Algorithmic Model Theory
SS 2010

Prof. Dr. Erich Grädel

Mathematische Grundlagen der Informatik
RWTH Aachen

Contents

2 Finite Model Property

We study the finite model property for fragments of FO as a mean to show that these fragments are decidable, and also to better understand their expressive power and algorithmic complexity.

Recall that a class $X \subseteq \text{FO}$ has the *finite model property* if $\mathit{Sat}(X) = \mathit{Fin\text{-}sat}(X)$. Since for any decidable class X, $\mathit{Fin\text{-}sat}(X)$ is r.e. and $\mathit{Sat}(X)$ is co-r.e., it follows that $\mathit{Sat}(X)$ is decidable if X has the FMP. In many cases, the proof that a class has the finite model property provides a bound on the model's cardinality, and thus a complexity bound for the satisfiability problem. To prove completeness for complexity classes we make use of a bounded variant of the domino problem.

2.1 Ehrenfeucht-Fraïssé Games

2.1.1 Atomic Types

Definition 2.1. The *atomic k-type* of $a_1, \ldots, a_k$ in $\mathfrak{A}$ is defined as

$$\begin{aligned}\text{atp}_{\mathfrak{A}}(a_1, \ldots, a_k) := \{\gamma(x_1 \ldots, x_k) : \gamma \text{ atomic formula or negated} \\ \text{atomic formula such that } \mathfrak{A} \models \gamma(a_1, \ldots, a_k)\}.\end{aligned}$$

We assume that all structures contain unary or binary relations only. Hence, to describe a structure it suffices to define its universe and to specify the atomic 1-types and 2-types for all of its elements.

Example 2.2. Let $\mathfrak{A}$ be the structure $(A, E_1, \ldots, E_m)$ where the E_i are binary relations. Then for $a \in A$:

$$\text{atp}_{\mathfrak{A}}(a) = \{E_i xx : \mathfrak{A} \models E_i aa\} \cup \{\neg E_i xx : \mathfrak{A} \models \neg E_i aa\}.$$

Definition 2.3. Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures over the same signature and

$\bar{a} \subseteq A$ and $\bar{b} \subseteq B$. We say that $\mathfrak{A}, \bar{a}$ is locally isomorphic to $\mathfrak{B}, \bar{b}$ and write $\mathfrak{A}, \bar{a} \equiv_0 \mathfrak{B}, \bar{b}$ if $\bar{a}$ has the same atomic type in $\mathfrak{A}$ as $\bar{b}$ in $\mathfrak{B}$.

2.1.2 The Game $EF_m(\mathfrak{A}, \mathfrak{B})$

The Ehrenfeucht-Fraïssé game $EF_m(\mathfrak{A}, \mathfrak{B})$ is played by two players according to the following rules.

The *arena* consists of the structures $\mathfrak{A}$ and $\mathfrak{B}$. We assume that $A \cap B = \emptyset$. The players are called *Spoiler* and *Duplicator*, and a play of $EF_m(\mathfrak{A}, \mathfrak{B})$ consists of m moves.

In the *i-th move*, Spoiler chooses either an element $a_i \in A$ or an element $b_i \in B$. Duplicator answers by choosing an element in the other structure.

After m moves, elements $a_1, \ldots, a_m$ from $\mathfrak{A}$ and $b_1, \ldots, b_m$ from $\mathfrak{B}$ are chosen. Duplicator wins the play if $\mathfrak{A}, (a_1, a_2, \ldots, a_m) \equiv_0 \mathfrak{B}, (b_1, b_2, \ldots, b_m)$. Otherwise Spoiler wins.

After i moves in $EF_m(\mathfrak{A}, \mathfrak{B})$ are made, a position $(a_1, \ldots, a_i, b_1, \ldots, b_i)$ is reached. We denote the remaining subgame in which $m - i$ moves are left by $EF_{m-i}(\mathfrak{A}, a_1, \ldots, a_i, \mathfrak{B}, b_1, \ldots, b_i)$.

A *winning strategy* of Spoiler for such a subgame is a function which, for every reachable position, determines a move such that Spoiler wins each play which is consistent with this strategy, no matter how Duplicator plays. Winning strategies for Duplicator are defined analogously.

We say that *Spoiler (respectively, Duplicator) wins the game* $EF_m(\mathfrak{A}, \mathfrak{B})$ if this player has a winning strategy for $EF_m(\mathfrak{A}, \mathfrak{B})$. By induction on the number of moves it is easy to show that for every (sub)game exactly one of the two players has a winning strategy.

Example 2.4.

- Let $\mathfrak{A} = (\mathbb{Z}, <)$, $\mathfrak{B} = (\mathbb{R}, <)$. Then Duplicator wins $EF_2(\mathfrak{A}, \mathfrak{B})$, but Spoiler wins $EF_3(\mathfrak{A}, \mathfrak{B})$.
- For a relational signature $\tau = \{E, P\}$ (where P has arity one and E has arity two), consider the structures $\mathfrak{A}$ and $\mathfrak{B}$ in Figure 2.1. Spoiler wins the game $EF_3(\mathfrak{A}, \mathfrak{B})$, but Duplicator wins $EF_2(\mathfrak{A}, \mathfrak{B})$.

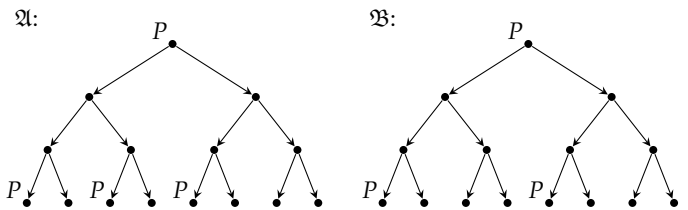


Figure 2.1. Two structures$\mathfrak{A}$ and $\mathfrak{B}$ with $\mathfrak{A} \equiv_2 \mathfrak{B}$ and $\mathfrak{A} \not\equiv_3 \mathfrak{B}$

2.1.3 The Game $EF(\mathfrak{A}, \mathfrak{B})$

An important variant is the Ehrenfeucht-Fraïssé game $EF(\mathfrak{A}, \mathfrak{B})$ in which plays of arbitrary length are possible. In each play, Spoiler first chooses an $m \in \mathbb{N}$, and then the players play the game $EF_m(\mathfrak{A}, \mathfrak{B})$.

