Journal of Logic, Language, and Information **6:** 427–439, 1997.

Modal Logic over Finite Structures

ERIC ROSEN*
Mathematische Grundlagen der Informatik, RWTH, Aachen, Germany
E-mail: erosen@informatik.rwth-aachen.de

(Received 20 January 1997)

Abstract. We investigate properties of propositional modal logic over the class of finite structures. In particular, we show that certain known preservation theorems remain true over this class. We prove that a class of finite models is defined by a first-order sentence and closed under bisimulations if and only if it is definable by a modal formula. We also prove that a class of finite models defined by a modal formula is closed under extensions if and only if it is defined by a $\diamond$-modal formula.

Key words: modal logic, finite model theory, preservation theorems

1. Introduction

In this paper, we discuss the finite model theory of propositional modal logic, PM. Throughout, we restrict our attention to finite models, unless indicated otherwise. Modal logic has been studied extensively in connection with philosophical logic. More recently, connections have emerged between modal logic and computational linguistics and certain areas of computer science. Below we will be interested in the "classical model theory" of modal logic, an approach taken by van Benthem and others. For example, PM satisfies certain preservation theorems that are analogous to classical theorems for first-order logic, FO. We show that, in contrast to more expressive logics (e.g. see Gurevich, 1984) PM remains well behaved over the class $\mathcal{F}$ of finite structures, as various classical results remain true over this class.

Van Benthem (1976) proved the following preservation theorem: a class of models defined by a first-order sentence is closed under bisimulations iff it can be defined by a modal formula. Below we prove that this result remains true over $\mathcal{F}$. We then show that an "existential" preservation theorem, due to van Benthem and Visser (see Andréka et al., 1995), also holds over the class of finite structures. Finally, we give an alternative proof, which does not use the compactness theorem, of Andréka et al.'s result (1995) establishing the modal analog of the Craig interpolation theorem.

* This research was supported in part by NSF-STC SBR-8920230 while the author was a graduate student at the University of Pennsylvania.

2. Background

In order to make this paper self-contained, we briefly describe the syntax and semantics of PM. Most of this material is well known, and more detailed descriptions can be found in many places (e.g. see van Benthem, 1985). The syntax of PM is obtained from that of simple sentential logic by adding the two modal operators $\Box\varphi$, *necessarily* φ, and $\Diamond\varphi$, *possibly* φ. We also assume that there is a 0-ary connective $\top$, *truth*. Over a signature of *proposition symbols*, $\sigma = \{p_1, \dots, p_k\}$, the class of formulas of PM(σ) is the smallest class containing each atomic formula p_i and closed under negation, conjunction, disjunction, and the operators $\Box$ and $\Diamond$. We will always assume that the signature is finite. A (Kripke) model of PM(σ) is a directed graph A with additional unary predicates, $\{P_1, \dots, P_k\}$, corresponding to each proposition symbol. The edge relation Rxy is often called the "accessibility relation," and we will say that b is accessible from a just in case Rab. Below, we will use A, etc., to refer to both a model and to its universe. The intended interpretation should always be clear from the context.

DEFINITION 1. Satisfaction for formulas of PM at a node (or world) is defined inductively. Let A be a model and let $a \in A$.

1. $(A, a) \models^{PM} p_i$ iff $A \models P_i(a)$.
2. The Boolean operations are given their standard interpretations.
3. For the modal operator necessarily, $(A, a) \models^{PM} \Box\varphi$ iff for all $b \in A$ such that $A \models Rab, (A, b) \models^{PM} \varphi$. Possibly is defined dually, $(A, a) \models^{PM} \Diamond\varphi$ iff there is some $b \in A$ such that $A \models Rab$ and $(A, b) \models^{PM} \varphi$.

This semantics suggest a natural interpretation of PM into FO. In fact, by reusing variables we can translate PM into the language L^2, the set of FO-formulas that only contain two (reusable) variables, x_0 and x_1. Since formulas of PM are evaluated at a node of the Kripke model, they naturally translate into FO-formulas with one free variable. In order to keep the image of the translation in L^2, we will simultaneously define two functions, $\mu_0(\varphi)$ and $\mu_1(\varphi)$ such that (i) $\mu_d(\varphi)$ contains x_d free; and (ii) for all $\varphi \in$ PM, $\mu_1(\varphi)$ is obtainable from $\mu_0(\varphi)$ by replacing every occurrence of x_0 by x_1, and vice versa. The functions $\mu_d(\varphi)$ from formulas of PM to formulas of L^2 are defined inductively as follows:

$$\mu_d(p_j) = P_j(x_d)$$

$$\mu_d(\varphi_1 \wedge \varphi_2) = \mu_d(\varphi_1) \wedge \mu_d(\varphi_2)$$

$$\mu_d(\varphi_1 \vee \varphi_2) = \mu_d(\varphi_1) \vee \mu_d(\varphi_2)$$

$$\mu_d(\neg\varphi) = \neg\mu_d(\varphi)$$

$$\mu_d(\Box\varphi) = \forall x_{1-d}(Rx_d x_{1-d} \rightarrow \mu_{1-d}(\varphi))$$

$$\mu_d(\Diamond\varphi) = \exists x_{1-d}(Rx_d x_{1-d} \wedge \mu_{1-d}(\varphi))$$

