SAT	finite $\mathcal{K}_\Phi$-SAT
inconsistent	P (TRIVIAL)	
consistent and bounded	FMP, NP-c	
unbounded, satisfies S2	NEXPTIME	
unbounded, satisfies S3	FMP, NP-c	
is unbounded, does not force long edges and …		
… forks at all levels and merges at some level	Lack of FMP, PSPACE-c	FMP, PSPACE-c
… forks at all levels and does not merge at any level	FMP, EXPTIME-c	FMP, PSPACE-c
… does not fork at some level	FMP, PSPACE-c	FMP, NP-c

Table 8.1.: A summary of a results for the finite satisfiability problems for modal logic defined by Horn formulas.

8.1.2. Global satisfiability

The case of global satisfiability is much more complicated. In the case of general satisfiability, it was enough to consider the behavior of a first order formula on $\mathcal{T}_\infty$ and $\mathcal{L}_\mathbb{Z}$. In the case of finite satisfiability, we need one more structure, called $\mathcal{X}$, that contains a world with in-degree 2.

Formally, we define the frame $\mathcal{X}$ as $\langle W_X, R_X\rangle$, where $W_X = \{\underline{i} | i \in \mathbb{Z}\} \cup \{\bar{i} | i \in \mathbb{Z} \setminus \{0\}\}$ and $R_X = \{(\underline{i}, \underline{i+1}) | i \in \mathbb{Z}\} \cup \{(\bar{i}, \overline{i+1}) | i \in \mathbb{Z} \setminus \{-1, 0\}\} \cup \{(\overline{-1}, \underline{0}), (\underline{0}, \overline{1})\}$. Fig. 8.1 contains a fragment of the structure $\mathcal{X}$.

We say that a formula Φ *merges* at a level $k < 0$ if in $\mathfrak{C}_\Phi(\mathcal{X})$ there is an edge from $\underline{k-1}$ to $\bar{k}$.

Example. Consider the formula $\Phi = xRz \wedge zRv \wedge yRv \Rightarrow xRy$ over $\mathcal{X}$. Clearly, Φ implies an edge from $\underline{-2}$ to $\overline{-1}$, so Φ merges at level 1. However, frames $\mathcal{T}_\infty$ and $\mathcal{L}_\mathbb{Z}$ satisfy Φ.

We consider three cases. For each formula Φ of UHF such that Φ does not force long edges, merges at some level and forks at all levels, we show that the modal logic over $\mathcal{K}_\Phi$ does not have the finite model property (Proposition 8.3), and that the global $\mathcal{K}_\Phi$–SAT is PSPACE–complete (Propositions 8.7 and 8.8). In the remaining cases, i.e. the case of formulas that do not force long edges, do not merge at any level and fork at all levels and and the case of formulas that do not force long edges and do not fork at all levels, the decidability follows from the finite model property

(Propositions 8.9 and 8.10).

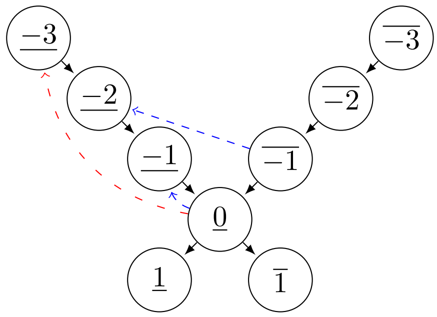

Figure 8.1.: A fragment of $\mathcal{X}$ structure (solid edges). Consider a formula $\Phi = yRw \wedge wRv \wedge xRv \Rightarrow xRy$ that forces an edge $(\overline{-1}, \underline{-2})$. When applied to $x = w = \underline{0}$, $y = \underline{-1}$ and $v = \underline{1}$, it implies edge $(\underline{0}, \underline{-1})$. Then, applied to $x = \underline{0}$, $v = \underline{-1}$, $w = \underline{-2}$ and $y = \underline{-3}$ it implies long edge $(\underline{0}, \underline{-3})$.

The following lemma shows an important regularity in models of formulas that merge.

Lemma 8.2. *Let Φ be an unbounded* UHF *formula that does not force long edges and merges at a level k, $\mathfrak{M}$ be a model of Φ, $v_1, v_2, \dots, v_l$ be a walk (i.e. a path, but not necessarily simple) in $\mathfrak{M}$ such that all v_i are ∞-inner.*

(i) If $v_l R v_{l-c}$ for some $c > 0$, then for all $i > c$, $v_i R v_{i-c}$.

(ii) If $v_{l-c} R v_l$ for some $c > 0$, then for all $i > c$, $v_{i-c} R v_i$.

Proof. Let $\dots, v_{-2}, v_{-1}, v_0, v_1$ and $v_l, v_{l+1}, \dots$ be infinite walks in $\mathfrak{M}$. Such walks exist since v_1 and v_l are ∞-inner.

We prove (i) in by induction. Assume that for some $i > 0$ for all $j > i$ we have $v_j R v_{j-c}$. We define a morphism h from $\mathcal{X}$ into $\mathfrak{M}$ as

$$h(w) = \begin{cases} v_{i+s+1} & \text{if } w = \underline{k+s} \text{ for some } s \leq 0 \\ v_{i-c+s} & \text{if } w = \underline{\underline{k+s}} \text{ for some } s > 0 \\ v_{i-c+s} & \text{if } w = \overline{k+s} \text{ for some } s \in \mathbb{Z} \end{cases}$$

A quick check shows that h is a morphism and since $\mathfrak{C}_\Phi(\mathcal{X})$ contains an edge from $\underline{k-1}$ to $\overline{k}$, $\mathfrak{M}$ has to contain edge from v_i to v_{i-c}.

The proof of (ii) is similar and thus omitted. □

Now we use the above lemma to show the lack of finite model property.

Proposition 8.3. *Each unbounded* UHF *formula Φ that does not force long edges, merges at a level $k < 0$ and forks at all levels lacks the finite model property in global case.*

Proof. Consider a formula τ defined as the conjunction of the following formulas.

1. $\bigvee_{i\in\{1,2,3,4\}} p_i \wedge \bigwedge_{i,j\in\{1,2,3,4\},i\neq j} \neg(p_i \wedge p_j)$ (each world satisfies exactly one of p_1, p_2, p_3, p_4).

2. $p_1 \Rightarrow (\Diamond p_2 \wedge \Box p_2)$

3. $p_2 \Rightarrow (\Diamond p_3 \wedge \Box p_3)$

4. $p_3 \Rightarrow (\Diamond p_2 \wedge \Diamond p_4 \wedge \Box(p_2 \vee p_4))$

5. $p_4 \Rightarrow (\Diamond p_1 \wedge \Box p_1)$

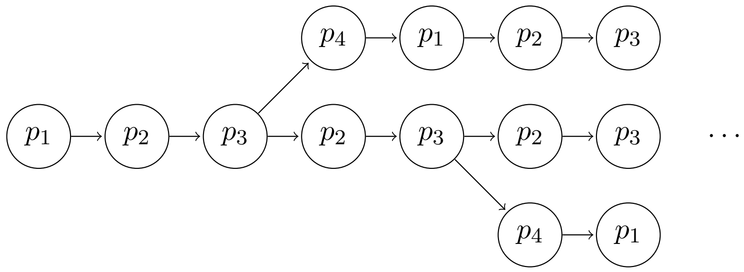

Figure 8.2.: An infinite model of τ.

An infinite model of this formula is presented in Fig. 8.2.

Assume that $\mathfrak{M}$ is a finite model of τ and let w be a world that satisfies p_1 in $\mathfrak{M}$. Quick check shows that such a world always exists. We define the path $w_0, w_1, \ldots, w_l$ such that all worlds of the form w_{2i+2} satisfy p_3 and all worlds of the form w_{2i+1} satisfy p_2. We do it recursively, starting from $w_0 = w$. Then, if i is odd, we find a successor v of w_{i-1} that satisfies p_2. Such a successor exists because of the parts 2 and 4 of τ. If i is even, we define v as a successor of w_{i-1} that satisfies p_3 — it exists because of 3. If v is already on the path, then we end the construction, otherwise we put $w_i = v$.

Since $\mathfrak{M}$ is finite, the above construction terminates. It means that there is some $r < l$ such that $\mathfrak{M} \models w_l K w_r$. Clearly, w_l and w_k are ∞-inner. It follows from Lemma 8.2 that $\mathfrak{M} \models w_{l-r} K w_0$. But w_{l-r} satisfies p_2 or p_3, w_0 satisfies p_0, and parts 3 and 4 of τ forbid such connections. Therefore there is no finite model of τ based on a frame from $\mathcal{K}_\Phi$. $\Box$

In order to show the decidability we provide some auxiliary lemmas. We start from a simple property of the global satisfiability problem — each satisfiable formula has a model, which is strongly connected.

Lemma 8.4. *Let $\Phi \in$ UHF, φ be a modal formula and $\mathfrak{M}$ be a finite $\mathcal{K}_\Phi$-based structure such that $\mathfrak{M} \models \varphi$. Then there is a $\mathcal{K}_\Phi$-based substructure $\mathfrak{N}$ of $\mathfrak{M}$ such that $\mathfrak{N} \models \varphi$ and the frame of $\mathfrak{N}$ is strongly connected.*

Proof. Let $\mathfrak{N}_1$, $\mathfrak{N}_2$, $\ldots \mathfrak{N}_k$ be a partition of $\mathfrak{M}$ into maximal (w.r.t. number of worlds) strongly connected components such that for any $i > j$ there is no path from (any world of) $\mathfrak{N}_i$ to $\mathfrak{N}_j$. Such a partition exists since the relation "there is a path from v to w" is a preorder and, while considered on maximal strongly connected components, it is antisymmetric (if there is a path from $\mathfrak{N}_i$ to $\mathfrak{N}_j$ and from $\mathfrak{N}_j$ to $\mathfrak{N}_i$, then $\mathfrak{N}_i$ and $\mathfrak{N}_j$ are not maximal) and therefore it is an order.

We put $\mathfrak{N} = \mathfrak{N}_k$. Clearly, $\mathfrak{N}$ satisfies Φ since Φ is universal. Moreover, since each world from $\mathfrak{N}$ has all its successors in $\mathfrak{N}$ (there is no path to worlds in other connected components), $\mathfrak{N}$ satisfies φ. $\square$

We say that a frame $\mathcal{M}$ is *k-periodic* if it consists of a pairwise disjoint, non-empty sets of worlds W_1, W_2, $\ldots$, W_k such that for each v, w from $\mathcal{M}$ there is an edge from v to w if and only if for some $i \leq k$, $v \in W_i$ and $w \in W_{(i \mod k)+1}$. Notice that 1-periodic frame is a clique. For each $k \in \mathbb{N}$ we define the cycle $\mathcal{C}_k$ as $\mathcal{I}_k$ with one additional edge, namely $(\underline{k-1}, \underline{0})$. Clearly, each $\mathcal{C}_k$ is k-periodic.

We are going to prove the decidability by showing that each satisfiable formula has a model that is k-periodic for some k. In order to do so, we have to prove two technical lemmas.

Lemma 8.5. *Let* $\Phi \in \mathsf{UHF}$*.*

(a) If Φ has a k-periodic model $\mathcal{M}$, then $\mathcal{C}_k$ is a model of Φ.

(b) If $\mathcal{C}_k$ is a model of Φ, then any k-periodic frame is a model of Φ.

(c) If $\mathcal{L}_{\mathbb{Z}}$ is a model of Φ, then for all $c > |\Phi|$, $\mathcal{C}_c$ is a model of Φ.

(d) If for some $c > |\Phi|$ the frame $\mathcal{C}_c$ is a model of Φ, then $\mathcal{L}_{\mathbb{Z}}$ is a model of Φ.

Proof. For (a), observe that if a periodic model $\mathcal{M}$ that consists of sets W_1, W_2, $\ldots$, W_k is a model of Φ, then $\mathcal{C}_k$ is isomorphic with an induced substructure of $\mathfrak{M}$ that consists of one world from every W_i.

In the rest of the proof we use one more definition. We say that a morphism $h : \mathcal{M} \to \mathcal{M}'$ is *complete* if for all v, v' we have $h(v)Rh(v')$ if and only if vRv'. Note that if there is a complete morphism $h : \mathcal{M} \to \mathcal{M}'$ and Φ does not hold in $\mathcal{M}$, then it does not hold in $\mathcal{M}'$.

For (b), assume that there is a periodic frame $\mathcal{M}$ that consists of sets W_1, W_2, $\ldots$, W_k and is not a model of Φ, but $\mathcal{C}_k$ is a model of Φ. We define a complete morphisms $f : \mathcal{M} \to \mathcal{C}_k$ as $f(v) = \underline{i}$ for $v \in W_i$. Since Φ does not hold in $\mathcal{M}$ and f is a complete morphism, Φ does not hold in $\mathcal{C}_l$ — a contradiction.

We prove (c) as follows. Let $c > |\Phi|$. Assume that there is a clause Ψ satisfied in $\mathcal{L}_{\mathbb{Z}}$ but not in $\mathcal{C}_c$, and let $v_1, v_2, \ldots, v_n$ be worlds of $\mathcal{C}_c$ such that $\Psi(v_1, \ldots, v_n)$ is false. Let k be such that no world among $v_1, \ldots, v_n$ is equal $\underline{\mathrm{k}}$. Consider the function $f : \mathcal{C}_{c \restriction \{v_1, \ldots, v_n\}} \to \mathcal{L}_{\mathbb{Z}}$ defined as

$$f(\underline{s}) = \begin{cases} \underline{s} & \text{for } s > k \\ \underline{c+s} & \text{for } s < k \end{cases}$$

A quick check shows that the function f is a complete morphism. Since $\Psi(v_1, \ldots, v_n)$ does not hold in $\mathcal{C}_c$, it follows that $\Psi(f(v_1), \ldots, f(v_n))$ does not hold in $\mathcal{L}_{\mathbb{Z}}$. But $\mathcal{L}_{\mathbb{Z}} \models \Psi$, a contradiction.