Spoiler wins the game $EF(\mathfrak{A}, \mathfrak{B})$ if and only if there exists an $m \in \mathbb{N}$ such that he wins the game $EF_m(\mathfrak{A}, \mathfrak{B})$. In other words: Duplicator wins $EF(\mathfrak{A}, \mathfrak{B})$ if and only if she has a winning strategy for each of the games $EF_m(\mathfrak{A}, \mathfrak{B})$.

Recall that two structures $\mathfrak{A}$ and $\mathfrak{B}$ are said to be *elementarily m-equivalent*, written $\mathfrak{A} \equiv_m \mathfrak{B}$, if no first-order formula of quantifier rank at most m can separate both structures. If $\mathfrak{A} \equiv_m \mathfrak{B}$ for all $m \in \mathbb{N}$ we write $\mathfrak{A} \equiv \mathfrak{B}$ and say that $\mathfrak{A}$ and $\mathfrak{B}$ are *elementarily equivalent*. The following theorem shows that elementary equivalence and Ehrenfeucht-Fraïssé games are in some sense equivalent concepts.

Theorem 2.5 (Ehrenfeucht, Fraïssé)**.** Let τ be finite and relational, and let $\mathfrak{A}, \mathfrak{B}$ be τ-structures.

(1) The following statements are equivalent:
   (i) $\mathfrak{A} \equiv \mathfrak{B}$.
   (ii) Duplicator wins the Ehrenfeucht-Fraïssé game $EF(\mathfrak{A}, \mathfrak{B})$.

(2) For all $m \in \mathbb{N}$ the following statements are equivalent:
   (i) $\mathfrak{A} \equiv_m \mathfrak{B}$.
   (ii) Duplicator wins $EF_m(\mathfrak{A}, \mathfrak{B})$.

In fact, even the following, somewhat stronger proposition holds (for a proof see the lecture notes of mathematical logic).

Theorem 2.6. Let $\mathfrak{A}, \mathfrak{B}$ be τ-structures, $\bar{a} = a_1, \ldots, a_r \in A$, $\bar{b} = b_1, \ldots, b_r \in B$. If there exists a formula $\psi(\bar{x})$ with $\mathrm{qr}(\psi) = m$ such that $\mathfrak{A} \models \psi(\bar{a})$ and $\mathfrak{B} \models \neg\psi(\bar{b})$ holds, then Spoiler has a winning strategy for the game $G_m(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$.

We use the above to prove finite model property of the following fragment of FO.

Theorem 2.7. If τ contains only unary predicates then FO$[\tau]$ has FMP.

Proof. Let $\mathfrak{A} = (A, P_1, \ldots, P_n)$ and let $\mathrm{qr}(\varphi) = m$. For each sequence of bits $\alpha = \alpha_1 \ldots \alpha_n$ we define $P_\alpha = Q_1 \cap Q_2 \cap \ldots \cap Q_n$, where $Q_i = P_i$ if $\alpha_i = 1$ and Q_i is the complement of P_i else.

Note that $\{\alpha \mid x \in P_\alpha\}$ determines all atomic types of x. We construct $\mathfrak{B}$ by taking $\min(|P_\alpha|, m)$ elements into each $P_\alpha^{\mathfrak{B}}$. Observe that $\mathfrak{B}$ is defined in this way (take $P_i^{\mathfrak{B}} = \bigcup_{\alpha|\alpha_i=1} P_\alpha^{\mathfrak{B}}$). We show that $\mathfrak{A} \equiv_m \mathfrak{B}$ using the Ehrenfeucht-Fraïssé Theorem.

The following is a winning strategy for Duplicator in $EF(\mathfrak{A}, \mathfrak{B})$: Answer each Spoiler's choice of an element with an element of the same atomic type in the other structure. Due to the construction it is possible to do that for m moves. It also follows from the construction that $\equiv_0$ is never violated and Duplicator wins the game. Q.E.D.

You can see from the proof that the constructed finite model $\mathfrak{B}$ is a sub-model of $\mathfrak{A}$. It is not always the case, sometimes it is not possible to find a finite sub-model, even for fragments with FMP.

2.2 FMP of Modal Logic

We proceed with proving that propositional modal logic (ML), which is an important fragment of FO^2, has the finite model property. In fact we establish an even stronger result showing that every satisfiable ML-formula has a finite model that is a tree. Hence, we prove that ML has the *finite tree model property*.

2.2.1 Modal Logic

Let us first briefly review the syntax and semantics of propositional modal logic (ML).

Definition 2.8. For a given set of actions A and atomic properties $\{P_i : i \in I\}$, the syntax of ML is inductively defined as:

- All propositional logic formulae with propositional variables P_i are in ML.
- If $\psi, \varphi \in$ ML, then also $\neg\psi$, $(\psi \wedge \varphi)$ and $(\psi \vee \varphi) \in$ ML.
- If $\psi \in$ ML and $a \in A$, then $\langle a \rangle \psi$ and $[a]\psi \in$ ML.

Remark 2.9. If there is only one action $a \in A$, we write $\Diamond\psi$ and $\Box\psi$ instead of $\langle a \rangle \psi$ and $[a]\psi$, respectively.

Definition 2.10. A *transition system* or *Kripke structure* with actions from a set A and atomic properties $\{P_i : i \in I\}$ is a structure

$$\mathcal{K} = (V, (E_a)_{a \in A}, (P_i)_{i \in I})$$

with a universe V of states, binary relations $E_a \subseteq V \times V$ describing transitions between the states, and unary relations $P_i \subseteq V$ describing the atomic properties of states.

A transition system can be seen as a labelled graph where the nodes are the states of $\mathcal{K}$, the unary relations are labels of the states, and the binary transition relations are the labelled edges.