To simplify the exposition, we add a single constant c to our FO-signature, to convert each formula with one free variable into a sentence. Let $\mu(\varphi)$ be the function from PM to L^2 such that for all $\varphi \in$ PM, $\mu(\varphi)$ is obtained from $\mu_0(\varphi)$ by replacing each *free* occurence of x_0 by c. Then each model is viewed as having a distinguished node, at which modal formula are evaluated. Let FO^M, the *modal fragment of first-order logic*, be the image of PM under the mapping $\mu(\varphi)$.

In his dissertation, van Benthem (1976) gave an algebraic characterization of classes defined by an FO sentence that are definable by a modal formula. He introduced the following important notion.

DEFINITION 2. Given two models A and B (with distinguished nodes c^A and c^B), a *bisimulation* between A and B, is a binary relation, $\sim$, contained in $A \times B$, such that

1. $c^A \sim c^B$
2. For all a, b such that $a \sim b$, if $A \models Raa'[B \models Rbb']$, then there is a $b' \in B[a' \in A]$ such that $a' \sim b'$ and $B \models Rbb'[A \models Raa']$.
3. For all a, b such that $a \sim b$, and all $P_j, A \models P_j(a)$ iff $B \models P_j(b)$.

We say that A bisimulates with B iff there is a bisimulation between the two models. We also write $(A, a) \sim (B, b)$ if there is a bisimulation $\sim$ between A and B such that $a \sim b$.

Bisimulation is an equivalence relation on structures, which can be seen as a modified, weak kind of partial isomorphism. It is easy to see that if there is a bisimulation between a pair of models, then they satisfy the same modal formulas.

We turn our attention now to modal definability.

DEFINITION 3. The quantifier rank of a formula, $qr(\varphi)$, is defined inductively.

1. $qr(P_i) = 0$
2. $qr(\neg\varphi) = qr(\varphi)$
3. $qr(\varphi_1 \wedge \varphi_2) = qr(\varphi_1 \vee \varphi_2) = \max(qr(\varphi_1), qr(\varphi_2))$
4. $qr(\Diamond\varphi) = qr(\Box\varphi) = qr(\varphi) + 1$.

Of course, there are no genuine quantifiers in PM; the choice of terminology emphasizes the connection between PM and FO. In particular, for all $\varphi \in$ PM, $qr(\varphi)$ equals the quantifier rank of the FO-sentence, $\mu(\varphi)$. Let PM^n be the set of formulas of quantifier rank $\leq n$. Given a model A, the PM^n-theory of A is then the set of formulas, of quantifier rank $\leq n$, satisfied by A.

LEMMA 1. *Let σ be a fixed signature.*

1. *For all n, up to logical equivalence, there are finitely many formulas of* PM^n.

2. *There is a recursive function $f(n)$ that generates a (finite) list of all formulas, up to logical equivalence, of quantifier rank $\leq n$.*
3. *For all A, the* PM^n*-theory of A is finitely axiomatizable. (That is, the theory is logically equivalent to a single formula.)*

Proof. We prove Part 1 by induction on n. The case $n = 0$ is obvious. For $n+1$, observe that every formula of quantifier rank $\leq n + 1$ is a Boolean combination of atomic formulas and formulas of the form $\Diamond\theta$, with $qr(\theta) \leq n$. Parts 2 and 3 follow easily from the proof of Part 1. □

The following notion was introduced in Hennessy and Milner (1985).

DEFINITION 4. We say that there is an *n-bisimulation* between A and B, written $A \sim_n B$, iff there is a sequence of relations $\approx_0, \ldots, \approx_n$, each on $A \times B$, such that:

1. $c^A \approx_0 c^B$
2. For all $m < n$, if $a \approx_m b$, and $A \models Raa'$ then there is a $b' \in B$ such that $B \models Rbb'$ and $a' \approx_{m+1} b'$[and vice-versa].
3. For all $m \leq n$, if $a \approx_m b$, then for all $P_j, A \models P_j(a)$ iff $B \models P_j(b)$.