For the proof of (d), let $k > |\Phi|$, $\Psi \Rightarrow \Psi'$ be satisfied in $\mathcal{C}_k$ but not in $\mathcal{L}_{\mathbb{Z}}$. Let $v_1 = \underline{s}, v_1 = \underline{t}, v_3 \ldots, v_n$ be worlds of $\mathcal{L}_{\mathbb{Z}}$ such that $\Psi(v_1, \ldots, v_n)$ is true, $\Psi'(v_1, \ldots, v_n)$ is not, and $|s-t|$ is minimal. Let $f(\underline{i}) = \underline{i \mod k}$ be a morphism from $\mathcal{L}_{\mathbb{Z}}$ onto $\mathcal{C}_k$. If $t - s \mod k \neq 1$, $\Psi \Rightarrow \Psi'(f(v_1), \ldots, f(v_n))$ does not hold and we have a contradiction. Otherwise, $|s-t| \geq k-1$ so there is a world $\underline{l}$ such that l is between s and t and $\underline{l}$ is different from all of $\underline{s}, \underline{t}, v_3, \ldots, v_n$. Then, morphism $g : \mathcal{L}_{\mathbb{Z}\restriction\{v_1,\ldots,v_n\}} \to \mathcal{L}_{\mathbb{Z}}$ defined as $g(\underline{s}) = \underline{s}$ for $s < l$ and $g(\underline{s}) = \underline{s-1}$ otherwise leads to the contradiction with the minimality of $|s-t|$. $\square$

Lemma 8.6. *Let Φ be an unbounded* UHF *formula that does not force long edges. Assume that for some $i, j < 0$, $\mathfrak{C}_\Phi(\mathcal{X})$ contains an edge $(\overline{i}, \underline{j})$ or $(\underline{i}, \overline{j})$. Then $j - i = 1$ and $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}}) = \mathcal{L}_{\mathbb{Z}}$.*

Proof. As $\mathcal{X}$ is symmetric, $\mathfrak{C}_\Phi(\mathcal{X}) \models \underline{i}R\overline{j}$ implies $\mathfrak{C}_\Phi(\mathcal{X}) \models \overline{i}R\underline{j}$. So we assume that $\mathfrak{C}_\Phi(\mathcal{X}) \models \underline{i}R\overline{j}$.

Let us consider a morphism f from $\mathcal{X}$ into $\mathcal{L}_{\mathbb{Z}}$ defined as

$$f(\underline{k}) = f(\overline{k}) = \underline{k}$$

If $|j-i| > 1$, then there is a long edge in $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ and it contradicts the assumption that $\mathcal{X}$ does not force long edges.

If $j - i = -1$, then the morphism f implies that there is an edge $(\underline{j}, \underline{j-1})$ in $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ and, since $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ is uniform, for all k there are edges $(\underline{k}, \underline{k-1})$ in $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$. We define another morphism g to show that then $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ contains a long edge. Let g be a morphism from $\mathcal{X}$ into $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ defined as

$$g(w) = \begin{cases} \underline{|k|} & \text{if } w = \underline{k} \text{ for some } k \\ \underline{-|k|} & \text{if } w = \overline{k} \text{ for some } k \end{cases}$$

It is not hard to see that g is a morphism and therefore that $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ contains a long edge $(\underline{|i|}, \underline{-|j|})$. An example is presented in Fig. 8.1.

If $j = i$, then the morphism f implies that there is a reflexive world in $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$, and therefore all worlds are reflexive. Consider a morphism h from $\mathcal{X}$ into $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ defined as

$$h(w) = \begin{cases} \underline{1} & \text{if } w = \overline{k} \text{ for some } k \leq i \\ \underline{0} & \text{otherwise} \end{cases}$$

Since all worlds in $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ are reflexive, h is indeed a morphism, so there is edge $(\underline{1}, \underline{0})$ in $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ and, as in the previous case, all edges in $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ are symmetric and therefore $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ contains a long edge.

For the proof of $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}}) = \mathcal{L}_{\mathbb{Z}}$, recall that if $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$ contains a symmetric or reflexive edge, then it contains long edges. But Φ does not force long edges, and therefore $\mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}}) = \mathcal{L}_{\mathbb{Z}}$. $\square$

In particular, the above lemma shows that if a first-order formula Φ does not force long edges and merges, then $\mathfrak{C}_\Phi(\mathcal{X})$ does not contain symmetric or reflexive edges.

Proposition 8.7. *Let Φ be an unbounded* UHF *formula that does not force long edges, merges at a level $k < 0$ and forks at all levels. Then finite global $\mathcal{K}_\Phi$-SAT is in* PSPACE.

Proof. Let φ be a modal formula and $\mathfrak{M}$ be a model of φ and Φ that is a strongly connected component. Such a model exists due to Lemma 8.4. Assume that $\mathfrak{M}$ contains at least two worlds.

We define a *characteristic cycle* of $\mathfrak{M}$ as a walk $v_0, v_1, \ldots, v_{l-1}$ that contains all worlds from $\mathfrak{M}$ and, moreover, in $\mathfrak{M}$ there is an edge from v_{l-1} to v_0. Note that such a characteristic cycle exists because $\mathfrak{M}$ is strongly connected. For the better readability, below we omit " mod l" in subscripts of vs.

Our aim is to show that $\mathfrak{M}$ is s-periodic for some s.

Let $\mathcal{X}_\mathfrak{M} \subseteq \mathbb{N}$ be such that $k \in \mathcal{X}_\mathfrak{M}$ if and only if there is v_i such that $\mathfrak{M} \models v_i R v_{i+k+1}$. Lemma 8.2 implies that for all v_i and $k \in \mathcal{X}_\mathfrak{M}$, $\mathfrak{M} \models v_i R v_{i+k+1}$.

We show that $\mathcal{X}_\mathfrak{M}$ is additively closed. Assume that $x, y \in \mathcal{X}_\mathfrak{M}$. It means that $\mathfrak{M}$ contains edges (v_{x+y+1}, v_{x+y+2}) , (v_{x+1}, v_{x+y+2}) and (v_0, v_{x+1}). We define a morphisms h from $\mathcal{X}$ to $\mathcal{M}$, the frame of $\mathfrak{M}$, as

$$h(w) = \begin{cases} v_s & \text{if } w = \underline{k-1+s} \text{ for all } s \leq 0 \\ v_{x+1} & \text{if } w = \underline{k} \\ v_{x+y+1+s} & \text{if } w = \underline{\underline{k+s}} \text{ for all } s > 0 \\ v_{x+y+1+s} & \text{if } w = \overline{k+s} \text{ for all } s \in \mathbb{Z} \end{cases}$$

We see that $h(\underline{k-1}) = v_0$ and $h(\overline{k}) = v_{x+y+1}$, and since in $\mathfrak{M}$ there is an edge from $\underline{k-1}$ to $\overline{k}$, $x + y \in \mathcal{X}_\mathfrak{M}$.

Let $\mathcal{X}^l_\mathfrak{M} = \{i \mod l | i \in \mathcal{X}_\mathfrak{M}\}$. By Fact 4.1, $\mathcal{X}^l_\mathfrak{M}$ can be represented as $\{i \cdot gcd(\mathcal{X}_\mathfrak{M}) \mod l | i \in \mathbb{N}\}$. Define $W_i = \{v_{i+j \cdot gcd(\mathcal{X}_\mathfrak{M})} | j \in \mathbb{N}\}$. It follows that all elements of W_i all successors in W_{i+1}, and therefore $\mathfrak{M}$ is $gcd(\mathcal{X}_\mathfrak{M})$-periodic.

We need to compress sets W_i. For each i and each subformula ψ of φ, if there is a world in W_i that satisfies ψ, we mark one such world. Then we remove unmarked worlds. It is easy to see that the types of worlds remain the same.

We have proved that all models of φ are s-periodic with the sets of the size bounded by $|\varphi|$, but the value of s can be arbitrary large. Now we show that there is a NPSPACE (=PSPACE) procedure that checks, for a given modal formula φ, if φ has a Φ-based finite global periodic model.

The NPSPACE algorithm works as follows. First, it checks if there is a single world or a single clique (1-periodic set) with the size bounded by φ that satisfy φ and Φ and if there is, it returns "Yes". Otherwise, it guesses a set W_1 with size bounded by $|\varphi|$ and then, recursively, it guesses the successive sets with size bounded by the same number, checking if guessed worlds are consistent with their predecessor, and returns "no" otherwise. The algorithm stops after $\binom{2^{|\varphi|}}{|\varphi|} + 1$ steps and returns "yes".

If there is a model of φ, then the algorithm returns "yes". Indeed, we showed that φ has single world model or a s-periodic model with size of sets bounded by $|\varphi|$, and the algorithm can simply guess this world or guess the consecutive sets of this model.

If the algorithm returns "yes", then it visited two sets satisfying the same subformulas, so there is a sequence of sets $V_1, V_2, \ldots, V_k, V_1$ with $k \leq 2^{|\varphi|}$ such that each set contains all witnesses needed by its predecessors. We build a s-periodic model that contains sets $V_1, \ldots, V_k$ repeated $\lceil |\Phi|/k \rceil + 1$ times. Clearly, the obtained structure satisfies φ. By Lemma 8.6 $\mathcal{L}_{\mathbb{Z}} = \mathfrak{C}_\Phi(\mathcal{L}_{\mathbb{Z}})$, and by Lemma 8.5 the obtained structure is a model of Φ. $\square$

The corresponding lower bound follows from the encoding of the bounded-space domino problem.

Proposition 8.8. *Let Φ be an unbounded* UHF *formula that does not force long edges, merges at a level $k < 0$ and forks at all levels. Then finite global $\mathcal{K}_\Phi$-SAT is* PSPACE-*hard.*

Proof. We encode the bounded-space domino problem, keeping whole row of a solution in one world. Let $\langle \mathcal{D}, V_\mathcal{D}, H_\mathcal{D} \rangle$ be an instance of the bounded bounded-space domino problem and $n = O(|\mathcal{D}|)$. We build a modal formula using variables p_i^d for $i < n$ and $d \in \mathcal{D}$. The intended meaning of p_i^d is that the point in the column i is tiled by d. We define a formula τ^l that guarantees that each point is tiled by exactly one element of $\mathcal{D}$ as follows.

$$\tau^l = \bigwedge_{i<n} (\bigvee_{d \in \mathcal{D}} p_i^d \wedge \bigwedge_{d,d' \in \mathcal{D}, d \neq d'} \neg (p_i^d \wedge p_i^{d'}))$$

Formula τ^h will ensure that the tilling is consistent with the relation $H_\mathcal{D}$.

$$\tau^h = \bigwedge_{i<n-1} \bigvee_{(d,d') \in H_\mathcal{D}} (p_i^d \wedge p_{i+1}^{d'})$$

We ensure that each world has a successor that describes the row which is consistent with the current one with respect to the relation $V_\mathcal{D}$.

$$\tau^v = \Diamond \top \wedge \bigwedge_{i<n} \bigvee_{(d,d') \in V_\mathcal{D}} (p_i^d \wedge \Box p_i^{d'})$$

Finally, we put $\tau = \tau^l \wedge \tau^h \wedge \tau^v$. If $\langle \mathcal{D}, V_\mathcal{D}, H_\mathcal{D} \rangle$ has a solution that consists of rows $r_1, r_2, \ldots$, then among first $n^n + 1$ of them some rows r_i, r_j with $i < j$ have to be the same. Let $l = c(j - i)$. We encode the solution on $\mathfrak{C}_\Phi(\mathcal{C}_l)$, such that $\underline{s}$ represents a configuration $i + (s \mod j)$. That ends the proof. $\square$

Now we prove that formulas that do not force long edges, fork at all levels and do not merge at any level, define logics with the finite model property. In the proof, we start from an infinite tree–based model $\mathfrak{M}$, and construct a very large structure that locally looks like a part of $\mathfrak{M}$, but is finite. We need to do it carefully in order not to violate the first–order formula.

Proposition 8.9. *Each unbounded* UHF *formula* Φ *that does not force long edges, forks at all levels and does not merge at any level* $k < 0$ *has the finite model property in the global case.*

Proof. Let $\mathfrak{M}^b$ be a tree-based model of φ and Φ based on a tree $\mathcal{T}^b$, $n = |\varphi|$ and $N = |\Phi|$. If there is a world in $\mathfrak{M}^b$ without a proper successors, then the structure that contains only this world is a model of φ and Φ. Otherwise, all worlds are ∞-followed. We assume that the degree of all worlds is equal n — if it is smaller, we can duplicate any subtree.

Let w be any $\mathfrak{g}(|\Phi|)$-inner world in $\mathcal{T}^b$, $\mathcal{T}$ be a subtree of $\mathcal{T}^b$ induced by w, and $\mathfrak{M}$ be a substructure of $\mathfrak{M}^b$ that consists of the worlds from $\mathcal{T}$. Clearly, $\mathfrak{M}$ satisfies Φ and Property (i) of Lemma 6.2 guarantees that it is a model of φ.

Let M be a set of worlds in $\mathfrak{M}$. For each $w \in M$, we define a tree $\mathcal{S}'_w$ as a subtree of $\mathcal{M}$ rooted in w, $\mathcal{S}_w$ as a structure that contains first $2N$ levels of $\mathcal{S}'_w$, and $\mathfrak{S}_w$ as a substructure of $\mathfrak{M}$ that contains the worlds from $\mathcal{S}_w$. Let $\mathcal{F}$ be a set of all types realized in $\mathfrak{M}$. For each type $t \in \mathcal{F}$, we pick one world w_t of this type and define $\mathfrak{S}_t = \mathfrak{S}_w$ and $\mathcal{S}_t = \mathcal{S}_w$

For each $\mathcal{S}_t$, we label leaves in $\mathcal{S}_t$ in a consecutive way, e.g. from left to right, such that leaves labeled with 1, 2, …, n have the same parent and so on.