Definition 2.11. Let $\mathcal{K} = (V, (E_a)_{a \in A}, (P_i)_{i \in I})$ be a transition system, $\psi \in$ ML a formula and v a state of $\mathcal{K}$. The *model relationship* $\mathcal{K}, v \models \psi$, i.e. ψ holds at state v of $\mathcal{K}$, is inductively defined:

- $\mathcal{K}, v \models P_i$ if and only if $v \in P_i$.
- $\mathcal{K}, v \models \neg\psi$ if and only if $\mathcal{K}, v \not\models \psi$.
- $\mathcal{K}, v \models \psi \vee \varphi$ if and only if $\mathcal{K}, v \models \psi$ or $\mathcal{K}, v \models \varphi$.
- $\mathcal{K}, v \models \psi \wedge \varphi$ if and only if $\mathcal{K}, v \models \psi$ and $\mathcal{K}, v \models \varphi$.
- $\mathcal{K}, v \models \langle a \rangle \psi$ if and only if there exists w such that $(v, w) \in E_a$ and $\mathcal{K}, w \models \psi$.
- $\mathcal{K}, v \models [a]\,\psi$ if and only if $\mathcal{K}, w \models \psi$ holds for all w with $(v, w) \in E_a$.

Definition 2.12. For a transition system $\mathcal{K}$ and a formula ψ we define the *extension*

$$[\![\psi]\!]^{\mathcal{K}} := \{v : \mathcal{K}, v \models \psi\}$$

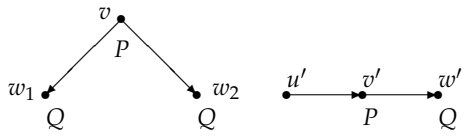
as the set of states of $\mathcal{K}$ where ψ holds.

2.2.2 Bisimulation

One of the most important notions in the analysis of modal logics is *bisimulation*. In fact bisimulation is closely related to logical equivalence of Kripke structures with respect to formulae from ML.

Definition 2.13. Let $\mathcal{K} = (V, (E_a)_{a\in A}, (P_i)_{i\in I})$ and $\mathcal{K}' = (V', (E'_a)_{a\in A}, (P'_i)_{i\in I})$ be transition systems. A *bisimulation* between $\mathcal{K}$ and $\mathcal{K}'$ is a relation $Z \subseteq V \times V'$ such that for all $(v, v') \in Z$
(Pred) $v \in P_i$ if and only if $v' \in P'_i$ for all $i \in I$,
(Forth) for all $a \in A$, $w \in V$ with $v \xrightarrow{a} w$ there exists a $w' \in V'$ with $v' \xrightarrow{a} w'$ and it is $(w, w') \in Z$,
(Back) for all $a \in A$, $w' \in V'$ with $v' \xrightarrow{a} w'$ there exists a $w \in V$ with $v \xrightarrow{a} w$ and it is $(w, w') \in Z$.

Example 2.14.



$Z = \{(v, v'), (w_1, w'), (w_2, w')\}$ is a bisimulation.

Definition 2.15. Let $\mathcal{K}, \mathcal{K}'$ be Kripke structures and let $u \in V$, $u' \in V'$. $(\mathcal{K}, u)$ and $(\mathcal{K}', u')$ are *bisimilar* (for short, $\mathcal{K}, u \sim \mathcal{K}', u'$), if there exists a bisimulation Z between $\mathcal{K}$ and $\mathcal{K}'$ such that $(u, u') \in Z$.

2.2.3 Bisimulation Invariance of Formulae of Modal Logic

The fundamental importance of bisimulation origins in the fact that formulae of modal logic are not able to distinguish between bisimilar

states. A more refined analysis considers the modal depth of formulae, i.e. the maximal depth of nesting of modal operators in a formula.

Definition 2.16. The *modal depth* of a formula $\psi \in$ ML is defined inductively by

(1) $\mathrm{md}(\psi) = 0$ for propositional formulae ψ,
(2) $\mathrm{md}(\neg\psi) = \mathrm{md}(\psi)$,
(3) $\mathrm{md}(\psi \circ \varphi) = \max(\mathrm{md}(\psi), \mathrm{md}(\varphi))$ for $\circ \in \{\wedge, \vee, \to\}$,
(4) $\mathrm{md}(\langle a\rangle\psi) = \mathrm{md}([a]\psi) = \mathrm{md}(\psi) + 1$.

Definition 2.17. Let $\mathcal{K}$ and $\mathcal{K}'$ be two Kripke structures and let $v \in \mathcal{K}$, $v' \in \mathcal{K}'$.

(1) $\mathcal{K}, v \equiv_{\mathrm{ML}} \mathcal{K}', v'$ if for all $\psi \in$ ML we have $\mathcal{K}, v \models \psi$ if and only if $\mathcal{K}', v' \models \psi$.
(2) $\mathcal{K}, v \equiv^n_{\mathrm{ML}} \mathcal{K}', v'$ if for all $\psi \in$ ML with $\mathrm{md}(\psi) \leq n$ we have $\mathcal{K}, v \models \psi$ if and only if $\mathcal{K}', v' \models \psi$.

One can refine the definition of the bisimilarity relation between transition systems as well. We say that $(\mathcal{K}, u)$ and $(\mathcal{K}', u')$ are *n-bisimilar* (for short, $\mathcal{K}, u \sim_n \mathcal{K}', u'$), if there exists a relation Z between $\mathcal{K}$ and $\mathcal{K}'$ such that $(u, u') \in Z$ and Z has the property 'Pred' and the 'Forth' and 'Back' property for all pairs of nodes $(v, v') \in Z$ with distance at most n from (u, u'). For a formal (game theoretical) definition, see the lectures notes of mathematical logic.

Theorem 2.18. For Kripke structures $\mathcal{K}$, $\mathcal{K}'$ and $u \in \mathcal{K}$, $u' \in \mathcal{K}'$ the following holds:

(1) $\mathcal{K}, u \sim \mathcal{K}', u' \Rightarrow \mathcal{K}, u \equiv_{\mathrm{ML}} \mathcal{K}', u'$.
(2) $\mathcal{K}, u \sim_n \mathcal{K}', u' \Rightarrow \mathcal{K}, u \equiv^n_{\mathrm{ML}} \mathcal{K}', u'$.

Statement (1) is called the *bisimulation invariance of modal logic*:

$$\text{If } \mathcal{K}, v \models \psi \text{ and } \mathcal{K}, v \sim \mathcal{K}', v', \text{ then } \mathcal{K}', v' \models \psi.$$

The reverse only holds for finitely branching systems. A transition system is *finitely branching* if for all states v and all actions a the set $vE_a := \{w : (v, w) \in E_a\}$ of a-successors of v is finite. (for proofs see the lecture notes of mathematical logic).