Intuitively, $A \sim_n B$ means that A and B bisimulate "up to depth n." Observe that $A \sim B$ implies $A \sim_n B$, for all n, and that $\sim_n$ also defines an equivalence relation on any class of structures. The next result can be established by a straightforward modification of the proof of the algebraic characterization of logical equivalence for FO, due to Ehrenfeucht and Fraïssé (e.g. see Hodges, 1993), or for finite variable logic, due to Barwise (1977).

PROPOSITION 1. *For all n, and A and B over some σ, the following conditions are equivalent:*

1. *There is an n-bisimulation between A and B.*
2. *For all modal formulas θ of quantifier rank $\leq n$, $A \models \theta$ iff $B \models \theta$.*

The next proposition follows easily from Proposition 1 and Lemma 1.

PROPOSITION 2. *Let $\mathcal{C}$ be any class of models, closed under isomorphism. Let $\mathcal{C}'$ be any subclass of $\mathcal{C}$, also closed under isomorphism. Then, for all n, the following conditions are equivalent:*

1. *For all $A \in \mathcal{C}', B \in \mathcal{C} - \mathcal{C}', A \not\sim_n B$.*
2. *There is a modal formula of quantifier rank $\leq n$ that defines the class $\mathcal{C}'$ over $\mathcal{C}$.*

Bisimulation and n-bisimulation are rather weak equivalence relations, in the sense that they determine relatively large equivalence classes. In other words, for every model A there are many other models with the same modal theory. Our proofs will exploit this feature repeatedly.

We fix the following terminology.

DEFINITION 5. The children of a in A are those b such that $A \models Rab$. We say that b is a descendent of a iff there is a directed path from a to b. For all n, b is an n-descendent of a if there is a path of length $\leq n$ from a to b. The family of a, written F^a is the submodel of A with universe $\{a\} \cup \{b \mid b \text{ is a descendent of } a\}$. For all a and b, we say that a and b are disjoint iff $F^a \cap F^b = \emptyset$.

The r-neighborhood of a point a, denoted $N_r(a)$, is defined inductively. $N_0(a)$ is the submodel of A with universe $\{a\}$. For all $r+1$, $b \in N_{r+1}(a)$ iff $b \in N_r(a)$ or there is an $a' \in N_r(a)$ such that $A \models Ra'b \vee Rba'$. An r-tree is a directed tree rooted at c of height $\leq r$. An r-pseudotree is a model such that $N_r(c)$ is a tree such that all distinct pairs of its leaves are disjoint, as defined above.

We now describe certain operations on models that produce either bisimilar or n-bisimilar models.

DEFINITION 6. Let A be a model and let $a \in A$. We say that A' *is obtained from* A *by adding a copy of the family of* a iff the following conditions hold.

1. The universe of A' is the disjoint union of A and $\widehat{F}^a$, a “copy” of F^a.
2. A' is an extension of A and the submodel $\widehat{F}^a \subseteq A'$ is isomorphic to F^a.
3. For all $b_0 \in A - F^a$ and $\widehat{b}_1 \in \widehat{F}^a$, $A' \models Rb_0\widehat{b}_1[R\widehat{b}_1b_0]$ iff $A \models Rb_0b_1[Rb_1b_0]$, where $\widehat{b}_1$ is the copy of b_1 in $\widehat{F}^a$.
4. For all $b \in F^a$ and $\widehat{b} \in \widehat{F}^a$, $A' \models \neg Rb\widehat{b} \wedge \neg R\widehat{b}b$.

Observe that the binary relation $\{(b, b') \mid b \in A, b' \in A' \text{ and } b = b' \text{ or } b' \text{ is a copy of } b\}$ witnesses that $A \sim A'$.

Another concept, introduced by Sahlqvist (1975), is that of *unravelling* a structure to produce a new structure with which it bisimulates. Before defining this notion, we give a simple illustration. Let A be the graph on one vertex with a loop, and let A' be the directed chain on $c = 0, 1, \ldots, n$ such that for all $m < n$, $A' \models Rm, m+1$ and $A' \models Rnn$. We can view A' as having been obtained from A by unravelling, or unwinding, the loop n times. The set $A \times A'$ is itself a bisimulation between A and A'. In general, any model A can be n-unravelled, so that the n-descendents of c form an n-tree. By ω-unravelling F^c in A we obtain a (possibly infinite) tree. Every unravelling of A bisimulates with A.