For each $s \in \{0, 1\}$, $p \in \mathbb{P}$ and $t \in \mathcal{F}$, we define $\mathfrak{T}^s_{t,p}$ as a copy of $\mathfrak{S}_t$. We define the finite structure $\mathfrak{M}_s$ as a disjoint union of all possible $\mathfrak{T}^s_{t,p}$. We say that a world w is *at a level* k in $\mathfrak{T}^s_{t,p}$ if it is a copy of a world that is at a level k in $\mathcal{S}_t$ and that it is *at a level* k in $\mathfrak{M}_s$ if it is at a level k in some tree of $\mathfrak{M}_s$. We say that a world v is a *parent* of v' in $\mathfrak{M}_k$ if wRv, v is at a level k and v' is at a level $k+1$ for some k. For any two worlds v, v' that are in the same tree, we define $lca(v, v')$ as the lowest common ancestor of v and v' (w.r.t. the relation parent). We define $llca(v, v')$ as the level of $lca(v, v')$ if such world exists and $llca(v, v') = -1$ otherwise.

We define a structure $\mathfrak{M}'$ as a disjoint union of $\mathfrak{M}_0$ and $\mathfrak{M}_1$ with additional edges defined as follows. Consider a tree $\mathfrak{T}^0_{t,p}$ and its leaf v labeled by p. Let w be a world in $\mathfrak{M}$ with the same type and $t_1, \ldots, t_k$ be types of successors of w in $\mathcal{T}$. For each $j \leq k$ we add an edge from v to the root of $\mathfrak{T}^1_{t_j,p}$ and, if some connection between w and its successors is symmetric, we make this edge symmetric as well. We do the same for the leaves from $\mathfrak{M}_1$, but we connect them with the roots from $\mathfrak{M}_0$.

It is not hard to see that all worlds in $\mathfrak{M}'$ satisfy φ. We prove that $\mathfrak{M}'$ satisfies Φ.

Assume otherwise and let $\Psi \Rightarrow \Psi'$ be the formula not satisfied in $\mathfrak{M}'$. There are worlds $v_1, \ldots, v_n$ such that $\Psi(v_1, \ldots, v_N)$ holds but $\Psi'(v_1, \ldots, v_N)$ does not.

We define a function $\nu_k : \mathfrak{M}' \to \{0, \ldots, 4N-1\}$ as

$$\nu_k(v) = \begin{cases} s - k & \text{for each } v \text{ at a level } s \geq k \text{ in } M_0 \\ s + 2N - k & \text{for each } v \text{ at a level } s \text{ in } M_1 \\ s + 4N - k & \text{for each } v \text{ at a level } s < k \text{ in } M_0 \end{cases}$$

Let k be such that no world among $v_1, \ldots, v_n$ is at a level k in $\mathfrak{M}_0$ and $\mathfrak{M}_1$. A function $f : \mathfrak{M}'{\restriction_{\{v_1,\ldots,v_n\}}} \to \mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ defined as $f(v) = \underline{\nu_k(v)}$ is a morphism.

It is not possible that $\Psi' = \bot$, because then Φ would not be satisfied in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ and since Φ is unbounded, $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ is a model of Φ . Similarly, if $\Psi' = xRx$, then some world in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ would be reflexive and, since all worlds in $\mathfrak{M}$ are $\mathfrak{g}(|\Phi|)$-inner in $\mathfrak{M}^b$, all worlds in $\mathfrak{M}$ and so $\Psi'(v_1, \ldots, v_N)$ would be satisfied.

The only remaining case is $\Psi' = xRy$. Let v_1 be at a level l_1 in $\mathfrak{M}_{s_1}$ and v_2 be at a level l_2 is $\mathfrak{M}_{s_2}$. There are two cases: either $s_1 = s_2$ and $|l_1 - l_2| \leq 1$, or $s_1 \neq s_2$ and one of v_1, v_2 is a root and the other one is a leaf. Otherwise, Φ would force long edges.

Assume that $s_1 < s_2$ and let k be such that no world among $v_1, \ldots, v_n$ is at a level k in $\mathfrak{M}_0$. Consider a morphism $g : \mathfrak{M}'{\restriction_{\{v_1,\ldots,v_n\}}} \to \mathfrak{M}'$ defined as

$$g(v) = \begin{cases} v' & \text{if } v \text{ is at a level } i \geq k \text{ in } \mathfrak{M}_0 \text{ and } v' \text{ is a parent of } v \\ v & \text{otherwise} \end{cases}$$

It implies that Φ requires also an edge for some world that is not a leaf to some root, and so by a morphism f we can show that Φ forces long edges. The case when $s_1 > s_2$ is symmetric.

Assume that $s_1 = s_2 = 0$. If $v_1 = v_2$, then, by a morphism f, all worlds of $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ are reflexive and Ψ' would be satisfied, as before. If v_2 is a parent of v_1, then, by a morphism f, all edges in $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ are symmetric and Ψ' would be satisfied. So we can assume that v_1 and v_2 are not on the same path in $\mathfrak{M}_0$.

Assume that $l_1 \leq N$ and $l_2 \leq N$ and let $k > N$ be such that no world among $v_3, \ldots, v_N$ is at the level k in $\mathfrak{M}_0$. We define a morphism $h_1 : \mathfrak{M}'{\restriction_{\{v_1,\ldots,v_n\}}} \to \mathcal{T}_\infty$ as follows.

$$h_1(v) = \begin{cases} \underline{0^{\nu_k(v)}} & \text{if } \nu_k(v) < 4N - k \\ \underline{0^{4N-k+llca(v,v_1)}1^{s-llca(v,v_1)}} & \text{if } v \text{ at level } s \text{ and } \nu_k(v) \geq 4N - k \end{cases}$$

Let $m = llca(v_1, v_2)$. Since v_1 and v_2 are not on the same path, $m < min(l_1, l_2)$. Since $h_1(v_1) = \underline{0^{4N-k+l_1}}$ and $h_1(v_2) = \underline{0^{4N-k+m}1^{l_2-m}}$ and h_1 is a morphism, it implies that Φ does not fork a the level $\underline{0^{4N-k+m}}$ — a contradiction.

Now consider the case when $l_1 \geq N$ and $l_2 \geq N$. Let $k < N$ be such that no world among $v_3, \ldots, v_N$ is at the level k in $\mathfrak{M}_0$.

If $llca(v_1, v_2) \leq k$, then Φ merges at some level. We prove it using the following morphism $h_2 : \mathfrak{M}'{\restriction_{\{v_1,\ldots,v_n\}}} \to \mathcal{X}$. Let $\mathfrak{T}^0_{t,p}$ be the tree that contains v_1.

$$h_2(v) = \begin{cases} \underline{\underline{s - 2N}} & \text{if } v \text{ at a level } s \geq k \text{ in } \mathfrak{M}_0 \text{ and } llca(v_1, v) > k \\ \underline{s - 2N} & \text{if } v \text{ at a level } s \geq k \text{ in } \mathfrak{M}_0 \text{ and } llca(v_1, v) \leq k \\ \underline{s} & \text{if } v \text{ at a level } s \text{ in } \mathfrak{M}_1 \\ \underline{2N + s} & \text{if } v \text{ at a level } s \text{ in } \mathfrak{M}_0 \end{cases}$$

It is readily checkable that h_2 is a morphism and it implies that Φ merges at some level.

Let $llca(v_1, v_2) > k$. We prove that Φ does not fork at some level. To this end, let k' be such that no world among $v_3, \ldots, v_N$ is at the level k' in $\mathfrak{M}_1$. We define $V_1 = V_{M0} \cup V_{M1}$ as follows. Set $v \in V_{M0}$ if and only if v is at a level $s > k$ in $\mathfrak{M}_0$ and $lcm(v_1, v) \in \{v_1, v\}$ (in other worlds, v is an ancestor or descendant of v_1 in $\mathfrak{M}_0$). Finally, for each leaf w from V_{M0} labeled by m and each $t \in \mathcal{F}$, V_{M1} contains all worlds from levels less than k' in $\mathfrak{T}^1_{t,m}$.

Let $t = llca(v_1, v_2) - k$, We define a morphism $h_3 : \mathfrak{M}'{\restriction_{\{v_1,\ldots,v_n\}}} \to \mathcal{T}_\infty$.

$$h_3(v) = \begin{cases} \underline{0^{\nu_k(v)}} & \text{if } v \in V_1 \text{ or } \nu_k(v) < t \\ \underline{0^t 1^{\nu_k(v)-t}} & \text{otherwise} \end{cases}$$

It is readily checkable that h_3 is a morphism and it implies that Φ does not fork at the level t.

The case when $s_1 = s_2 = 1$ is symmetric. □

In the case of formulas that do not force long edges and do not fork at some level, the finite model property follows from the fact that each satisfiable formula has a k-periodic model for some k.

Proposition 8.10. *Each unbounded* UHF *formula Φ that does not force long edges and does not fork at some level $k > 0$ has the finite model property in global case.*

Proof. First, observe that $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) = \mathcal{L}_\mathbb{Z}$ and, since Φ is unbounded, $\mathcal{L}_\mathbb{Z}$ is a model of Φ. Let v be a world at a level $\mathfrak{g}(|\Phi|)$ and $\mathfrak{M}'$ be a model that consists of all descendants of v at levels greater than $2\mathfrak{g}(|\Phi|)$. By Lemmas 6.12, all worlds in $\mathfrak{M}'$ at the same level are equivalent. Since the number of types is bounded, there exist two levels k, l in $\mathfrak{M}'$ such that $k - l > |\Phi| + 1$ and the sets of types realized at a level k and l are equal. We create model $\mathfrak{M}''$ by removing all worlds at a level greater than or equal to k and connecting all worlds from level $k-1$ to worlds from level l. Finally, we define a model $\mathfrak{M}'''$ by taking for each level one world of each type realized at this level. A quick check shows that all $\mathfrak{M}'$, $\mathfrak{M}''$, and $\mathfrak{M}'''$ satisfies φ and that $\mathfrak{M}'''$ is finite.

Now we justify that $\mathfrak{M}'''$ is a model of Φ. Since $\mathcal{L}_\mathbb{Z}$ is a model of Φ, Lemma 8.5 shows that C_{k-l} is a model of Φ, and the same lemma shows that therefore any $k-l$–periodic model is a model of Φ. Model $\mathfrak{M}'''$ is obviously $k-l$-periodic. □

8.2. Formulas that force long edges

For the formulas that satisfy S3, we proved the polynomial model property in Section 7. The rest of this section is devoted to formulas $\Phi \in$ UHF that satisfy S2.

First, observe that the modal logic defined by Φ can lack the finite model property. Consider, for example, $(xRz_1 \wedge z_1Ry \Rightarrow xRy) \wedge (xRx \Rightarrow \bot)$ and a modal formula $\Diamond\top \wedge \Box\Diamond\top$. A quick check shows that all models of these formulas are infinite (in local and global cases). On the other hand, a modal logic defined by a formula $xRx \wedge (xRz_1 \wedge z_1Ry \Rightarrow xRy)$ has the finite model property.

We show that modal logic over $\mathcal{K}_\Phi$ has the following property: if a formula φ has a finite model, then it has a model of size bounded by $|\varphi|^{O(|\varphi|)}$. Clearly, it leads to a NExpTime algorithm that simply guesses such a model and verifies it.

Consider a modal formula φ and its $\mathcal{K}_\Phi$-based model $\mathfrak{M}$ with universe M. We say that a world w is *redundant* for φ and model $\mathfrak{M}$ if $\mathfrak{M}_{\restriction M \setminus \{w\}}$ is a model of φ. We prove the following lemma by showing that a model that is large enough has to contain a redundant world.

Lemma 8.11. *Let Φ be unbounded* UHF *formula that forces long edges. If φ has a finite $\mathcal{K}_\Phi$-based model, then it has a model of the size bounded by $|\varphi|^{O(|\varphi|)}$.*

Proof. Let Φ be an unbounded UHF formula that satisfies S2 for some l and $a_1, \ldots, a_l$ and φ be a modal formula with a $\mathcal{K}_\Phi$-based model $\mathfrak{M}$.

Let $c = a_1$. Observe that for all $i \in \mathbb{Z}$ and $k \geq 0$ we have $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z}) \models \underline{i} R \underline{i + kc + 1}$.

We start from bounding the number of worlds that are not $\mathfrak{g}(|\Phi|)$-preceded. We use the standard selection technique [4] — we start from an arbitrary world that satisfies φ, and then recursively for each world added in the previous stage we pick at most $|\varphi|$ witnesses. Let $\mathfrak{M}'$ be a model obtained this way. We define *the royal part* of $\mathfrak{M}'$ as the set of worlds that contain all worlds that are not $\mathfrak{g}(|\Phi|)$-preceded and *the court* as the set of $\mathfrak{g}(|\Phi|)$-preceded worlds that were added as witnesses for some worlds from the royal part. Clearly, the sizes of the royal part and the court can be bounded by $|\varphi|^{\mathfrak{g}(|\Phi|)+1}$.

Let w be a $\mathfrak{g}(|\Phi|)$-inner world not from the court such that for each subformula $\Diamond\psi$ of φ such that ψ is satisfied in w there exists a $\mathfrak{g}(|\Phi|)$-inner world $w_\psi \neq w$ that satisfies ψ and that there is a path from w to w_ψ with the length cj for some j. We show that w is redundant.

Consider any predecessor w' of w. If w' is not $\mathfrak{g}(|\Phi|)$-preceded, then it has all the required witnesses in the court and the royal part. Otherwise, let ψ be a subformula of φ such that w satisfies ψ. We show that there is an edge from w' to w_ψ. To this end, consider a path $v_1, v_2, \ldots, v_{\mathfrak{g}(|\Phi|)},\ w', w,\ v'_1, v'_2, \ldots, v'_{cj},\ w_\psi, v''_1, v''_2, \ldots, v''_{\mathfrak{g}(|\Phi|)}$. Such a path exists since w' is $\mathfrak{g}(|\Phi|)$-preceded and w_ψ is $\mathfrak{g}(|\Phi|)$-inner, and there is a straightforward morphism from $\mathcal{I}_{2\mathfrak{g}(|\Phi|)+2+cj}$ into this path. So it is enough to show that there is an edge from $\underline{\mathfrak{g}(|\Phi|) + 1}$ to $\underline{\mathfrak{g}(|\Phi|) + 1 + cj + 1}$ in $\mathfrak{C}(\mathcal{I}_{2\mathfrak{g}(|\Phi|)+2+j})$. By earlier observations, $\mathfrak{C}_\Phi(\mathcal{L}_\mathbb{Z})$ contains an edge from $\underline{\mathfrak{g}(|\Phi|) + 1}$ to $\underline{\mathfrak{g}(|\Phi|) + 1 + cj + 1}$, and Lemma 6.10 implies that there is an edge from $\underline{\mathfrak{g}(|\Phi|) + 1}$ to $\underline{\mathfrak{g}(|\Phi|) + 1 + cj + 1}$ in $\mathfrak{C}(\mathcal{I}_{2\mathfrak{g}(|\Phi|)+2+cj})$.