Theorem 2.19. Let $\mathcal{K}, \mathcal{K}'$ be finitely branching transitions systems. Then

$$\mathcal{K}, u' \sim \mathcal{K}', u' \text{ if and only if } \mathcal{K}, u \equiv_{\mathrm{ML}} \mathcal{K}', u'.$$

2.2.4 Tree Model Property

Definition 2.20. A transition system $\mathcal{K} = (V, (E_a)_{a\in A}, (P_i)_{i\in I})$ with a marked node w is a *tree* if

(1) $E_a \cap E_b = \emptyset$ for all actions $a \neq b$,
(2) (V, E) is a (directed) tree with root w in the graph theoretical sense, where $E = \bigcup_{a\in A} E_a$.

Definition 2.21. Let Φ be a set of formulae (of some logic, e.g. of modal logic or first-order logic) over a signature which contains at most binary relations and no functions.

(1) Φ has the *finite model property* (FMP) if every satisfiable formula $\varphi \in \Phi$ has a finite model.
(2) Φ has the *tree model property* (TMP) if every satisfiable formula in Φ has a tree as a model.
(3) Φ has *finite tree model property* if every satisfiable formula in Φ has a finite tree as a model.

We shall prove that formulae of modal logic have the finite tree model property. For that consider *unfoldings* of transition systems. The unfolding of $\mathcal{K}$ from state v consists of all paths in $\mathcal{K}$ that start with v. Hereby every path is considered as a distinguished object, i.e. even if two paths intersect, the unfolding $\mathcal{T}$ contains several copies of the intersection points and each state from $\mathcal{K}$ that is reachable from v via a path is added to the unfolding, no matter whether is has already been reached. Self-loops in $\mathcal{K}$ correspond thus to infinite paths in the unfolding. Formally, unfoldings are defined as follows.

Definition 2.22. Let $\mathcal{K} = (V^{\mathcal{K}}, (E_a^{\mathcal{K}})_{a\in A}, (P_i^{\mathcal{K}})_{i\in I})$ be a Kripke structure and let $v \in V^{\mathcal{K}}$. The *unfolding* of $\mathcal{K}$ from v is the Kripke structure $\mathcal{T}_{\mathcal{K},v} = (V^{\mathcal{T}}, (E_a^{\mathcal{T}})_{a\in A}, (P_i^{\mathcal{T}})_{i\in I})$ with

$$V^{\mathcal{T}} = \{\bar{v} = v_0 a_0 v_1 a_1 v_2 \dots v_{m-1} a_{m-1} v_m : m \in \mathbb{N},$$

$$v_0 = v,\ v_i \in V^{\mathcal{K}},\ a_i \in A,\ (v_i, v_{i+1}) \in E_{a_i}^{\mathcal{K}} \text{ for all } i < m\}$$
$$E_a^{\mathcal{T}} = \{(\bar{v}, \bar{w}) \in V^{\mathcal{T}} \times V^{\mathcal{T}} : \bar{w} = \bar{v}aw \text{ for some } w \in V^{\mathcal{K}}, a \in A\}$$
$$P_i^{\mathcal{T}} = \{\bar{v} = v_0 a_0 \dots v_m \in V^{\mathcal{T}} : v_m \in P_i^{\mathcal{K}}\}.$$

We write $\mathrm{End}(\bar{v})$ for the last state on the path $\bar{v}$, so we have $\bar{v} \in P_i^{\mathcal{T}}$ if and only if $\mathrm{End}(\bar{v}) \in P_i^{\mathcal{K}}$.

Lemma 2.23. For all Kripke structures $\mathcal{K}$ and all states v in $\mathcal{K}$ we have $\mathcal{K}, v \sim \mathcal{T}_{K,v}, v$.

Proof. $Z := \{(w, \bar{w}) \in V^{\mathcal{K}} \times V^{\mathcal{T}} : \mathrm{End}(\bar{w}) = w\}$ is a bisimulation from $\mathcal{K}$ to $\mathcal{T}_{\mathcal{K},v}$ with $(v, v) \in Z$. Q.E.D.

Theorem 2.24. ML has the tree model property.

Proof. Let ψ be an arbitrary satisfiable formula from ML. Then there is a model $\mathcal{K}, v \models \psi$. Let $\mathcal{T} := \mathcal{T}_{\mathcal{K},v}$ be he unfolding of $\mathcal{K}, v$. As $\mathcal{K}, v \sim \mathcal{T}, v$, due to the bisimulation invariance of modal logic we have $\mathcal{T}, v \models \psi$. Thus ψ has a tree model. Q.E.D.

The same argument shows that *every* class of bisimulation invariant formulae has the tree model property.

2.2.5 Finite Model Property

For ML, we can prove a stronger result. For this, we use the notion of the closure $C(\psi)$ of a formula ψ.

Definition 2.25. For every formula $\psi \in \mathrm{ML}$ we inductively define for all $n \in \mathbb{N}$ the sets of formulae $C_n(\psi)$ as follows:

(1) $\psi \in C_0(\psi)$.
(2) If $\neg\varphi \in C_n(\psi)$ then also $\varphi \in C_n(\psi)$.
(3) If $(\varphi \wedge \vartheta) \in C_n(\psi)$ or $(\varphi \vee \vartheta) \in C_n(\psi)$ then also $\varphi \in C_n(\psi)$ and $\vartheta \in C_n(\psi)$.
(4) If $\langle a\rangle\varphi \in C_n(\psi)$ or $[a]\varphi \in C_n(\psi)$ then $\varphi \in C_{n+1}(\psi)$.

Finally let $C(\psi) := \bigcup_{n\in\mathbb{N}} C_j(\psi)$.

The closure $C(\psi)$ contains those formulae from ML that are substantial for the evaluation of ψ; $C_j(\psi)$ are hereby formulae that appear in ψ within j nested modal operators. Notice that $|C(\psi)| \leq 2|\psi|$ (negated formulas are added) and that $C_n(\psi) = \emptyset$ for all $n > \mathrm{md}(\psi)$.