DEFINITION 7. For all models A and all $n \in \omega$, we define the n-*unravelling* of A, an n-pseudotree denoted A', as follows. We first describe the $(n-1)$-tree portion of A', which we will call A_0. The universe of A_0 is the set of all paths in A that begin at c^A and have length $\leq n-1$. That is, $A_0 = \{\overline{a} \mid \overline{a} = (c = a_0, \ldots, a_s), s \leq n-1, a_i \in A$ for all $i \leq s$ and for all $i < s$, $A \models Ra_i a_{i+1}\}$. For each such $\overline{a}, A_0 \models P_j(\overline{a})$ iff $A \models P_j(a_s)$. Given a path $\overline{a} \in A_0$ and an element $a' \in A$, let $\overline{a} * a'$ denote their concatenation, that is, the sequence $(a_0, a_1, \ldots, a_s, a')$. In A_0, there is an R edge from $\overline{a}$ to $\overline{a}_1$ iff $\overline{a}_1 = \overline{a} * a'$, for some $a' \in A$. This completes the description of A_0. We define $c^{A'}$ to be the root of this tree.

The universe of A' is the disjoint union, $A_0 \cup \{(\overline{a}, b) \mid \overline{a}$ is a path of length n rooted at c and $b \in F^{a_n}$, where a_n is the last element of $\overline{a}\}$. That is, the elements of $A' - A_0$ are ordered pairs. A' is an extension of A_0, and for all $(\overline{a}, b) \in A'$ and all $P_i(x), A' \models P_i((\overline{a}, b))$ iff $A \models P_i(b)$. All that remains is to complete the definition of the relation R on A'. For $(\overline{a}, b), (\overline{a}', b') \in A' - A_0$, $A' \models R((\overline{a}, b), (\overline{a}', b'))$ iff $\overline{a} = \overline{a}'$ and $A \models R(b, b')$. For $(\overline{a}, b) \in A' - A_0$ and $\overline{a}' \in A_0, A' \models \neg R((\overline{a}, b), \overline{a}')$ and also $A' \models R(\overline{a}', (\overline{a}, b))$ iff $\overline{a} = \overline{a}' * b$.

For later use we collect some facts that are easy to verify.

PROPOSITION 3. *For all A, (1) $A \sim F^c$. (2) A bisimulates with a tree rooted at c, its ω-unravelling. (3) A bisimulates with an n-pseudotree, its n-unravelling. (4) A n-bisimulates with an n-tree, a submodel of its n-unravelling. (5) Over a fixed signature σ, there is a recursive function $f(x)$ such that for all modal formulas φ of quantifier rank $\leq n$, if φ is satisfiable, by a finite or infinite model, then it is satisfiable by an n-tree of cardinality $\leq f(n)$. (6) For all finite A, the modal theory of A is finitely axiomatizable iff F^c is acyclic.*

Proof. We provide proofs of Facts 5 and 6. From Fact 4 and Proposition 1, it is clear that for all $\varphi \in \mathrm{PM}^n$, φ is satisfiable iff it is satisfied by an n-tree. Given a fixed finite signature σ, we now define an effective procedure that maps each natural number n into a finite set of n-trees $\mathcal{T}^n$ such that for all $\varphi \in \mathrm{PM}(\sigma)$ of quantifier rank $\leq n$, if φ is satisfiable, then it is satisfied in some $A \in \mathcal{T}^n$. This will suffice to establish the claim. The sets $\mathcal{T}^n$ are defined inductively. $\mathcal{T}^0$ contains every model, up to isomorphism, with exactly one element, and has cardinality $= 2^{|\sigma|}$. For $n+1, A \in \mathcal{T}^{n+1}$ iff $A \in \mathcal{T}^n$ or A is an $n+1$-tree rooted at c with children $a_1, \ldots, a_k$ satisfying the following properties: (i) for all $i \leq k$, the family F^{a_i} is isomorphic to some tree $B \in \mathcal{T}^n$; and (ii) for all $i \neq j \leq k, F^{a_i} \not\cong F^{a_j}$. It is easy to verify both that there is a recursive bound on the size of models in each $\mathcal{T}^n$ and that every n-tree bisimulates with an n-tree in $\mathcal{T}^n$. This establishes Fact 5.

We now prove Fact 6. Suppose that F^c is acylic. We show, by induction on the "height" n of F^c, that A is axiomatized by a formula of quantifier rank $= n+1$. (Here, we define the *height* to be the length of the longest path that begins at c.) For $n = 0$, let $\theta = (\bigwedge_{P \in \tau} P \wedge \bigwedge_{Q \in \sigma - \tau} \neg Q) \wedge (\neg\Diamond P' \wedge \Box P')$, where τ is the set of proposition symbols satisfied at c, and P' is any proposition symbol in σ. For $n \geq 1$, and each child a_i of c, by the induction hypothesis let θ_i be a formula that axiomatizes the family F^{a_i}. Then let $\theta = (\bigwedge_{P \in \tau} P \wedge \bigwedge_{Q \in \sigma - \tau} \neg Q) \wedge (\bigwedge_i \Diamond \theta_i) \wedge (\Box \bigvee_i \theta_i)$. It is clear that θ axiomatizes the modal theory of A. In the other direction, let A be such that F^c contains a cycle, and let θ be a modal formula of quantifier rank n. Let B be an n-tree that verifies θ. It is easy to show that there is a modal formula, ψ, of quantifier rank $= n+1$ true in A but not in B. For example, for any $P \in \sigma$, let $\psi = \Diamond(\ldots \Diamond(P \vee \neg P) \ldots)$ contain a string of $n+1$ $\Diamond$'s. Therefore the modal theory of A is not axiomatized by any formula of quantifier rank n, and hence is not finitely axiomatizable. $\Box$