By iterating the above argument we can remove all $\mathfrak{g}(|\Phi|)$—inner worlds except for at most $|\varphi|^{c \cdot |\varphi|}$ worlds. Finally, we again use the selection technique to bound the number of worlds that are not $\mathfrak{g}(|\Phi|)$-followed by $|\varphi|^{c \cdot |\varphi|} \cdot |\varphi|^{\mathfrak{g}(|\Phi|)}$. Since Φ is not a part of an instance, we reduced the number of worlds to $|\varphi|^{O(|\varphi|)}$. $\square$

The above lemma leads to the following result.

Proposition 8.12. *If Φ is unbounded* UHF *formula that forces long edges, then finite $\mathcal{K}_\Phi$-SAT and finite global $\mathcal{K}_\Phi$-SAT are in* NExpTime.

Part III.

Halpern–Shoham logic

9. Overview

9.1. Main theorems

Our contribution consists of the proofs of the following two theorems:

Theorem 9.1. *The satisfiability problem for the formulae of the logic of subintervals, over models which are suborders of the order $\langle \mathbb{Z}, \leq \rangle$, is undecidable.*

Since truth value of a formula is defined with respect to a model and an initial interval in this model (see Preliminaries), and since the only allowed operator is D, which means that the truth value of a formula in a given interval depends only on the labeling of this interval and its subintervals, Theorem 9.1 can be restated as: *The satisfiability problem for the formulae of the logic of subintervals, over finite models is undecidable*, and it is this version that will be proved in Section 10.

Theorem 9.2. *The satisfiability problem for the formulae of the logic of subintervals, over all discrete models, is undecidable.*

An overview of the proofs. One possible source of undecidability, and the one we are making use of, is the interaction of regularity and measurement. Consider the following example proposition:

Proposition 9.3. *The problem:*

For a given regular language $L \subseteq \Sigma^*$ and a given set $B \subseteq \Sigma^2$, do there exist a natural number n and a word $w \in L$ such that $|w|$ (the length of w) is greater than n and for each sub-word avb of w (where $a, b \in \Sigma$), if the length of avb is n, then $\langle a, b \rangle \in B$?

is undecidable.

The proposition is obvious – if we can make sure that any two symbols in the word, which are at distance n, are a "correct pair", then we can easily encode a Turing machine.

In Sections 10.2 and 10.3, we show how is it possible, in the logic of subintervals, to encode any regular language.

But encoding the measurement is not that simple. The logic of subintervals is not able – as far as we know – to measure the length of each sub-word of w. We need to mark each endpoint of the measured interval by a symbol that does not occur inside this interval. This means that we can only afford a bounded number of measurements taking place at the same time. Imagine we had four identical hourglasses, which we are free to turn at any moment while reading consecutive symbols of the word. This would not be enough to directly encode a Turing machine, but still enough for undecidability. In Section 10.1 we describe a class of regular languages (one regular language for each Minsky machine) for which the possibility of such four simultaneous measurements leads to an undecidable non-emptiness problem. This property is stated in Lemma 10.2 which is a counterpart of Proposition 9.3.

In Section 10.4, we define our measuring tool, which we call a cloud, and in Section 10.5, which completes the proof of Theorem 9.1, we show how to use it. Actually, the idea here is very simple: how much you see is a monotonic function of how high you are.

In the proof of Theorem 9.1, the measuring device, the cloud, is existentially quantified. Its role is identical with the role of the number n in Proposition 9.3. "They" provide it, together with the word w, and we only check that all the specification conditions are met. This approach would not work in the situation of Theorem 9.2. The reason for that is that the logic of subintervals gives no means (that we are aware of) to specify the requirement that all the intervals of the cloud are finite (i.e., contain a finite number of elements of the order). Or – using other words – that time periods measured by the hourglasses are finite. This could lead to pathologies that we do not even want to think about. Instead, in Section 11, we build our own hourglass which we call the parabola. It is not as good as the cloud – its size increases from time to time. But a closer look at Lemma 10.2 shows that we can live with it. And, unlike the cloud, the parabola does not suffer from the possible pathologies of discrete orderings.

Related work. The results presented in this part are based on the LICS paper [25]. However, this thesis contains some additional results and more detailed proofs.

9.2. Preliminaries

Orderings. Originally, Halpern–Shoham logic was defined for any order that satisfy the "linear interval property", i.e. for each a, c_1, c_2, b if $a \leq c_1$, $a \leq c_2$, $c_1 \leq b$, and $c_2 \leq b$, then $c_1 \leq c_2$ or $c_2 \leq c_1$. In such orderings, when we restrict our attention to the operators that look only "inside" of an initial interval, such as D, the reachable part of the ordering is totally ordered. For that reason in the rest of this paper we consider only the total orderings.

As in [15], we say that a total order $\langle \mathbb{D}, \leq \rangle$ is *discrete* if each element is either minimal (maximal) or has a predecessor (successor); in other words for all $a, b \in \mathbb{D}$ if $a < b$, then there exist points a', b' such that $a < a'$, $b' < b$ and there exists no c with $a < c < a'$ or $b' < c < b$.

Semantics of the D fragment of logic HS (logic of subintervals). Let $\langle \mathbb{D}, \leq \rangle$ be a discrete ordered set [1]. An *interval* over $\mathbb{D}$ is a pair $[a, b]$, with $a, b \in \mathbb{D}$ and $a \leq b$. A *labeling* is a function $\gamma : \mathrm{I}(\mathbb{D}) \to \mathcal{P}(\mathcal{V}ar)$, where $\mathrm{I}(\mathbb{D})$ is the set of all

[1]To keep the notation light, we will identify the order $\langle \mathbb{D}, \leq \rangle$ with its set $\mathbb{D}$

intervals over $\mathbb{D}$ and $\mathcal{V}ar$ is a finite set of propositional variables. A structure of the form $\mathfrak{M} = \langle \mathrm{I}(\mathbb{D}), \gamma \rangle$ is called a *model*.

We say that an interval $[a, b]$ is a *leaf* iff it has no subintervals (i.e., $a = b$).

The truth values of formulae are determined by the following (natural) semantic rules:

1. For all $v \in \mathcal{V}ar$, we have $\mathfrak{M}, [a, b] \models v$ iff $v \in \gamma([a, b])$.
2. $\mathfrak{M}, [a, b] \models \neg\varphi$ iff $\mathfrak{M}, [a, b] \not\models \varphi$.
3. $\mathfrak{M}, [a, b] \models \varphi_1 \wedge \varphi_2$ iff $\mathfrak{M}, [a, b] \models \varphi_1$ and $\mathfrak{M}, [a, b] \models \varphi_2$.
4. $\mathfrak{M}, [a, b] \models \langle D \rangle\varphi$ iff there exists an interval $[a', b']$ such that $\mathfrak{M}, [a', b'] \models \varphi$, $a \leq a'$, $b' \leq b$, and $[a, b] \neq [a', b']$.

Boolean connectives $\vee, \Rightarrow, \Leftrightarrow$ are introduced in the standard way. We abbreviate $\neg\langle D \rangle\neg\varphi$ by $[D]\varphi$ and $\varphi \wedge [D]\varphi$ by $[G]\varphi$.

Note that we use the proper subinterval relation D (the prefixes and suffixes are treat as subintervals), but our technique works also in the strict case, where instead of $[a, b] \neq [a', b']$ we assume that $a \neq a'$ and $b \neq b'$ — see Section 12.3. On the other hand, if we remove the condition $[a, b] \neq [a', b']$, then the problem is known to be decidable[30].

A formula φ is said to be *satisfiable* in a class of orderings $\mathcal{D}$ if there exist a structure $\mathbb{D} \in \mathcal{D}$, a labeling γ, and an interval $[a, b]$, called *the initial interval*, such that $\langle \mathrm{I}(\mathbb{D}), \gamma \rangle, [a, b] \models \varphi$. A formula is satisfiable in a given ordering $\mathbb{D}$ if it is satisfiable in $\{\mathbb{D}\}$.

Useful formulae. We will often use the formulae of the form λ_i that are satisfied in the intervals with the specific length. We define those formulae as $\lambda_i = \langle D^i \rangle\top \wedge \neg\langle D^{i+1} \rangle\top$, where the exponentiation D^k is defined as usual:
$\langle D^0 \rangle\phi = \phi$,
$\langle D^k \rangle\phi = \langle D \rangle\langle D^{k-1} \rangle\phi$ for all $k > 0$.
It is readily checked that λ_k is as required — for example, $\lambda_0 = \top \wedge [D]\bot$ — it is satisfied in the intervals with no subintervals, i.e. the intervals with the length 0.

10. Proof of Theorem 9.1

In Section 10 we only consider finite orderings.
Our representation. We imagine the Kripke structure of intervals of a finite ordering as a directed acyclic graph, where intervals are vertices and each interval $[a, b]$ of length greater than 0 has two successors: $[a + 1, b]$ and $[a, b - 1]$. Each level of this representation contains intervals of the same length (see Fig. 10.1). As usual, each vertex is associated with a subset of propositional variables.

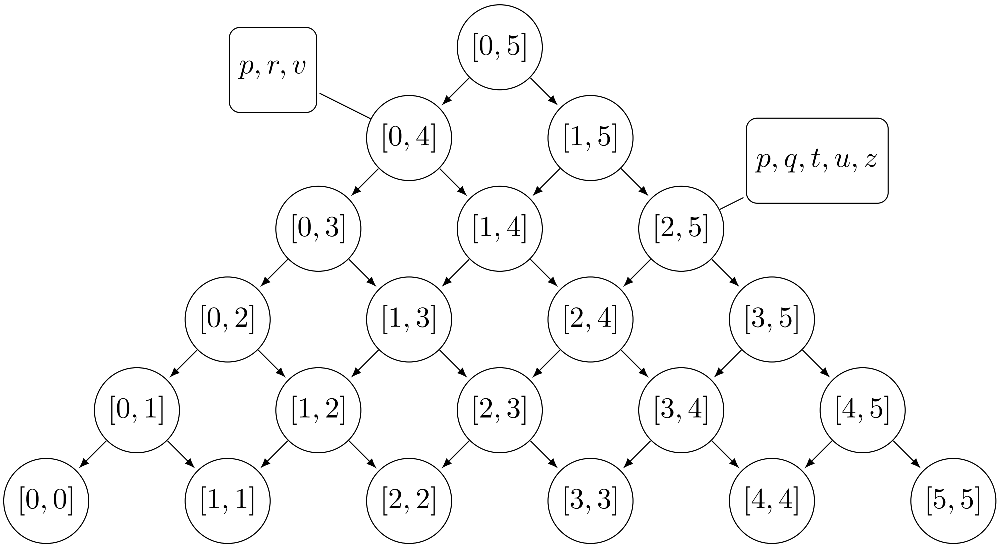

Figure 10.1.: Our representation of order $\langle\{0, 1, \ldots, 5\}, \leq\rangle$.

10.1. The Regular Language L_A

In this section, for a given two-counter finite automaton (Minsky machine) A we will define a regular language L_A. There is nothing about the logic of subintervals in this section – we are just preparing an undecidable problem which will be handy to encode.

Let Q be the set of states of A, and let $Q' = \{q' : q \in Q\}$. Define $B = \{f, f_r, s, s_r\}$ and $B' = \{b' : b \in B\}$

The alphabet Σ of L_A will consist of all the elements of $Q \cup Q'$ (jointly called *states*), symbols x and x' (jointly called X-symbols) and of all the subsets (possibly empty) of B and of B'. Talking about the subsets of B and B', we will not respect types, saying for example "f_r occurs in the word v" rather than "there is a symbol in v that contains f_r".

Symbols (of Σ containing) f and f' (s and s') will be called *first* (resp., *second*) *counters*. Symbols f_r and f'_r (s_r and s'_r) will be called *first* (resp., *second*) *shadows* (or *shadows of the first/the second counter*).

The language L_A consists of the words w over Σ that satisfy the following seven conditions:

C1. The first symbol of w is the initial state q_0 of A and the last symbol of w is either q or q', where q is one of the final states of A.

By a *configuration*, we will mean a maximal sub-word[1] of w, whose first element is a state (called *the state of the configuration*) and which contains exactly one state (so that w is split into disjoint configurations). A configuration will be called *even* if its state is from Q and *odd* if it is from Q'.

C2. Odd and even configurations alternate in w. All the non-state symbols occurring in even configurations are subsets of B and all the non-state symbols occurring in odd configurations are subsets of B'.

C3. Each configuration, except for the last one (which only consists of a state) contains exactly one first counter and exactly one second counter.

We want a word from L_A to encode a sequence of configurations of A which, once an additional distance constraint is satisfied (see Lemma 10.2), will be a correct accepting computation of A. So, except for a state of A, in each configuration, we need to remember the values of the two counters. We define the *value of the first counter* of a configuration as the number of symbols (strictly) between the state of the configuration and its first counter. The same applies to the second counter.

Example. A configuration with the state q, the first counter set to 3, and the second counter set to 4 can be stored as a word $q\emptyset\emptyset\emptyset\{f, f_r, s_r\}\{s\}\emptyset\emptyset x$ (the meaning of f_r, s_r, and x will be defined later).

Using this language, we can state:

C4. In the first configuration, the value of both the counters is zero.

Which can also be read as: The second symbol of w contains f and s.