Theorem 2.26. For every satisfiable formula $\psi \in \mathrm{ML}$ there is a finite tree structure $\mathcal{T}, v$ of depth $\leq \mathrm{md}(\psi)$ and branching factor $\leq |C(\psi)|$ such that $\mathcal{T}, v \models \psi$. Thus ML has finite tree model property.

Proof. Without loss of generality we can assume that ψ is in negation normal form. As ψ is satisfiable, there exists a tree model $\mathcal{T}, u \models \psi$. The depth of a node of $\mathcal{T}$ is its distance from the root. We define now a labelling function S which assigns a subset of $C_m(\psi)$ to every node v of $\mathcal{T}$ of depth m, namely

$$S(v) := \{\varphi \in C_m(\psi) : \mathcal{T}, v \models \varphi\}.$$

We transform $\mathcal{T}$ in a finite tree structure by successively deleting unnecessary subtrees. Let $\mathcal{T}' \subseteq \mathcal{T}$ be some subtree of $\mathcal{T}$ and let v be a node of $\mathcal{T}'$. Notice that $\mathcal{T}, v \models S(v)$. The following lemma provides a sufficient condition for $\mathcal{T}', v \models S(v)$.

Lemma 2.27. Let the subtree $\mathcal{T}' \subseteq \mathcal{T}$ be constructed in a way that the following conditions are fulfilled.

(1) For every successor w of v in $\mathcal{T}'$ we have $\mathcal{T}', w \models S(w)$.
(2) For every formula of the form $\langle a \rangle \varphi \in S(v)$ there exists an a-successor $w_{\langle a \rangle \varphi}$ of v in the tree $\mathcal{T}'$ such that $\mathcal{T}', w_{\langle a \rangle \varphi} \models \varphi$.

Then it is $\mathcal{T}', v \models S(v)$.

Proof. Each formula in $S(v)$ is a combination of formulae of the form $P_i, \neg P_i, \langle a \rangle \varphi$ and $[a]\varphi$ that are built with $\wedge$ and $\vee$. So it suffices to show for every formula ϑ of this form that $\mathcal{T}, v \models \vartheta$ implies $\mathcal{T}', v \models \vartheta$. For $\vartheta = P_i$ and $\vartheta = \neg P_i$ this is clear as the atomic properties of the node v are the same in $\mathcal{T}$ and $\mathcal{T}'$. For formulae $[a]\varphi$ this follows from condition (1) and for formulae $\langle a \rangle \varphi$ from condition (2). Q.E.D.

Now we can construct a finite subtree $\mathcal{T}'$ as follows. First, let v be the root of $\mathcal{T}$. For every formula of the form $\langle a \rangle \varphi \in S(v)$ we choose an a-successor $w_{\langle a \rangle \varphi} \in vE_a$ such that $\mathcal{T}, w_{\langle a \rangle \varphi} \models \varphi$ holds and delete all not chosen successor nodes of v and the trees that have those nodes as roots from $\mathcal{T}$. We continue this process for all remaining nodes of depth $1, 2, \ldots$. As the labelling $S(v)$ of nodes of depth $m = \mathrm{md}(\psi)$ only consists of formulae P_i and $\neg P_i$, the resulting tree has depth at most m. Every node v has at most $|S(v)| \leq C(\psi)$ successors such that the branching factor of $\mathcal{T}'$ is bounded by $|C(\psi)|$.

It follows by inductively proceeding from leaves to the root of $\mathcal{T}'$ that $\mathcal{T}', v \models S(v)$, in particular, $\mathcal{T}', v \models \psi$. Q.E.D.

2.3 Finite Model Property of FO^2

We denote relational first-order logic over k variables by FO^k, i.e.

$$\mathrm{FO}^k := \{\varphi \in \mathrm{FO} : \varphi \text{ relational, } \varphi \text{ only contains } k \text{ variables}\}.$$

One result of the previous chapter was that $[\forall\exists\forall, \mathrm{all}, (0)] \subseteq \mathrm{FO}^3$ is a conservative reduction class. We now prove that FO^2 has the finite model property and is thus decidable. Note that FO^k formulae are not necessarily in prenex normal form. A further motivation for the study of FO^2 is that propositional modal logic can be viewed as a fragment of FO^2 (in fact ML can be proven to be precisely the bisimulation invariant fragment of FO^2).

Before we proceed to prove the finite model property for FO^2, as a first step we establish a normal form for formulae in FO^2.

Lemma 2.28 (Scott)**.** For each sentence $\psi \in \mathrm{FO}^2$ one can construct in polynomial time a sentence $\varphi \in \mathrm{FO}^2$ of the form

$$\varphi := \forall x \forall y \alpha \wedge \bigwedge_{i=1}^{n} \forall x \exists y \beta_i$$

such that $\alpha, \beta_1, \ldots, \beta_n$ are quantifier free and such that ψ and φ are satisfiable over the same universe. Moreover, we have $|\varphi| = \mathcal{O}(|\psi| \cdot \log |\psi|)$.

Proof. First of all, we can assume that formulae $\varphi \in \mathrm{FO}^2$ only contain unary and binary relation symbols. This is no restriction since relations of higher arity can be substituted by introducing new binary and unary relation symbols. For example, if R is a relation of arity three, one could add a unary relation R_x and three binary relations $R_{x,x,y}$, $R_{x,y,x}$ and $R_{x,y,y}$ and replace each atom $R(x,x,x)$ (or $R(y,y,y)$) by $R_x(x)$ (or $R_x(y)$) and atoms as $R(x,x,y)$ or $R(x,y,x)$ by $R_{x,x,y}(x,y)$ and $R_{x,y,x}(x,y)$ respectively. By adding appropriate new subformulae one can ensure that the semantics are preserved, i.e. that the newly introduced relations partition a ternary relation in the intended sense. For example we would introduce as a new subformula $\forall x(R_x(x) \leftrightarrow R_{x,x,y}(x,x))$.