Observe that Fact 5 implies some well-known results. One, a modal formula is satisfiable iff it is satisfiable by a finite Kripke model. Two, it is decidable whether a formula is satisfiable, both over the class of all structures and over $\mathcal{F}$.

3. Preservation Theorems

In this section, we show that two modal preservation theorems remain valid over the class $\mathcal{F}$. The arguments do not use finiteness in any essential way; therefore they also give alternative proofs of the theorems in the general case that do not rely on the Compactness theorem. Finally, we show how these methods can be used to reprove the modal version of the Craig interpolation theorem without employing compactness.

DEFINITION 8. Let $A \equiv^n B$ mean that for all $\varphi \in$ FO, with qr$(\varphi) \leq n$, $A \models \varphi$ iff $B \models \varphi$.

PROPOSITION 4. *The bisimulation preservation theorem for modal formulas remains true in the finite case. That is, a class $\mathcal{C}$ is defined by an FO sentence and closed under bisimulations iff it is definable by a modal formula.*

Proof. We only prove the non-trivial direction. Let $\mathcal{C}$ be defined by an FO sentence and closed under bisimulations. Suppose that $\mathcal{C}$ is not definable by a modal formula. By Proposition 2, this implies that for all n, there are $A \in \mathcal{C}$ and $B \notin \mathcal{C}$ such that $A \sim_n B$. (Of course, since $\mathcal{C}$ is closed under bisimulations, we have that $A \not\sim B$.) We will show that this condition implies that for all n, there are actually $A \in \mathcal{C}$ and $B \notin \mathcal{C}$ such that $A \equiv^n B$. This immediately implies that $\mathcal{C}$ is not defined by any FO sentence, a contradiction.

More specifically, we show that there is a function $l(x)$ such that, for all n, if $A \sim_{l(n)} B$, then there are A' and B' such that $A \sim A', B \sim B'$ and $A' \equiv^n B'$. By choosing $A \in \mathcal{C}$ and $B \notin \mathcal{C}$, we get $A' \in \mathcal{C}$ and $B' \notin \mathcal{C}$. Given A and B, we find A' and B' by modifying A and B in a sequence of steps, as described in the following lemmas.

LEMMA 2. *Let A and B be such that $A \sim_t B$. Then there are t-pseudotrees A' and B' such that $A \sim A', B \sim B'$, and $A' \sim_t B'$.*

Let A' and B' be the t-unravellings of A and B. Then A' and B' are t-pseudotrees such that $A \sim A'$ and $B \sim B'$. By the transitivity of $\sim_t$, this implies that $A' \sim_t B'$.

LEMMA 3. *Let A and B be t-pseudotrees such that $A \sim_t B$. Then there are t-pseudotrees A' and B' such that $A \sim A'$, $B \sim B'$, and $N_t(c^{A'}) \cong N_t(c^{B'})$.*

The proof describes an algorithm for modifying the two models in a sequence of steps that yields models with isomorphic t-neighborhoods of c. After each step $s, s \leq t$, we have models A_s and B_s such that $A \sim A_s$ and $B \sim B_s$, and c^{A_s} and c^{B_s} have isomorphic s-neighborhoods. At each step $s+1$, A_{s+1} [resp. B_{s+1}] is obtained from A_s by adding copies of families of nodes of distance $s+1$ from c.