Now the role of shadows is going to be revealed:

C5. There are no shadows in the first and the last configuration. Each configuration, except for the first and the last, contains exactly one first shadow and exactly one second shadow.

[1] By a sub-word, we mean a sequence of consecutive elements of a word, an infix or a prefix or a suffix.

In reading the next condition, it is good to have in mind that the position of a shadow in a given configuration, relative to the state of the configuration, will be enforced, by the distance constraints of Lemma 10.2, to be the same as the position of the corresponding counter in the previous configuration.

Since the format of an instruction of A is:

```
If the state is $q$
 and the first counter equals/does not equal 0
 and the second counter equals/does not equal 0
then change the state to $q_1$
 and decrease/increase/keep unchanged the first counter
 and decrease/increase/keep unchanged the second counter.
```

it is clear what we mean, saying that *configuration C in word w matches the assumption of the instruction I.*

C6. If C and C_1 are consecutive configurations in w, and C matches the assumption of an instruction I, then:

- If I changes the state into q_1, then the state of C_1 is q_1.
- If I orders the first (second) counter to remain unchanged, then the first (resp., second) counter in C_1 coincides with the first (resp., second) shadow in C_1.
- If I orders the first (second) counter to be decreased, then the first (resp., second) counter in C_1 is the immediate predecessor of the first (resp., second) shadow in C_1.
- If I orders the first (second) counter to be increased, then the first (resp., second) counter in C_1 is the immediate successor of the first (resp., second) shadow in C_1.

One remaining condition is the following:

C7. There is exactly one x in each even configuration. All the counters and shadows of the same configuration are to the left of x. Each x is followed by a state symbol. The same holds for odd configurations and x'.

This completes the definition of the language L_A. It is clear that it is regular – each of the seven conditions above can be checked by a very small finite automaton. Before we formulate Lemma 10.2, which will be our main tool, we need one more definition:

Definition 10.1. *Let $w \in L_A$ and let cvd be a sub-word of w, (where $c, d \in \Sigma$). We will call cvd* an interesting infix *if there is exactly one X symbol in v and one of the following conditions holds:*

1. *c and d are states;*
2. *c is a first counter and d is a shadow of the first counter;*
3. *c is a second counter and d is a shadow of the second counter.*

Notice that the condition that there is exactly one X symbol in v is a way of saying that positions of c and d belong to two consecutive configurations.

Lemma 10.2. *The following two conditions are equivalent:*

(i) Two-counter automaton A, starting from the initial state q_0 and empty counters, accepts.

(ii) There exists a word $w \in L_A$ and a natural number n such that the length of all the interesting infixes of w is n.

Proof. For the $\Rightarrow$ direction consider an accepting computation of A and take n as any number greater than all the numbers that appear on the two counters of A during this computation plus 3 (this is for X-symbols, states and the counters). For the $\Leftarrow$ direction, notice that the distance constraint from (ii) implies that the distance between a state and the subsequent first (second) shadow equals the value of the first (resp., second) counter in the previous configuration. Together with condition 5, defining L_A, this implies that the subsequent configurations in $w \in L_A$ can indeed be seen as subsequent configurations in the valid computation of A.

□

Since the halting problem for two-counter automata is undecidable, the proof of Theorem 9.1 will be completed once we write, for a given automaton A, a formula Ψ of the language of the logic of subintervals which is satisfiable (in a finite model) if and only if condition (ii) from Lemma 10.2 holds. Actually, what the formula Ψ is going to say is, more or less, that the word written (with the use of the labeling function γ) in the leaves of the model is a word w as described in Lemma 10.2, condition (ii).

In the following subsections, we are going to write formulae Φ_{orient}, Φ_{L_A}, Φ_{cloud}, and Φ_{length} such that $\Phi_{\text{orient}} \wedge \Phi_{L_A} \wedge \Phi_{\text{cloud}} \wedge \Phi_{\text{length}}$ will be the formula Ψ we want.

10.2. Orientation

As we said, we want to write a formula saying that the word written in the leaves of the model is the w described in Lemma 10.2, condition (ii).

The first problem we need to overcome is the symmetry of D – the operator does not see a difference between past and future, or between left and right, so how can we distinguish between the beginning of w and its end? We deal with this problem by introducing five variables L, R, s_0, s_1, s_2 and writing a formula Φ_{orient} which will be satisfied by an interval $[a, b]$ if $[a, a]$ is the only subinterval of $[a, b]$ that satisfies L and

$[b, b]$ is the only subinterval of $[a, b]$ that satisfies R, or $[b, b]$ is the only subinterval of $[a, b]$ that satisfies L and $[a, a]$ is the only subinterval of $[a, b]$ that satisfies R, and all the following conditions hold:

- any interval that satisfies L or R satisfies also one of s_0, s_1, or s_2;
- each leaf is labeled either with s_0 or with s_1 or with s_2;
- each interval labeled with s_0 or with s_1 or with s_2 is a leaf;
- if c, d, e are three consecutive leaves of $[a, b]$ and if s_i holds in c, s_j holds in d and s_k holds in e then $\{i, j, k\} = \{0, 1, 2\}$;
- the initial interval has the length at least 3.

If $[a, b] \models \Phi_{\text{orient}}$, then the leaf of $[a, b]$ where L holds (resp., where R holds) will be called the left (resp., the right) end of $[a, b]$.

Let $exactly\_one\_of(Y) = \bigvee_{y \in Y}(y \wedge \bigwedge_{y' \in Y \setminus \{y\}} \neg y')$ be a formula saying (which is not hard to guess) that exactly one variable from the set Y is true in the current interval. Φ_{orient} is the conjunction of the following formulae:

(i) $\langle D \rangle \langle D \rangle \langle D \rangle \top$

(ii) $[G]((\lambda_0 \Rightarrow exactly\_one\_of(\{s_0, s_1, s_2\})) \wedge (s_0 \vee s_1 \vee s_2 \Rightarrow \lambda_0))$

(iii) $[G](\lambda_2 \Rightarrow \langle D \rangle s_0 \wedge \langle D \rangle s_1 \wedge \langle D \rangle s_2)$

(iv) $[G](L \vee R \Rightarrow s_0 \vee s_1 \vee s_2)$

(v) $\langle D \rangle R \wedge \langle D \rangle L$

(vi) $[G](L \Rightarrow \neg R)$

(vii) $\bigvee_{i \in \{0,1,2\}}(\langle D \rangle (L \wedge s_i) \wedge [D](\lambda_1 \wedge \langle D \rangle L \Rightarrow \neg \langle D \rangle s_{(i-1) \bmod 3}))$

(viii) $\bigvee_{i \in \{0,1,2\}}(\langle D \rangle (R \wedge s_i) \wedge [D](\lambda_1 \wedge \langle D \rangle R \Rightarrow \neg \langle D \rangle s_{(i+1) \bmod 3}))$

Formulae (i), (ii), (iii), and (iv) express the property defined by the conjunction of the five items above (notice, that λ_0 means that the current interval is a leaf).

Formula (v) says that there exists an interval labeled with R and an interval labeled with L.

Formula (vi) states that no interval satisfies both L and R.

Formula (vii) guarantees that no interval containing exactly 2 leaves, which is a superinterval of an interval labeled with L and s_i, can contain a subinterval labeled with $s_{(i-1) \bmod 3}$. It implies that an interval labeled with L can only have one superinterval containing exactly 2 leaves — if there were two, then their common superinterval containing 3 leaves would not have a subinterval labeled with $s_{(i-1) \texttt{ mod } 3}$, thus contradicting (iii).

Finally, formula (viii) works like (vii) but for R.

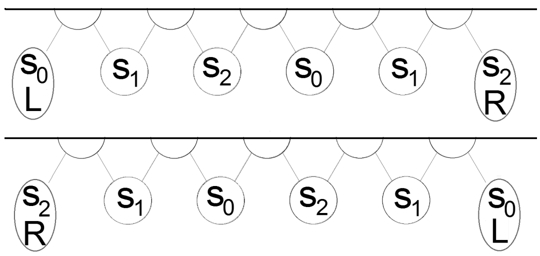

Figure 10.2.: Two possible models that satisfy the formulae from Section 10.2.

In the rest of paper, we restrict our attention to models satisfying formula Φ_{orient}, and treat the leaf labeled with L as the leftmost element of the model.

Notice that everything we did above can be applied not only to the whole model, but also to any subinterval of the model. We will say that set U *marks the left endpoint of interval* $[c, d]$ if some $u \in U$ holds in $[c, c]$ and no $u' \in U$ holds in any other subinterval of $[c, d]$. Analogously we define what it means that a set *marks the right endpoint of an interval.* What we proved in this section is:

Lemma 10.3. *There exists a formula $mle(U)$ (and $mre(U)$) which is true in interval $[a, b]$ if and only if U marks the left (resp. the right) end of $[a, b]$.*

Notice that we only know how to express the fact that $u \in U$ is valid in the left end of $[a, b]$ if u does not occur anywhere else in this interval.

10.3. Encoding a Finite Automaton

In this section, we show how to make sure that consecutive leaves of the model, read from L to R, are labeled with variables that represent a word of a given regular language.

Lemma 10.4. *Let $\mathcal{A} = \langle \Sigma, \mathcal{Q}, q^0, \mathcal{F}, \delta \rangle$, where $q^0 \in \mathcal{Q}$, $\mathcal{F} \subseteq \mathcal{Q}$, $\delta \subseteq \mathcal{Q} \times \Sigma \times \mathcal{Q}$ be a finite–state automaton. There exists a formula $\psi_{\mathcal{A}}$ of the D fragment of Halpern–Shoham logic over alphabet $\mathcal{Q} \cup \Sigma$ that is satisfiable (with respect to the valuation of the variables from $\mathcal{Q}$) if and only if the word, over the alphabet Σ written in the leaves of the model, read from L to R, belongs to the language accepted by $\mathcal{A}$.*

Proof. It is enough to write a conjunction of the following properties.

1. In every leaf, exactly one letter from Σ is satisfied (so there is indeed a word, written in the leaves). Moreover, the letters from Σ are true at leaves only.

2. Each leaf is labeled with exactly one variable from $\mathcal{Q}$. Moreover, the variable from $\mathcal{Q}$ are true at leaves only.

3. For each interval whose length is 1, if this interval contains an interval labeled with s_i, with $a \in \Sigma$, and with $q \in \mathcal{Q}$, and another interval labeled with $s_{(i+1) \bmod 3}$ and with $q' \in \mathcal{Q}$, then $\langle q, a, q' \rangle \in \delta$ (notice that we rely here on the assumption that Φ_{orient} holds in the model).

4. The interval labeled with R is labeled with such $q \in \mathcal{Q}$ and $a \in \Sigma$ such that $\langle q, a, q' \rangle \in \delta$ for some $q' \in \mathcal{F}$.

5. The interval labeled with L is labeled with q^0.

Clearly, a model satisfies properties 1-5 if and only if its leaves are labeled with an accepting run of $\mathcal{A}$ on the word over Σ written in its leaves. The formulae of the D fragment of Halpern–Shoham logic expressing properties 1-5 are not hard to write:

1. $[G]((\lambda_0 \Rightarrow \mathit{exactly\_one\_of}(\Sigma)) \wedge (\bigvee \Sigma \Rightarrow \lambda_0))$

2. $[G]((\lambda_0 \Rightarrow \mathit{exactly\_one\_of}(\mathcal{Q})) \wedge (\bigvee \mathcal{Q} \Rightarrow \lambda_0))$

3. $[G](\lambda_1 \wedge \langle D \rangle s_i \wedge \langle D \rangle s_{i+1 \bmod 3} \Rightarrow \bigvee_{\langle q,a,q' \rangle \in \delta} \langle D \rangle (s_i \wedge q \wedge a) \wedge \langle D \rangle (s_{i+1 \bmod 3} \wedge q'))$, for each $i \in \{0, 1, 2\}$

4. $[G](R \Rightarrow \bigvee_{\langle q,a,q' \rangle \in \delta, q' \in \mathcal{F}} (q \wedge a))$

5. $[G](L \Rightarrow q^0)$

Now, let $\mathcal{A}$ be a finite automaton recognizing language L_A from Section 10.1. We put $\Phi_{L_A} = \psi_{\mathcal{A}}$. $\square$

10.4. A Cloud

We still need to make sure that there exists n such that each configuration (but the last one) has length $n-1$ and that each interesting infix has length exactly n. Let us start with:

Definition 10.5. *Let $\mathfrak{M} = \langle \mathrm{I}(\mathbb{D}), \gamma \rangle$ be a model and p a propositional variable. We call p* a cloud *if there exists $k \in \mathbb{N}$ such that $p \in \gamma([a,b])$ if and only if the length of $[a,b]$ is exactly k.*

So one can view a cloud as a set of all intervals of some fixed length. Notice, that if the current interval has length k then exactly $k+1$ leaves are reachable from this segment with the operator D.

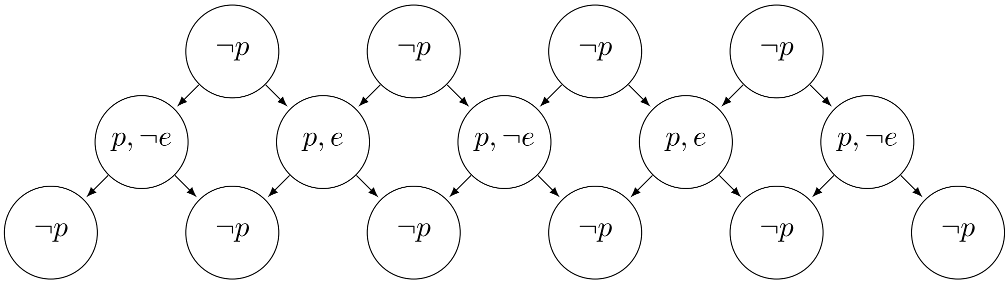

Figure 10.3.: An example of a cloud.

We want to write a formula in the language of the D fragment of Halpern-Shoham logic saying that p is a cloud. In order to do that, we use an additional variable e. The idea is that an interval $[a, a+n]$ satisfies e iff $[a+1, a+n+1]$ does not.