With ψ containing at most binary relations, we iterate the following steps until ψ has the desired form. We choose a subformula $Qy\eta$ of ψ ($Q \in \{\forall, \exists\}$, η quantifier free) and add a new unary relation R:

$$\begin{array}{rcl} \psi' & := & \psi[Qy\eta / Rx] \\ \psi & \mapsto & \psi' \wedge \forall x(Rx \leftrightarrow Qy\eta). \end{array}$$

R captures those x that satisfy $Qy\eta$. The resulting formula φ is not yet of the desired form, but it is equivalent to the following:

(a) if $Q = \exists$, then

$$\varphi \equiv \psi' \wedge \forall x \forall y(\eta \to Rx) \wedge \forall x \exists y(Rx \to \eta)$$

(b) else if $Q = \forall$, then

$$\varphi \equiv \psi' \wedge \forall x \forall y(Rx \to \eta) \wedge \forall x \exists y(\eta \to Rx)$$

Now use that conjunctions of $\forall\forall$-formulae are equivalent to a $\forall\forall$-formula and obtain $\psi \equiv \forall x \forall y \alpha \wedge \bigwedge_{i=1}^{n} \forall x \exists y \beta_i$. Q.E.D.

Theorem 2.29. FO^2 has the finite model property. In fact, every satisfiable formula $\psi \in \mathrm{FO}^2$ has a model with at most $2^{|\psi|}$ elements.

Proof. The proof strategy is as follows: we start with a model $\mathfrak{A}$ of ψ and

proceed by constructing a new model $\mathfrak{B}$ of ψ such that $|\mathfrak{B}| \leq 2^{\mathcal{O}(|\psi|)}$. For the construction the following definitions will be essential.

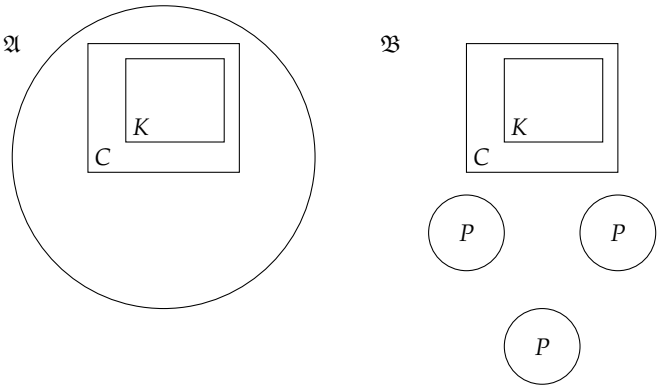
An element $a \in A$ is said to be a *king of* $\mathfrak{A}$ if its atomic 1-type is unique in $\mathfrak{A}$, i.e. if $\mathrm{atp}_{\mathfrak{A}}(b) \neq \mathrm{atp}_{\mathfrak{A}}(a)$ for all $b \neq a$. We let

- $K := \{a \in A : a \text{ is a king of } \mathfrak{A}\}$ be the set of kings of $\mathfrak{A}$, and
- $P := \{\mathrm{atp}_{\mathfrak{A}}(a) : a \in A, a \notin K\}$ be the set of atomic 1-types which are realized at least twice in $\mathfrak{A}$.

Since $\mathfrak{A} \models \forall x \exists y \beta_i$ for $i = 1, \ldots, n$, there exist (Skolem) functions $f_1, \ldots, f_n : A \to A$ such that $\mathfrak{A} \models \beta_i(a, f_i a)$ for all $a \in A$. The *court* of $\mathfrak{A}$ is defined as

$$C := K \cup \{f_i k : k \in K, i = 1, \ldots, n\}.$$

Let $\mathfrak{C}$ be the substructure of $\mathfrak{A}$ induced by C. We construct a model $\mathfrak{B} \models \psi$ with universe $B = C \cup (P \times \{1, \ldots, n\} \times \{0, 1, 2\})$.



To specify $\mathfrak{B}$ we set $\mathfrak{B}|_C = \mathfrak{C}$ and for all other elements we specify the 1- and 2-types (in this way fixing $\mathfrak{B}$ on the remaining part). However,

(1) This must be done consistently:
   - $\mathrm{atp}_{\mathfrak{A}}(b, b')$ and $\mathrm{atp}_{\mathfrak{A}}(b, b'')$ must agree on $\mathrm{atp}_{\mathfrak{A}}(b)$, and
   - $\gamma(x,y) \in \mathrm{atp}_{\mathfrak{B}}(b, b') \Leftrightarrow \gamma(y,x) \in \mathrm{atp}_{\mathfrak{B}}(b', b)$.

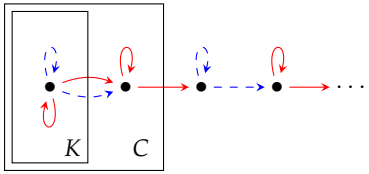
(2) Of course we have to ensure that $\mathfrak{B} \models \psi$.

We illustrate the construction with the following example.

Example 2.30. Consider the formula ψ over the signature $\tau = \{R, B\}$ (red edges and blue edges).

$$\begin{aligned}
\psi \quad &= \quad \exists x(Rxx \wedge Bxx) \\
&\wedge \quad \forall x\forall y((Rxx \wedge Bxx \wedge Ryy \wedge Byy \to x = y) \\
&\qquad \wedge(Rxx \vee Bxx) \\
&\qquad \wedge(Rxy \wedge Ryx \to x = y) \\
&\qquad \wedge(Bxy \wedge Byx \to x = y) \\
&\qquad \wedge(Bxy \wedge x \neq y \to Ryy)) \\
&\wedge \quad \forall x\exists y(x \neq y \wedge (Rxx \to Rxy) \\
&\wedge \quad (Bxx \to Bxy)).
\end{aligned}$$

Let $\mathfrak{A} \models \psi$, then $\mathfrak{A}$ looks like follows:



In this case $P = \{\{Rxx, \neg Bxx\}, \{\neg Rxx, Bxx\}\}$ and the universe of $\mathfrak{B}$ is $B = C \cup (P \times \{1\} \times \{0, 1, 2\})$.

We proceed to construct $\mathfrak{B}$ by specifying the 1-types and 2-types of its elements as follows.

(1) The atomic 1-types of elements (p, i, j) are set to $\mathrm{atp}_{\mathfrak{B}}((p, i, j)) = p$.