Let $\{a_1, \ldots a_l, b_1, \ldots, b_m\}$ be the set of the children of c in A and B. The relation $\sim_{t-1}$ induces an equivalence relation on this set such that each equivalence class has at least one member in each of A and B. To obtain A_1 and B_1 with isomorphic 1-neighborhoods of c that bisimulate with A and B, it suffices to add enough copies of families of the c-children a_i and b_j such that each equivalence class has an equal number of members in A_1 and B_1. For example, renumbering the indices of c-children if necessary, suppose that $\{a_1, \ldots, a_i; b_1, \ldots b_j\}$ is one such equivalence class. Also, without loss of generality, assume that $i \leq j$. Then A_1 will contain $j-i$ additional copies of the family F^{a_i}. Let $g_1(x)$ be a bijection between the c-children in A_1 and B_1 such that for all a_i, $(A_1, a_i) \sim_{t-1} (B_1, g_1(a_i))$. By iterating this procedure, at each step $s+1$, we obtain A_{s+1} and B_{s+1}, and a bijection g_{s+1} between nodes of distance $s+1$ from c^A and c^B with the following properties. For all nodes a_i in A_s of distance s from c, the bijection g_{s+1} maps the children of a_i to those of $g_s(a_i)$, and for all $a \in dom(g_{s+1}), (A_{s+1}, a) \sim_{t-(s+1)} (B_{s+1}, g_{s+1}(a))$. Finally, we choose A' and B' to be the models A_t and B_t.

Together, these lemmas establish that there are models $A \in \mathcal{C}$ and $B \notin \mathcal{C}$ that look rather similar. In particular, for all t, there are t-pseudotrees $A \in \mathcal{C}$ and $B \notin \mathcal{C}$ such that $N_t(c^A) \cong N_t(c^B)$. Although these models have isomorphic t-neighborhoods of c, we still know nothing about the other part of each model, which might make A and B “look very different” in FO. The final step of the proof takes care of this by using a version of Hanf’s lemma due to Fagin et al. (1995).

PROPOSITION 5 (Fagin et al., 1995). *For each signature σ, there is a function $f(x)$ with the following property. For all n, A and B, if there is a bijection $h : A \to B$ such that for all $a \in A$, $N_{f(n)}(a) \cong N_{f(n)}(h(a))$, (with a and $h(a)$ distinguished), then $A \equiv^n B$.*

LEMMA 4. *Let A and B be $(2f(n))$-pseudotrees with $N_{2f(n)}(c^A) \cong N_{2f(n)}(c^B)$, where $f(x)$ is the Hanf function. Then there are A' and B' such that $A \sim A'$, $B \sim B'$, and $A' \equiv^n B'$.*

Each of A' and B' will be obtained from A and B, respectively, by extending the original model by adding disjoint components in such a way that it will be obvious that A' and B' possess the same $f(n)$-neighborhoods. It is clear that extending models in this way does not affect bisimulations. We define A' [B'] to be the disjoint union of A and B [B and A], with the constant c still interpreted as the root of the A [B] component. Thus, A' and B' are identical except for the interpretation of c.

Let $g(x)$ be an isomorphism from $N_{2f(n)}(c^A)$ and $N_{2f(n)}(c^B)$. Let $h(x)$ be the following bijection from A' to B'. For all $a \in A'$, if $a \in N_{f(n)}(c^A)$ [$a \in N_{f(n)}(c^B)$], then $h(a) = g(a)$ [$h(a) = g^{-1}(a)$]. That is, $h(x)$ maps all $a \in A'$ whose distance to the connected root is $\leq n$ to the opposite component in B'. For all other $a \in A', h(a) = a$. It is easy to see that h is a bijection that preserves $f(n)$-neighborhoods. Thus $A' \equiv^n B'$ as desired.

To complete the proof, all that remains is to combine the above results. Suppose that $\mathcal{C}$ is defined by an FO sentence and closed under bisimulations, but not definable by a modal formula. Then by Lemmas 2, 3, and 4, for all n, there are $A \in \mathcal{C}$ and $B \notin \mathcal{C}$ such that $A \equiv^n B$. But this implies that $\mathcal{C}$ is not defined by an FO sentence, a contradiction. This proves the proposition. □

The next preservation theorem that we consider characterizes those formulas whose classes of models are closed under extensions. Before stating the main result, we define some terminology and prove a few preliminary lemmas.

DEFINITION 9.

1. A $\diamond$-formula is a modal formula built up from atomic propositions and negated atomic propositions using $\wedge$, $\vee$, and $\diamond$.
2. For all A and B, we write $A \rightsquigarrow_\diamond B$ iff for all $\diamond$-formula φ, if $A \models \varphi$, then $B \models \varphi$.
3. Given a model A, the $\diamond$-theory of A is the set of $\diamond$-formulas satisfied by A.

Observe that the $\diamond$-formulas are precisely those $\varphi \in$ PM such that $\mu(\varphi)$ is an existential FO sentence. In particular, the class of models defined by any $\diamond$-formula is closed under extensions.