Let Φ_{cloud} be the conjunction of the following formulae.

1. $\langle D\rangle(p \wedge \langle D\rangle L)$ — there exists an interval that satisfies p and this interval contains the leftmost element of the model.

2. $[G](p \Rightarrow [D]\neg p)$ — intervals labeled with p cannot contain intervals labeled with p.

3. $[G](\langle D\rangle p \Rightarrow \langle D\rangle(p \wedge e) \wedge \langle D\rangle(p \wedge \neg e))$ — each interval that contains an interval labeled with p actually contains at least two such intervals — one labeled with e and one with $\neg e$.

Lemma 10.6. *If $\mathfrak{M}, [a_{\mathfrak{M}}, b_{\mathfrak{M}}] \models \Phi_{\text{cloud}}$, where $a_{\mathfrak{M}}$ and $b_{\mathfrak{M}}$ are endpoints of $\mathfrak{M}$, then p is a cloud.*

Proof. We will prove that if an interval $[x, y]$ is labeled with p, then also $[x+1, y+1]$ is labeled with p. A symmetric proof shows that the same holds for $[x-1, y-1]$, so all the intervals of length equal to m, where m is the length of $[x, y]$, are labeled with p.

This will imply that no other intervals can be labeled with p and p is indeed a cloud. This is because each such interval either has a length greater than m, and thus contains an interval of length m, and as such labeled with p, or has a length smaller than m, and is contained in an interval labeled by p, in both cases contradicting 2.

Consider an interval $[x, y]$ labeled with p. Interval $[x, y+1]$ contains an interval labeled with p, so it has to contain two different intervals labeled with p – one labeled with e and the other one with $\neg e$. Suppose, without loss of generality, that $[x, y]$ is the one labeled with e, and let us call the second one $[u, t]$. If $t < y+1$, then $[u, t]$ is a subinterval of $[x, y]$ and is labeled with p, a contradiction. So $t = y+1$.

Let us assume that $u > x+1$. The interval $[u-1, y+1]$ must contain two different intervals labeled with p. One of them is $[u, y+1]$, and it cannot contain another interval labeled with p, so the other one must be $[u-1, y]$ or one of its subintervals.

But then it is a subinterval of $[x, y]$ (because $u - 1 > x + 1 - 1 = x$) which also is labeled with p, but this leads to a contradiction. So $u = x + 1$. $\square$

10.5. Using a cloud.

10.5.1. An example

Before we proceed into the technical aspects of encoding the automata, let us show a simple example of the usage of a cloud.

Let $\mathcal{A}_{'}$ be a finite automaton that recognizes the language defined by the regular expression $(ac^*bc^*)^*$. Consider the formula $\Phi_{\text{orient}} \wedge \Phi_{L_{A_0}}$. In each model of ρ, leaves contain some word w accepted by $\mathcal{A}$. Our goal is to force the following property.

(e) There exists $n \in \mathbb{N}$ such that each maximal block c that is between a and b has the length n.

We use a cloud. In fact, n will be equal to the length of the intervals in the cloud. Property (e) can be expressed as a conjunction of Φ_{cloud} and the following formulae.

(i) $[G](p \Rightarrow \neg(\langle D \rangle a \wedge \langle D \rangle b))$

(ii) $[G](p \Rightarrow \langle D \rangle (a \vee b))$

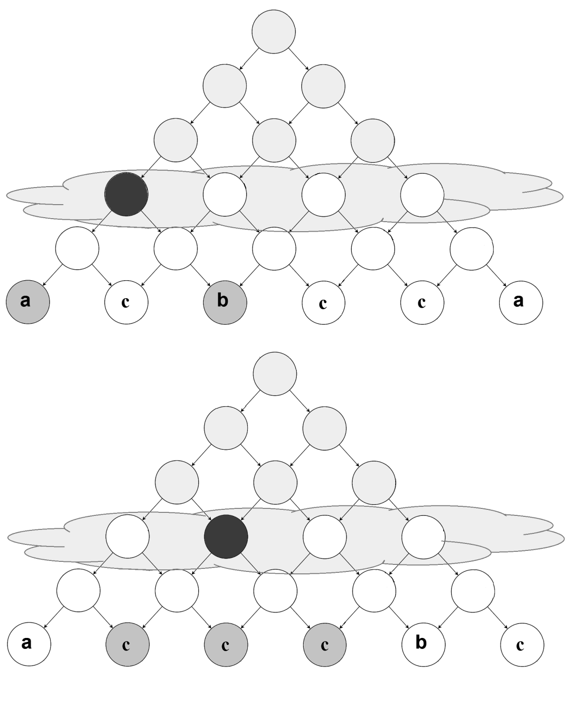

Figure 10.4.: An example of using a cloud.

Now, consider two examples in Figure 10.4. At the left side, the red vertex contains no subinterval that satisfies a or b — it contradicts Formula (ii). At the right side, the red vertex contain a subinterval that satisfies a and a subinterval that satisfies b — it contradicts Formula (i). Therefore, this block of c has to have the length 2, equal to the length of the intervals in the cloud.

10.5.2. Encoding two-counter automaton

Let us now concentrate on models which satisfy $\Phi_{\text{orient}} \wedge \Phi_{L_A} \wedge \Phi_{\text{cloud}}$. Since Φ_{cloud} is satisfied, then p is a cloud. Let n denote the number of leaves contained in the intervals that form the cloud. Since Φ_{L_A} is satisfied, we know that the word written in the leaves of the model must belong to L_A. What remains to be done is writing a formula Φ_{length} that would guarantee that the distance constraints from Lemma 10.2 are satisfied in this word.

The following lemma is just a restatement of Definition 10.1 in the language of the last paragraph of Section 10.1:

Lemma 10.7. *Let $w \in L_A$ and let v be a sub-word of w. Then v is an interesting infix if it contains exactly one X-symbol and one of the following conditions holds:*

- *one of the endpoints of v is marked with a state from Q and the other endpoint is marked with a state from Q';*
- *the left endpoint of v is marked with f (f', s, s') and the right endpoint is marked with f'_r (f_r, s'_r, s_r, resp.).*

Using the formulae mle and mre from Section 10.2, it is straightforward to translate the conditions of the lemma into a formula $Phi_{\text{interesting}}$ saying that the current interval is interesting:

$$
\begin{aligned}
& [G]((mle(Q) \wedge mre(Q') \vee (mle(Q') \wedge mre(Q)) \Rightarrow interesting) \\
\wedge\ & [G](mle(\{l \in \Sigma | f \in l\}) \wedge mre(\{l \in \Sigma | f'_r \in l\}) \Rightarrow interesting) \\
\wedge\ & [G](mle(\{l \in \Sigma | f' \in l\}) \wedge mre(\{l \in \Sigma | f_r \in l\}) \Rightarrow interesting) \\
\wedge\ & [G](mle(\{l \in \Sigma | s \in l\}) \wedge mre(\{l \in \Sigma | s'_r \in l\})) \Rightarrow interesting) \\
\wedge\ & [G](mle(\{l \in \Sigma | s' \in l\}) \wedge mre(\{l \in \Sigma | s_r \in l\}) \Rightarrow interesting)
\end{aligned}
$$

Note that the part about containing exactly one X-symbol comes for free here from the definition of the language and the properties of mle and mre. Now, we are ready to write Φ_{length}.

$$\Phi_{\text{length}} = \Phi_{\text{interesting}} \wedge [G](interesting \Rightarrow p)$$

which means that if what you see is exactly an interesting interval, then you are exactly on the level of the cloud.

This ends the proof of Theorem 9.1.

11. Proof of Theorem 9.2

The idea of the proof of Theorem 9.2 is exactly the same as of Theorem 9.1. But, because of the possible pathologies of discrete orders, almost all the details of the proof will now be much more complicated.

11.1. Damage assessment

Let us see which of the constructions form Section 10 can be saved in the new context.

Orientation In the new situation we still can, as we did in Section 10.2, write formulae enforcing that the model has its left endpoint, marked with L, and its right endpoint, marked with R. But the trick with labeling each three consecutive elements with s_0, s_1 and s_2, which we used to define direction inside the model will, in the discrete case, orient only the locally finite fragments of the model (i.e., those maximal sets C of elements of the ordering such that, for each $a, b \in C$ the interval $[a, b]$ contains only finitely many elements).

On the other hand, if the model is infinite, then the left endpoint has its successor, which has a successor, etc. so that we have a copy of the ordered set of natural number as an initial fragment of the model. We will identify elements of this fragment with natural numbers. If formula Φ_{orient} is satisfied, then the set of natural numbers is oriented as in Section 10.2.

It also turns out that we can actually force the model to be infinite. To do that, take a new variable nat, and write a formula Φ_{nat} saying that:

- nat only holds at leaves;
- L implies nat;
- if an interval contains two leaves, and in some of those two leaves nat holds, then it holds in all of them;
- there is an leaf where nat does not hold.

Those properties can by expressed using D in the following way:

- $[G](nat \Rightarrow \lambda_0)$
- $[G](L \Rightarrow nat)$;
- $[G](\lambda_1 \wedge \langle D\rangle nat \Rightarrow [D]nat)$;

- $\langle D \rangle(\lambda_0 \wedge \neg nat)$

Let now Φ^d_{orient} be the formula $\Phi_{\mathrm{orient}} \wedge \Phi_{nat}$. From now on, we assume that all the models under consideration satisfy Φ^d_{orient}.

The regular language L^d_A and the finite automaton. In the finite satisfiability case, the set of satisfiable formulae was recursively enumerable. Now, as we will see later, it is coRE-hard. This means that we now need, for a given Minsky machine A, to write a formula Ψ^d, of the logic of subintervals, which will be satisfiable if and only if A **does not** accept. We can assume that A has only one final (accepting) state q_f and that the machine runs forever if this state is not reached. So the formula we are going to write in this chapter should be satisfiable if and only if the machine A runs forever and never reaches q_f.

Since we still want to represent the computation of A as a word written in atoms of the model (to be more precise, in the atoms that are natural numbers), we must be ready to deal with an infinite word. The method the transition function of an automaton is encoded in Section 10.3 still works, so we can encode any automaton on infinite words with a "safety accepting condition", which means that it accepts a word if no forbidden state is entered during a run. Let L^d_A be the language of infinite words satisfying conditions [C1] – [C7] from Section 10.1 (with the obvious exception of the parts of conditions [C1], [C3] and [C5] which concern the final configuration) and additionally

C8. The third symbol of w is its first X-symbol.

Clearly, L^d_A can be recognized by an automaton with safety accepting condition. So we can write a formula $\Phi_{L^d_A}$ which will be satisfied in a model if and only if the word written in atoms being natural numbers belongs to L^d_A. Notice that $\Phi_{L^d_A}$ will be satisfied also by some words which only consist of finitely many configurations (last of them ending with infinitely many empty symbols). This cannot be prevented by a safety automaton, and we will need to find another way to forbid such words.

Lemma 10.2. The last remark leads to one change in Lemma 10.2. Another change will result from the fact that, in the new context we do not have the cloud anymore – the method it was defined does not translate to discrete orderings. So we no longer will be able to make sure that all the interesting infixes have the same length. But it turns out that we do not really need that much.

Definition 11.1. *We say that an infinite word w is* nice *if for each pair v, u of interesting infixes such that v begins earlier than u, if k is the number of X-symbols between the left endpoint of v and the left endpoint of u then $|v| + k = |u|$.*

The following version of Lemma 10.2 is easy to prove:

Lemma 11.2. *The following two conditions are equivalent:*

(i) Two-counter automaton A, started from the initial state q_0 and empty counters, runs forever.

(ii) There exists a nice word $w \in L_A^d$ with infinitely many X-symbols in w.

Notice that if there are two interesting infixes of a nice word, whose left ends are in the same configuration, then their lengths are equal. This is exactly what we need to be sure that the values of counters in each configuration are correctly reflected by the positions of shadows of the counters in the following configuration.

The second consequence of the fact that a word is nice is that the length of a subsequent configuration is always one plus the length of the previous one. This means that, if the first configuration was long enough to contain the values of the counters, then each configuration will be long enough, regardless of the possible unbounded growth of those values. And it follows from condition [C8] that the first configuration is long enough.

Having the idea on mind, proving Lemma 11.2 is straightforward.

11.2. The parabola

Let us remind that we identify the initial fragment of the model with the set $\mathbb{N}$.

Definition 11.3. *Let $\mathfrak{M} = \langle \mathrm{I}(\mathbb{D}), \gamma \rangle$ be a model and p, p_E, x, x' be a quadruple of variables. We call the quadruple p, p_E, x, x'* the parabola *if:*

(i) only leaves are labeled with x or with x', and the leaf $[3, 3]$ is labeled with x;

(ii) $[1, 4]$ is labeled with p;

(iii) if $[i, j]$ is labeled with p and $[i, i]$ is not labeled with x or with x', then $[i+1, j+1]$ is labeled with p;

(iv) each interval labeled with p contain a subinterval labeled with x or with x', but no such interval marks the right endpoint with x or with x';

(v) all intervals labeled with p_E mark the right endpoint with x or with x';

(vi) no subinterval of an interval labeled with p can contain a subinterval with x and a subinterval with x';

(vii) if $[i, j]$ is labeled with p and x (resp., x') marks the left endpoint of $[i, j]$, then $[i + 1, j + 1]$ is labeled by p_E (see Figure 11.2);

(viii) if $[i, j]$ is labeled with p_E, then $[i, j + 1]$ is labeled with p;

(ix) no other interval whose left endpoint is a natural number are labeled with p.

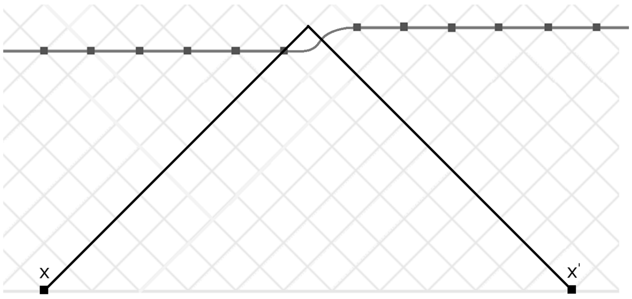

Figure 11.1.: A fragment of the parabola.