(2) The atomic 2-types $\mathrm{atp}_{\mathfrak{B}}(b, b')$ will be set so that $\mathfrak{B} \models \forall x\exists y\beta_i$ for $i = 1, \ldots, m$.
Choose for each $p \in P$ an element $h(p) \in A$ with $\mathrm{atp}_{\mathfrak{A}}(h(p)) = p$.
Find for each $b \in \mathfrak{B}$ and each i a suitable element b' such that $\mathfrak{B} \models \beta_i(b, b')$ (by defining $\mathrm{atp}_{\mathfrak{B}}(b, b')$ appropriately).

(a) If b is a king, set $b' := f_i(b) \in C \subseteq B$. Then $\mathfrak{B} \models \beta_i(b, b')$.

(b) If $b \in C \setminus K$ (non-royal member of the court), distinguish:

- If $f_i(b) \in K$, then set $b' := f_i(b) \in K \subseteq B$.

- Otherwise it holds that $\mathrm{atp}_{\mathfrak{A}}(f_i(b)) = p \in P$.
  In this case, set $b' := (p, i, 0)$. Now set $\mathrm{atp}_{\mathfrak{B}}(b, b') := \mathrm{atp}_{\mathfrak{A}}(b, f_i(b))$. Thus $\mathfrak{B} \models \beta_i(b, b')$ since $\mathfrak{A} \models \beta_i(b, f_i(b))$.

(c) If $b = (p, j, \ell)$ for some $p \in P, j \in \{1, \ldots, n\}, \ell \in \{0, 1, 2\}$, let $a := h(p)$ and consider $f_i(a)$.
If $f_i(a) \in K$, set $b' = f_i(a)$ and $\mathrm{atp}_{\mathfrak{B}}(b, b') := \mathrm{atp}_{\mathfrak{A}}(a, b')$.
If $f_i(a) \notin K$, then $\mathrm{atp}_{\mathfrak{A}}(f_i(a)) = p' \in P$.
Set $b' := (p', i, (\ell + 1) \pmod 3)$.
Then set $\mathrm{atp}_{\mathfrak{B}}(b, b') := \mathrm{atp}_{\mathfrak{A}}(a, f_i(a))$, and thus $\mathfrak{B} \models \beta_i(b, b')$.

To complete the construction of $\mathfrak{B}$, let $b_1, b_2 \in B$ be such that $\mathrm{atp}_{\mathfrak{B}}(b_1, b_2)$ is not yet specified. Choose $a_1, a_2 \in A$ so that

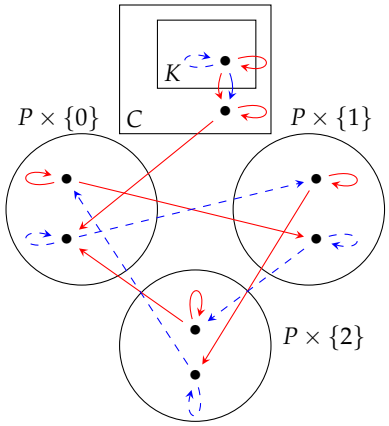
$$\begin{aligned}
\mathrm{atp}_{\mathfrak{A}}(a_1) &= \mathrm{atp}_{\mathfrak{B}}(b_1) \text{ and} \\
\mathrm{atp}_{\mathfrak{A}}(a_2) &= \mathrm{atp}_{\mathfrak{B}}(b_2)
\end{aligned}$$

and set

$$\mathrm{atp}_{\mathfrak{B}}(b_1, b_2) := \mathrm{atp}_{\mathfrak{A}}(a_1, a_2).$$

Since $\mathfrak{A} \models \alpha(a_1, a_2)$, also $\mathfrak{B} \models \alpha(b_1, b_2)$.

For the previously considered example, $\mathfrak{B}$ looks as follows:



Overall, we obtain $\mathfrak{B} \models \forall x\forall y\alpha \wedge \bigwedge_{i=1}^{n} \forall x\exists y\beta_i = \psi$, and the size of B

is restricted by

$$|B| = \underbrace{|C|}_{\leq |K|(n+1)} + 3n|P| = \mathcal{O}(n \cdot \#\,(\text{atomic 1-types}))\,.$$

For k relation symbols, there are 2^k atomic 1-types, hence $|B| = 2^{\mathcal{O}(|\psi|)}$.

Q.E.D.

This result implies that $Sat(\mathrm{FO}^2)$ is in NEXPTIME (indeed it is NEXPTIME-complete), since we can simply guess a finite structure $\mathfrak{A}$ of exponential size (in the length of ψ) and verify that $\mathfrak{A} \models \psi$.

Corollary 2.31. $Sat(\mathrm{FO}^2) \in \mathrm{NEXPTIME} = (\bigcup_k \mathrm{NTIME}(2^{n^k}))$.

This is a typical complexity level for decidable fragments of FO. In fact, $Sat(\mathrm{FO}^2)$ is even complete for NEXPTIME. For showing this, we reduce a bounded version of the domino problem to $Sat(\mathrm{FO}^2)$.

Definition 2.32. Let $\mathcal{D} = (D, H, V)$ be a domino system and let $Z(t)$ denote $\mathbb{Z}/t\mathbb{Z} \times \mathbb{Z}/t\mathbb{Z}$. For a word $w = w_0, \ldots, w_{n-1} \in D^n$ we say that $\mathcal{D}$ tiles $Z(t)$ with initial condition w if there is $\tau : Z(t) \to D$ such that

- if $\tau(x, y) = d$ and $\tau(x+1, y) = d'$ then $(d, d') \in H$ for all $(x, y) \in Z(t)$,
- if $\tau(x, y) = d, \tau(x, y+1) = d'$ then $(d, d') \in V$ for all $(x, y) \in Z(t)$ and
- $\tau(i, 0) = w_i$ for all $i = 0, \ldots, n-1$.

Let $\mathcal{D}$ be a domino system and $T : \mathbb{N} \to \mathbb{N}$ a mapping. Define

$$\mathrm{DOMINO}(\mathcal{D}, T) := \{w \in D^* : \mathcal{D} \text{ tiles } Z(T(|w|)) \text{ with initial condition } w\}\,.$$

As before we describe a computation of a (in this case non-deterministic) Turing machine by a domino tiling in such a way that the input condition of the domino problem relates to the initial configuration of the Turing machine. The restrictions on the size of the tiled rectangle correspond to the time and space restrictions of the Turing

machine. To prove that a problem A is NEXPTIME-hard, it suffices to show that $\mathrm{DOMINO}(\mathcal{D}, 2^n) \leq_p A$.