LEMMA 5. *Let A be an n-tree, rooted at c.*

1. *For all $\diamond$-formulas, φ, of quantifier rank $\geq n+1$, $A \not\models \varphi$.*
2. *The $\diamond$-theory of A is axiomatized by a formula of quantifier rank $= n$.*

Proof. Part 1 is obvious, since A does not contain any paths of length $n+1$. By Lemma 1, let $\theta_1, \ldots \theta_k$ be the set of all $\diamond$-formulas of quantifier rank $\leq n$, up to equivalence, satisfied in A. By Part 1, it is clear that $\theta = \bigwedge \theta_i$ axiomatizes the $\diamond$-theory of A. □

LEMMA 6. *Given a fixed signature, there is a finite set of n-trees, $\mathcal{T}^n = \{B_1, \ldots, B_v\}$ such that for all A, there is a $u \leq v$ such that $A \sim_n B_u$. Furthermore, $\mathcal{T}^n$ can be obtained effectively.*

Proof. This result follows easily from Fact 5 of Proposition 3. Let $\mathcal{T}^n$ be the same set that was defined in the proof of this Fact, such that every satisfiable formula φ of quantifier rank $\leq n$ is satisfied by some $B \in \mathcal{T}^n$. Let A be any model, and let $\theta_n \in \text{PM}^n$ axiomatize its PM^n-theory, again using Lemma 1. By Fact 5, there is a $B \in \mathcal{T}^n$ such that $B \models \theta_n$. This now implies that $A \sim_n B$. □

The next result can be viewed as the modal version of the Łos–Tarski theorem for finite structures. We use $\text{Mod}_f(\varphi)$ [$\text{Mod}(\varphi)$] to denote the class of finite [all] models of φ. EXT is the set of classes of finite models that are closed under extensions.

PROPOSITION 6. *The existential preservation theorem for modal logic remains true over $\mathcal{F}$. That is, for all φ, if* $\text{Mod}_f(\varphi) \in \text{EXT}$, *then φ is equivalent to a $\Diamond$-formula θ. Moreover, there is an effective procedure for finding the equivalent $\Diamond$-formula.*

Proof. Let $\mathcal{C} \in \text{EXT}$ be defined by some modal formula φ, with quantifier rank n. Let $\mathcal{C}^n = \mathcal{C} \cap \mathcal{T}^n = \{D_1, \ldots, D_k\}$. For each $D_i, i \leq k$, let θ_i axiomatize the $\Diamond$-theory of D_i. By Lemma 5, $qr(\theta_i) \leq n$. Let $\theta = \bigvee_{i \leq k} \theta_i$. We claim that φ is equivalent to θ.

First we show that φ implies θ. Suppose that $A \models \varphi$. We claim that there is a $D \in \mathcal{C}^n$ such that $A \sim_n D$. By Lemma 6, there is a $B \in \mathcal{T}^n$ such that $A \sim_n B$. Since $\mathcal{C}$ is closed under $\sim_n$-equivalence, B must actually be in $\mathcal{C}$, and hence in $\mathcal{C}^n$. Let $D = B$. There is some θ_i, as defined above, such that $D \models \theta_i$. Since $qr(\theta_i) \leq n$, this implies that $A \models \theta_i$, and hence $A \models \theta$.

Now we prove the opposite direction, θ implies φ. Suppose that $A \models \theta$. Then $A \models \theta_i$, for some $i \leq k$. By Lemma 6, there is a $B \in \mathcal{T}^n$ such that $A \sim_n B$. Observe that $D_i \rightsquigarrow_\Diamond B$. We want to show that there is an A' such that (i) $B \sim A'$, and hence $A \sim_n A'$; and (ii) $D_i \subseteq A'$. As $D_i \in \mathcal{C}$, and $\mathcal{C} \in \text{EXT}$, (i) and (ii) imply that $A' \in \mathcal{C}$. Since $\mathcal{C}$ is closed under $\sim_n$-equivalence, $A \in \mathcal{C}$, as desired. Thus, it suffices to establish the following lemma.

LEMMA 7. *Let B, D be trees such that $D \rightsquigarrow_\Diamond B$. Then there is a m-tree A', $m =$ height of B, such that $B \sim A'$ and $D \subseteq A'$.*

By induction, on the height n of D. For $n = 0$, it is obvious that $D \subseteq B$, since D is just the single node c^D, and for all predicate symbols p, $D \models p$ iff $B \models p$. Let $A' = B$.

Consider $n > 0$. Let $\{d_1, \ldots, d_s\}$ and $\{b_1, \ldots b_t\}$ be the children of c^D and c^B, respectively. We claim that for each d_p, there is a b_r such that $F^{d_p} \rightsquigarrow_\Diamond F^{b_r}$. Let ψ, with $\text{qr}(\psi) \leq n$, axiomatize the $\Diamond$-theory of F^{d_p}. Then $D \models \Diamond\psi$, and therefore $B \models \Diamond\psi$. Thus there is a b_r such that $F^{b_r} \models \psi$, as desired.