Notice that the x, x' from the parabola coincide with the X-symbols from Σ, which means that if p, p_E, x, x' is the parabola then there are infinitely many X-symbols in w, and, in consequence, w consists of infinitely many configurations — a property that could not be enforced by a safety automaton alone.

The variable p_E is auxiliary in some sense - we use it to mark in each row that contains p, the first column after the columns with p. In other words, an interval $[i, j]$ will be labeled with p_E if $[i-1, j-1]$ is labeled with p and $[i-1, i-1]$ is labeled with x or x' (so, therefore, $[i, j+1]$ should be labeled with p) — see Figure 11.2.

The role of variable p is as the role of cloud — we will use it to guarantee that the interesting infixes have the right length.

Lemma 11.4. *Let $\mathfrak{M} = \langle \mathrm{I}(\mathbb{D}), \gamma \rangle$ be a model and p, p_E, x, x' be the parabola in $\mathfrak{M}$. Let $w \in L_A^d$ be the infinite word of symbols written in the natural numbers of $\mathfrak{M}$. Then the following two conditions are equivalent:*

(i) w is a nice word with infinitely many X-symbols;

(ii) each interesting infix of w is (as an interval) labeled with p.

Proof. Condition [C8] implies that the fourth symbol of w is a state symbol, and consequently, that $[1, 4]$ is the first interesting infix of w, and condition (ii) from the definition of the parabola implies that it is labeled by p.

Notice that conditions from the definition of parabola guarantee that if $[i, j]$ is labeled with p, then there exists an interval labeled with p that begins in $i + 1$ and has the length greater by one than $[i, j]$ if there is an X-symbol in $[i, i]$ (conditions (vi)-(viii)) and has the same length as $[i, j]$ otherwise (condition (iii)). It implies that the length of intervals labeled with p is increased by one with each X-symbol, so the length of two intervals labeled with p whose left ends are separated by k X-symbols differs by k — exactly as in the definition of a nice word. The condition (iv) guarantee that there are infinitely many X-symbols. $\square$

The idea behind the formula Φ^d_{par} is simple. At first, note that we still can use formulae *mle* and *mre* defined exactly as in the definition of a cloud to define the parabola.

Let us Φ^d_{par} be the conjunction of the following formulae (they correspond to the properties from Definition 11.3).

(i) $[G](x \vee x' \Rightarrow \lambda_0) \wedge \langle D\rangle(\lambda_3 \wedge \langle D\rangle L \wedge mre(\{x\}))$;

(ii) $\langle D\rangle(p \wedge \langle D\rangle L \wedge \lambda_3)$ (recall that by definition L is satisfied in [1, 1]);

(iii) $[G]((\langle D\rangle p \wedge \neg mle(\{x\}) \wedge \neg mle(\{x'\})) \Rightarrow \langle D\rangle(p \wedge e) \wedge \langle D\rangle(p \wedge \neg e))$ (we use an auxiliary variable e in the same way as for cloud);

(iv) $[G]((mre(\{x\}) \vee mre(\{x'\})) \Rightarrow \neg p) \wedge [G](p \Rightarrow \langle D\rangle(x \vee x'))$;

(v) $[G](p_E \Rightarrow mre(\{x\}) \vee mre(\{x'\}))$;

(vi) $[G](p \Rightarrow [D]([D]\neg x \vee [D]\neg x'))$;

(vii) $[G]((mle(\{y\}) \wedge \langle D\rangle(p \wedge mle\{y\}) \wedge [D][D]\neg p) \Rightarrow \langle D\rangle p_E \wedge [D][D]\neg p_E)$ for each $y \in \{x, x'\}$;

(viii) $[G](\langle D\rangle p_E \wedge [D][D]\neg p_E \wedge \neg(mle(\{x\}) \vee mre(\{x'\})) \Rightarrow p)$;

(ix) $[G](p \Rightarrow [D]\neg p)$ (it works as for cloud).

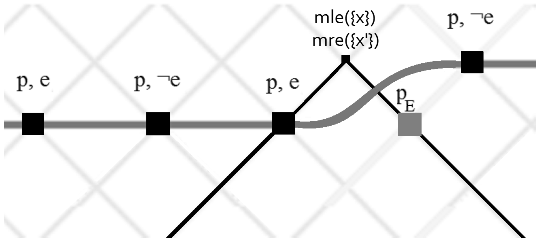

Figure 11.2.: A place where the parabola from Figure 11.1 is lifted, zoomed in.

It turns out that we can simply put $\Phi^d_{\text{length}} = \Phi_{\text{length}}$. Now we can finish the proof of Theorem 9.2 defining

$$\Psi^d = \Phi^d_{\text{orient}} \wedge \Phi_{L^d_A} \wedge \Phi^d_{\text{par}} \wedge \Phi^d_{\text{length}}$$

12. More results

12.1. Superinterval relation

Our theorems hold also for $\bar{D}$ instead of D. But the changes to the proof need to be significant. Consider, for example, formulae $[D]\langle D\rangle\top$ and $[\bar{D}]\langle\bar{D}\rangle\top$. The first one is not satisfied in the discrete case, but the second one is satisfied over, e.g., $\mathbb{N}$.

Another important difference between the models for the logic with D and for the logic with $\bar{D}$ is that in linear discrete orders all intervals have either 0 or 2 subintervals with the maximal size, while an interval may have only one superinterval with the minimal size (e.g. over the naturals, the interval $[1,5]$ has only one such superinterval — $[1,6]$).

In other words, the leaves are not always well-defined in case of $\bar{D}$. There can be no interval that satisfies $[\bar{D}]\top$ in a model, and even if there is such interval, then it is the only one. Therefore we have to pay more attention while encoding a regular language.

To handle it, we use a cloud to define *pseudo-leaves* — a set of intervals on the same level that satisfies a special variable $leaf$. Once we defined leaves, we guarantee that nothing wrong happens above the leaves, i.e. $[\bar{D}](leaf \Rightarrow [\bar{D}]\bigwedge_{v\in\mathcal{V}ar}\neg v)$. The idea is that having pseudo-leaves, we can simply adopt the formula from the proof of Theorem 2 to prove undecidability, replacing D by $\bar{D}$ and λ_i by λ'_i, where $\lambda'_i = \langle D^i\rangle leaf \wedge \neg\langle D^{i+1}\rangle leaf$.

As the *successor* (*predecessor*) function is well defined in discrete linear orders for all points except for the maximal (minimal, resp.) one, it should be clear what we mean by $c+k$ ($c-k$, resp.), where c is a point and k is a natural number — it is just a successor (predecessor, resp.) function iterated k times. Now we are going to write a formula Φ^s_{leaves} that is satisfied only in models with *pseudo-leaves*. More precisely, $\mathfrak{M},[a,b]\models\Phi^s_{\text{leaves}}$ iff there exists a superinterval $[c,d]$ of $[a,b]$ such that:

(i) $\mathfrak{M},[c,d]\models L\wedge leaf\wedge\neg l_e$

(ii) For each even $k\in\mathbb{Z}$ such that $[c+k,d+k]$ is in the model, $\mathfrak{M},[c+k,d+k]\models leaf\wedge\neg l_e$

(iii) For each odd $k\in\mathbb{Z}$ such that $[c+k,d+k]$ is in the model, $\mathfrak{M},[c+k,d+k]\models leaf\wedge l_e$

So all the leaves are at the same level at least in some fragment of the model that contains an interval which labeled with L (so the place where a word from a regular language starts) and is long enough to contain the infinite world.

The technical aspects of Φ^s_{leaves} are not surprising — there are almost the same as for cloud. Note that the last property is a little bit stronger, as explained before (below $[G]\varphi$ stands for $\varphi \wedge [\bar{D}]\varphi$).

- $\langle \bar{D} \rangle (leaf \wedge L)$
- $[G](\langle \bar{D} \rangle leaf \Rightarrow \langle \bar{D} \rangle (leaf \wedge l_e) \wedge \langle \bar{D} \rangle (leaf \wedge \neg l_e))$
- $[G](leaf \Rightarrow [\bar{D}] \bigwedge_{v \in \mathcal{V}ar} \neg v)$

Now, for a given automaton $\mathcal{A}$, we define Φ^s_{orient} ($\Phi^s_{L^d_A}$, Φ^s_{par}, Φ^s_{length}) as Φ^d_{orient} ($\Phi_{L^d_A}$, Φ^d_{par}, Φ^d_{length}, resp.) by replacing every occurrence of D by $\bar{D}$ and every occurrence of λ_i by λ'_i. Let $\Psi^s = \Phi^s_{\text{leaves}} \wedge \Phi^s_{\text{orient}} \wedge \Phi^s_{L^d_A} \wedge \Phi^s_{\text{par}} \wedge \Phi^s_{\text{length}}$. It is easy to check that the formula Ψ^s is satisfiable if and only if $\mathcal{A}$ started from the initial state q_0 and empty counters, runs forever. We conclude this discussion with the following theorem.

Theorem 12.1. *The satisfiability problem for the formulae of the logic of superintervals, over all discrete models, is undecidable.*

12.2. Global satisfiability

The *global satisfiability* problem is defined as follows. For a given formula φ, does there exist a structure $\mathbb{D}$ such that every point of $\mathbb{D}$ satisfies φ? The question about the global satisfiability of formulae has been studied in literature, e.g., in [11] or [17]. For a basic modal logic the global satisfiability problem is ExpTime-complete, while the classical (local) satisfiability problem is PSpace-complete.

We show that the global satisfiability problem is very easy for D, a little bit more complicated for $\bar{D}$ (but still in NP), but becomes undecidable as soon as we allow both D and $\bar{D}$.

We start with the following proposition.

Proposition 12.2. *The global satisfiability problem for the logic of subintervals is* NP*-complete.*

Proof. Let φ be a formula and $\mathfrak{M} = \langle \mathrm{I}(\mathbb{D}), \gamma \rangle$ a model of φ such that for all $w \in \mathrm{I}(\mathbb{D})$ we have $\mathfrak{M}, w \models \varphi$. Let $i \in \mathbb{D}$. We define $\mathfrak{M}' = \langle \mathrm{I}(\{i\}), \gamma' \rangle$ where $\gamma'([i,i]) = \gamma([i,i])$. It is easy to see that $\mathfrak{M}', w \models \varphi$.

Therefore in this case we have a single-interval model property, and checking for the existence such a model can be done in NP. The NP lower bound comes from a trivial reduction from SAT. □

Now we prove that if we have both D and $\bar{D}$, then the global satisfiability problem is much harder.

Proposition 12.3. *The global satisfiability problem for the fragment $D\bar{D}$ of the Halpern–Shoham logic is undecidable.*

Proof. For a given two-counter automaton $\mathcal{A}$, let Ψ^d be a formula defined in section 11. Let u be a fresh variable. Define

$$\Psi^g = (u \Rightarrow \Psi^d) \wedge (\neg u \Rightarrow \langle D\rangle\langle \bar{D}\rangle u \vee \langle \bar{D}\rangle\langle D\rangle u)$$

Clearly, any model $\mathfrak{M}$ that globally satisfies Ψ^g contains an interval w labeled by u and $\mathfrak{M}, w \models \Psi^d$. For the other direction, if we have a model $\mathfrak{M}$ and an interval w such that $\mathfrak{M}, w \models \Psi^d$, then we can create a model $\mathfrak{M}'$ from $\mathfrak{M}$ by labeling w by u and all other intervals by $\neg u$. Then $\mathfrak{M}'$ globally satisfies Ψ^g. □

Both proofs are almost straightforward. Now we prove less obvious proposition about $\bar{D}$.

Proposition 12.4. *The global satisfiability problem for the logic of superintervals is* NP*-complete.*

Proof. Let φ be a formula and $\mathfrak{M} = \langle \mathrm{I}(\mathbb{D}), \gamma\rangle$ a model of φ such that for all $w \in \mathrm{I}(\mathbb{D})$ we have $\mathfrak{M}, w \models \varphi$. If there is a maximal interval in $\mathrm{I}(\mathbb{D})$, then we simply proceeder like for D and obtain a single-point model.

Suppose that there is no maximal interval in $\mathrm{I}(\mathbb{D})$. We define a *modal type* of an interval w, denoted by mt_w, as a set of subformulae of φ of the form $\langle \bar{D}\rangle\varphi'$ satisfied in w. Note that if w is a superinterval of w', then $mt_w \subseteq mt_{w'}$. Therefore there exists an interval $[a_0, b_0]$ such that for all superintervals v' of v we have $mt_v = mt_{v'}$.

We define a set of intervals $\mathcal{T}$ in the following way. For each ψ, which is a subformula of φ, if there exists a superinterval of $[a_0, b_0]$ that satisfies ψ, then we put one such interval in $\mathcal{T}$. Observe that the size of is polynomial in the size of φ.

Let $\{t_0, t_1, \dots, t_{k-1}\}$ be a set $\mathcal{T}$ written more explicit and $f : \mathrm{I}(\mathbb{Z}) \to$ be a function such that for each $w \in \mathrm{I}(\mathbb{Z})$ we have $f(w) = t_{|[a,b]| \mod k}$.

We define a structure $\mathfrak{M}' = \langle \mathrm{I}(\mathbb{Z}), \gamma'\rangle$, where $\gamma'([a, b]) = \gamma(f([a, b]))$. We claim that $\mathfrak{M}'$ is a model of φ.

We prove by the induction of the structure of φ, that for all ψ, which is a subformula of φ, and all w we have $\mathfrak{M}', w \models \psi$ iff $\mathfrak{M}, f(w) \models \psi$.

Let $[a', b']$ be an interval of $\mathfrak{M}'$.