Our goal is to show that $\mathrm{DOMINO}(\mathcal{D}, 2^n)$ reduces to $Sat(X)$ for relatively simple classes $X \subseteq \mathrm{FO}$. Set

$$X = \{\varphi \in \mathrm{FO}^2 : \varphi = \forall x \forall y\, \alpha \wedge \forall x \exists y\, \beta, \text{ s.t. } \alpha, \beta \text{ quantifier-free, without } =, \text{ and with only monadic predicates}\}\,.$$

We show that $Sat(X)$ is NEXPTIME-complete and hence also $Sat(\mathrm{FO}^2)$ is NEXPTIME-complete.

Lemma 2.33. For each domino system $\mathcal{D} = (D, H, V)$ there exists a polynomial time reduction $w \in D^n \mapsto \psi_w \in X$ such that $\mathcal{D}$ tiles $Z(2^n)$ with initial condition w if and only if ψ_w is satisfiable.

Proof. The intended model of ψ_w is a description of a tiling $\tau : Z(2^n) \to D$ in the universe $Z(2^n)$.

Let $z = (a, b) \in Z(2^n)$ with $a = \sum_{i=0}^{n-1} a_i 2^i$ and $b = \sum_{i=0}^{n-1} b_i 2^i$. Encode the tuple as $(a_o, \ldots, a_{n-1}, b_0, \ldots, b_{n-1}) \in \{0,1\}^{2n}$.

To encode the tiling, we define ψ_w with the monadic predicates X_i, X_i^*, Y_i, Y_i^*, N_i for $0 \leq i < n$ and $P_d (d \in D)$ with the following intended meaning:

$$\begin{aligned}
X_i z \quad &\text{iff} \quad a_i = 1.\\
X_i^* z \quad &\text{iff} \quad a_j = 1 \text{ for all } j < i.\\
Y_i z \quad &\text{iff} \quad b_j = 1.\\
Y_i^* z \quad &\text{iff} \quad b_j = 1 \text{ for all } j < i.\\
N_i z \quad &\text{iff} \quad z = (i, 0).\\
P_d z \quad &\text{iff} \quad \tau(z) = d.
\end{aligned}$$

ψ_w will have the form $\psi_w = \forall x \forall y \alpha \wedge \forall x \exists y \beta$, where β accounts for the correct interpretation of $X_i, X_i^*, Y_i, Y_i^*, N_i$ and ensures that every element has a successor, and α accounts for the description of a correct tiling.

Now β is the the following formula:

$$\begin{aligned}
\beta = {} & X_0^* x \wedge Y_0^* x \\
\wedge \quad & \bigwedge_{i=1}^{n-1} X_i^* x \leftrightarrow (X_{i-1}^* x \wedge X_{i-1} x) \\
\wedge \quad & \bigwedge_{i=1}^{n-1} Y_i^* x \leftrightarrow (Y_{i-1}^* x \wedge Y_{i-1} x) \\
\wedge \quad & \bigwedge_{i=0}^{n-1} X_i y \leftrightarrow (X_i x \oplus X_i^* x) \\
\wedge \quad & \bigwedge_{i=0}^{n-1} Y_i y \leftrightarrow (Y_i x \oplus (Y_i^* x \wedge X_{n-1} x \wedge X_{n-1}^* x)) \\
\wedge \quad & N_0 x \leftrightarrow (\bigwedge_{i=0}^{n-1} \neg X_i x \wedge \neg Y_i x) \\
\wedge \quad & \bigwedge_{i=0}^{n-1} N_i x \leftrightarrow N_{i+1} y.
\end{aligned}$$

We define the following shorthands for use in α:

$$\begin{aligned}
H(x,y) \quad &:= \quad \bigwedge_{i=0}^{n-1} (Y_i y \leftrightarrow Y_i x) \wedge \bigwedge_{i=0}^{n-1} (X_i y \leftrightarrow (X_i x \oplus X_i^* x)) \\
V(x,y) \quad &:= \quad \bigwedge_{i=0}^{n-1} (X_i y \leftrightarrow X_i x) \wedge \bigwedge_{i=0}^{n-1} (Y_i y \leftrightarrow (Y_i x \oplus Y_i^* x)).
\end{aligned}$$

Now α is defined to be

$$\begin{aligned}
\alpha = {} & \bigwedge_{d \neq d'} \neg (P_d x \wedge P_{d'} x) \\
\wedge \quad & (H(x,y) \rightarrow \bigvee_{(d,d') \in H} (P_d x \wedge P_{d'} y)) \\
\wedge \quad & (V(x,y) \rightarrow \bigvee_{(d,d') \in V} (P_d x \wedge P_{d'} y)) \\
\wedge \quad & (\bigwedge_{i=i}^{n-1} (N_i x \rightarrow P_{w_i} x)).
\end{aligned}$$

Claim 2.34. ψ_w is satisfiable if and only if $\mathcal{D}$ tiles $Z(2^n)$ with initial condition w.

Proof. We show both directions.

($\Leftarrow$) Consider the intended model, ψ_w holds in it.
($\Rightarrow$) Consider $\mathfrak{C} = (C, X_1, \ldots) \models \psi_w$ and define a mapping

$$\begin{aligned}
f : \quad C \quad &\rightarrow Z(2^n) \\
c \quad &\mapsto (a,b) \equiv (a_0, \ldots, a_{n-1}, b_0, \ldots, b_{n-1})
\end{aligned}$$

$$\begin{aligned}
\text{with } a_i = 1 \quad &\text{iff} \quad \mathfrak{C} \models X_i c \quad \text{and} \\
b_i = 1 \quad &\text{iff} \quad \mathfrak{C} \models Y_i c.
\end{aligned}$$

As $\mathfrak{C} \models \forall x \exists y \beta$, f is surjective. Choose for each $z \in Z(2^n)$ an element $c \in f^{-1}(z)$ and set $\tau(z) = d$ for the unique d that satisfies $\mathfrak{C} \models P_d c$. Then τ is a correct tiling with initial condition w. Q.E.D.

Since the length of ψ_w is $|\psi_w| = O(n \log n)$, the above claim completes the proof of the lemma. Q.E.D.