- If ψ is a propositional variable p, then $\mathfrak{M}, f([a', b']) \models \psi$ iff $p \in \gamma(f([a', b'])) = \gamma'([a', b'])$ iff $\mathfrak{M}', [a', b'] \models \psi$.
- The cases when $\psi = \neg\psi'$ or $\psi = \psi' \vee \psi''$ are straightforward.
- If $\psi = \langle D\rangle\psi'$ and $\mathfrak{M}', [a', b'] \models \psi$, then there exists a superinterval $[c', d']$ of $[a', b']$ such that $\mathfrak{M}', [c', d'] \models \psi'$. By the inductive assumption, $\mathfrak{M}, f([c', d']) \models \psi'$. The intervals $f([c', d'])$ is a subinterval of $[a_0, b_0]$, and therefore $mt_{[a_0,b_0]}$ contains ψ. Since $mt_{[a_0,b_0]} = mt_{f([a',b'])}$, we have $\mathfrak{M}, f([a', b']) \models \psi$.

  If $\psi = \langle D\rangle\psi'$ and $\mathfrak{M}, f([a', b']) \models \psi$, then there exists an superinterval $[a, b]$ of $f([a', b'])$ such that $\mathfrak{M}, [a, b] \models \psi'$. ψ' is a subformula o φ, so there is an interval $[c, d] \in \mathcal{T}$ that satisfies ψ'. Let $[c', d']$ be a superinterval of $[a', b']$ such that $f([c', d']) = [c, d]$. By the inductive assumptions, $[c', d']$ satisfies ψ' and therefore $\mathfrak{M}', [a', b'] \models \psi$.

The existence of a model of the form $\langle \mathrm{I}(\mathbb{Z}), \gamma' \rangle$ can be checked in NP. The algorithm simply guesses a modal type of all intervals in the model and a set $\mathcal{T}$, and then simply verify that the modal type is consistent with the types in $\mathcal{T}$.

□

12.3. Strict D

The strict D, denoted as $D_\subset$, is defined as follows.

$\mathfrak{M}, [a, b] \models \langle D_\subset \rangle \varphi$ iff there exist an interval $[a', b']$ such that $\mathfrak{M}, [a', b'] \models \varphi$ and $a < a' \leq b' < b$.

Our undecidability result holds also for $D_\subset$, there are just some minor technical details to handle. Here we will only describe how we label the leaves with a special variable l in that case — the remaining modifications are similar and are left to the reader.

In the D case, the labeling of leaves is easy — the formula λ_0 does it. But in the $D_\subset$ case, a similar formula would label also the intervals of length 1. To avoid it, we use auxiliary variables a, b, c, A, B and the conjunction of the following properties:

(i) Each interval of length at most one is labeled with exactly one of a, b, c, A, B.

(ii) No interval of length greater than 1 is labeled by a, b, c, A, or B.

(iii) Each interval of length at least 2 contains an interval labeled with a, b, or c.

(iv) Each interval of length at least 4 contains intervals with all five auxiliary symbols.

Condition (iii) guarantees that the intervals of length 0 are cannot be labeled with A or B. Observe that the intervals of length 4 contain exactly 2 strict subintervals of length 1 and exactly 3 strict subintervals of length 0, and due to (iv) those 5 intervals have to contain 5 auxiliary symbols, so the intervals of length 1 have to be labeled with A and B. Figure 12.1 contains an example of such labeling.

The formulae expressing properties (i)-(iv) are easy to express using $D_\subset$. Finally, we can define $\lambda_0'' = a \vee b \vee c$.

12.4. Arbitrary orderings

The question whether the D fragment is decidable over the class of all (total) orderings is still open. However, our technique can be used to proof the following proposition.

Proposition 12.5. *The satisfiability problem for the formulae of the $D\bar{D}$ fragment of Halpern–Shoham logic over the class of all total orderings is undecidable.*

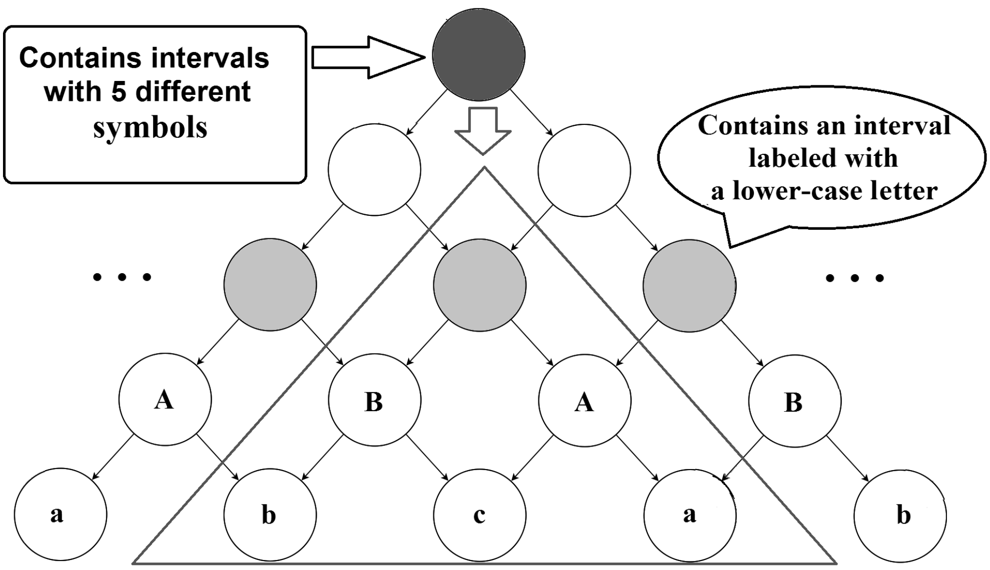

Figure 12.1.: A labeling of leaves in the strict case.

Proof. For the strict D case, consider the formula φ defined as follows

$$[G]([D_{\subset}]\bot \Rightarrow \langle \bar{D}_{\subset}\rangle([D][D]\bot)).$$

This formula is satisfiable in orderings such that for each reachable interval $[a, a]$ there exist b, c such that $b < a < c$ and the interval $[b, c]$ contains at most 4 points (including a, b, c). It implies that a has both predecessor and successor. Therefore the reachable part of the ordering is discrete.

Now we would like to say that all discrete orderings satisfy φ, no matter which initial interval we choose. It is not entirely true — the formula is not satisfied if the interval $[x, x]$ is reachable, where x is the maximal or the minimal point. But this is the case only if the interval $[x, x]$ is initial, so we can simply fix that: let $\varphi' = \varphi \vee ([D_{\subset}]\bot \wedge [\bar{D}_{\subset}]\bot)$.

Now we can use the formula $\Psi^d \wedge \varphi'$ (where Ψ^d is the formula from the proof of the undecidability of the D fragment in the discrete case) to proof the undecidability.

The proper D case can be solved in the same way, however the proof is much more technical. $\square$

The proof bases on the fact that we allow the intervals of the form $[a, a]$. The question of what happens if we exclude such intervals remains open.

Bibliography

[1] Allen, J. F.: Maintaining knowledge about temporal intervals, *Commun. ACM*, **26**(11), November 1983, 832–843.

[2] Bárány, V., Bojanczyk, M.: Finite satisfiability for guarded fixpoint logic, *Inf. Process. Lett.*, **112**(10), 2012, 371–375.

[3] Berger, R.: The undecidability of the domino problem, *Mem. AMS*, **66**, 1966.

[4] Blackburn, P., de Rijke, M., Venema, Y.: *Modal Logic*, vol. 53 of *Cambridge Tracts in Theoretical Comp. Sc.*, Cambridge University Press, Cambridge, 2001.

[5] Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G.: Decidable and Undecidable Fragments of Halpern and Shoham's Interval Temporal Logic: Towards a Complete Classification, in: *Logic for Programming, Artificial Intelligence, and Reasoning*, vol. 5330 of *LNCS*, Springer Berlin / Heidelberg, 2008, 590–604.

[6] Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G.: Undecidability of the Logic of Overlap Relation over Discrete Linear Orderings, *Electron. Notes Theor. Comput. Sci.*, **262**, May 2010, 65–81, ISSN 1571-0661.

[7] Bresolin, D., Goranko, V., Montanari, A., Sala, P.: Tableau-based decision procedures for the logics of subinterval structures over dense orderings, *Journal of Logic and Computation*, 2008.

[8] Bresolin, D., Goranko, V., Montanari, A., Sciavicco, G.: Propositional interval neighborhood logics: Expressiveness, decidability, and undecidable extensions, *Annals of Pure and Applied Logic*, **161**, 2009, 289–304.

[9] Bresolin, D., Montanari, A., Sala, P., Sciavicco, G.: Optimal Tableaux for Right Propositional Neighborhood Logic over Linear Orders, *Logics in Artificial Intelligence*, 2008.

[10] Goranko, V., Montanari, A., Sciavicco, G.: A Road Map of Interval Temporal Logics and Duration Calculi, *Journal of Applied Non-classical Logics*, **14**, 2004, 9–54.

[11] Goranko, V., Passy, S.: Using the Universal Modality: Gains and Questions, *Journal of Logic and Computation*, **2**(1), 1992, 5–30.

[12] Grädel, E.: On the restraining power of guards, *J. Symbolic Logic*, **64**, 1999, 1719–1742.

[13] Grädel, E., Walukiewicz, I.: Guarded fixed point logic, *Fourteenth Annual IEEE Symposium on Logic in Computer Science*, 1999.

[14] Gurevich, Y. S., Koryakov, I. O.: Remarks on Berger's paper on the domino problem, *Siberian Mathematical Journal*, **13**, 1972, 319–321.

[15] Halpern, J. Y., Shoham, Y.: A propositional modal logic of time intervals, *Journal of The ACM*, **38**, 1991, 935–962.

[16] Hamblin, C. L.: Instants and intervals, *Studium Generale*, **27**, 1971, 127–134.

[17] Hemaspaandra, E.: The Price of Universality, *Notre Dame Journal of Formal Logic*, **37**(2), 1996, 174–203.

[18] Hemaspaandra, E., Schnoor, H.: On the Complexity of Elementary Modal Logics, *STACS* (S. Albers, P. Weil, Eds.), 1, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2008.

[19] Hemaspaandra, E., Schnoor, H.: A Universally Defined Undecidable Unimodal Logic, *MFCS* (F. Murlak, P. Sankowski, Eds.), 6907, Springer, 2011.

[20] Kieroński, E., Michaliszyn, J., Otop, J.: Modal Logics Definable by Universal Three-Variable Formulas, *FSTTCS*, 13, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011.

[21] Kieroński, E., Michaliszyn, J., Pratt-Hartmann, I., Tendera, L.: Extending two-variable first-order logic with equivalence closure, *LICS '12: Proceedings of the 29th IEEE symposium on Logic in Computer Science*, 2012.

[22] Kieronski, E., Tendera, L.: On Finite Satisfiability of the Guarded Fragment with Equivalence or Transitive Guards, *LPAR* (N. Dershowitz, A. Voronkov, Eds.), 4790, Springer, 2007, ISBN 978-3-540-75558-6.

[23] Ladner, R. E.: The Computational Complexity of Provability in Systems of Modal Propositional Logic, *Siam Journal on Computing*, **6**, 1977, 467–480.

[24] Lodaya, K.: Sharpening the Undecidability of Interval Temporal Logic, *Asian Computing Science Conference*, 2000.

[25] Marcinkowski, J., Michaliszyn, J.: The Ultimate Undecidability Result for the Halpern-Shoham Logic, *LICS*, IEEE Computer Society, 2011.

[26] Marcinkowski, J., Michaliszyn, J., Kieronski, E.: B and D Are Enough to Make the Halpern-Shoham Logic Undecidable, *ICALP (2)*, 6199, Springer, 2010.

[27] Michaliszyn, J.: Decidability of the Guarded Fragment with the Transitive Closure, *ICALP (2)*, 5556, Springer, 2009.

[28] Michaliszyn, J., Kieroński, E.: Finite Satisfiability of Universally-Horn Definable Modal Logics, *Advances in Modal Logic*, 2012.

[29] Michaliszyn, J., Otop, J.: Decidable Elementary Modal Logics, *LICS '12: Proceedings of the 29th IEEE symposium on Logic in Computer Science*, 2012.

[30] Montanari, A., Pratt-Hartmann, I., Sala, P.: Decidability of the Logics of the Reflexive Sub-interval and Super-interval Relations over Finite Linear Orders, *17th Int. Symposium on Temporal Representation and Reasoning*, 2010, 27–34.

[31] Montanari, A., Puppis, G., Sala, P.: A decidable spatial logic with cone-shaped cardinal directions, *Proceedings of the 23rd CSL*, CSL'09/EACSL'09, Springer-Verlag, Berlin, Heidelberg, 2009.

[32] Montanari, A., Puppis, G., Sala, P.: Maximal Decidable Fragments of Halpern and Shoham's Modal Logic of Intervals., *ICALP (2)*, 6199, Springer, 2010.

[33] Mortimer, M.: On languages with two variables, *Mathematical Logic Quarterly*, **21**(1), 1975, 135–140, ISSN 1521-3870.

[34] Moszkowski, B. C.: *Reasoning about digital circuits*, Ph.D. Thesis, Stanford University, Stanford, CA, USA, 1983.

[35] Otto, M.: Two Variable First-Order Logic over Ordered Domains, *Journal of Symbolic Logic*, **66**, 1998, 685–702.

[36] Parikh, R.: A decidability result for a second order process logic, *Proceedings of the 19th Annual Symposium on Foundations of Computer Science*, IEEE Computer Society, Washington, DC, USA, 1978.

[37] Pratt, V. R.: Process logic: preliminary report, *Proceedings of the 6th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*, POPL '79, ACM, New York, NY, USA, 1979.

[38] Sahlqvist, H.: Completeness and correspondence in the first and second order semantics for modal logic, *Proceedings of the Third Scandinavian Logic Symposium*, 1973.

[39] Savitch, W. J.: Relationships between nondeterministic and deterministic tape complexities, *J. Comput. Syst. Sci.*, **4**, April 1970, 177–192, ISSN 0022-0000.

[40] Schwentick, T., Zeume, T.: Two-Variable Logic with Two Order Relations, in: *Computer Science Logic* (A. Dawar, H. Veith, Eds.), vol. 6247 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2010, 499–513.

[41] Vardi, M. Y.: Why is modal logic so robustly decidable?, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science,*, **31**, 1997, 149–184.