By adding extra copies of families of the children of c^B to B, if necessary, we get B^0 such that $B \sim B^0$ and there is an injection $h : \{d_1, \ldots d_s\} \longrightarrow \{b_1^0, \ldots, b_{t'}^0\}$, $b_j^0 \in B^0$, such that $F^{d_i} \rightsquigarrow_\Diamond F^{h(d_i)}$. By the induction hypothesis, each such $F^{h(d_i)}$

bisimulates with an $(n-1)$-tree, $T^{h(d_i)}$, such that $F^{d_i} \subseteq T^{h(d_i)}$. Let A' be obtained from B^0 by replacing each subtree $F^{h(d_i)} \subseteq B^0$, with the tree $T^{h(d_i)}$. It is easy to see that $B \sim A'$ and $D \subseteq A'$.

This also completes the proof of the proposition. □

COROLLARY 1. *For every formula φ, there is a decision procedure that determines whether* $\mathrm{Mod}_f(\varphi)$ $[\mathrm{Mod}(\varphi)]$ *is closed under extensions. Therefore the set of formulas that defines such classes is recursive.*

Proof. By the proof of the previous proposition, if $\mathrm{Mod}_f(\varphi) \in \mathrm{EXT}$, then it is equivalent to a $\Diamond$-formula of quantifier rank $\leq qr(\varphi)$. By Lemma 1, one can effectively list, up to logical equivalence, all such formulas, $\psi_1, \ldots, \psi_l$. Then it suffices to test the validity of each formula, $\varphi \leftrightarrow \psi_i$, which is decidable. □

We now turn to an interpolation theorem, due to Andréka et al. (1995). It will be convenient to introduce briefly a fragment of second-order propositional modal logic, which allows quantification over propositions. We often use $\overline{P}$, etc., as shorthand for sequences, $(P_1, \ldots, P_n)$. We write $\psi(\overline{P})$ to indicate that the set of proposition symbols that occur in ψ equals $\overline{P}$. Also, by $\exists\overline{P}\psi(\overline{P}, \overline{Q})$ we mean the formula $\exists P_1 \ldots \exists P_n \psi(\overline{P}, \overline{Q})$.

DEFINITION 10. Let $\varphi(\overline{P}, \overline{Q})$ be a formula of PM, such that $\overline{P} \cap \overline{Q} = \emptyset$. Then $\exists\overline{Q}\varphi(\overline{P}, \overline{Q})$ is a Σ^1_1 modal formula; for all A, with signature $\sigma = \overline{P}$, $A \models \exists\overline{Q}\varphi(\overline{P}, \overline{Q})$ iff there is a B, an expansion of A with signature $\tau = \overline{P} \cup \overline{Q}$, such that $B \models \varphi(\overline{P}, \overline{Q})$. Π^1_1 modal formulas, of the form $\forall\overline{Q}\varphi(\overline{P}, \overline{Q})$, are defined similarly.

For all A, B, and n, we write $A \sim^{\overline{P}}_n B$ iff for all formulas $\varphi, qr(\varphi) \leq n$, that only contain proposition symbols from $\overline{P}$, $A \models \varphi$ iff $B \models \varphi$. Recall that every satisfiable modal formula is satisifed by a finite model; hence φ implies θ over the class of all models iff φ implies θ over $\mathcal{F}$. By this fact, the truth of the interpolation theorem in the general case immediately yields its truth over $\mathcal{F}$.

PROPOSITION 7 (Andréka et al., 1995). *Let φ and θ be formulas, with signatures σ_φ and σ_θ. If φ implies θ (over $\mathcal{F}$), then there is a formula ψ, with signature $\sigma_\psi \subseteq \sigma_\varphi \cap \sigma_\theta$, such that φ implies ψ and ψ implies θ. Furthermore, $qr(\psi) \leq$* $\max(qr(\varphi), qr(\theta))$.

Proof. Suppose that $\varphi(\overline{P}, \overline{Q})$ implies $\theta(\overline{P}, \overline{R})$, where $\overline{P}, \overline{Q}$, and $\overline{R}$ are pairwise disjoint sequences of propositions symbols. Equivalently, $\exists\overline{Q}\varphi(\overline{P}, \overline{Q})$ implies $\forall\overline{R}\theta(\overline{P}, \overline{R})$. If $\overline{P} = \emptyset$, then either $\varphi(\overline{P}, \overline{Q})$ must be unsatisfiable, or $\theta(\overline{P}, \overline{R})$ must be valid. In the first case, we can let $\psi = \neg\top$; in the second case, let $\psi = \top$. Thus, we consider models over the signature $\sigma = \overline{P}$, $\overline{P}$ non-empty.

Let $n = \max(qr(\varphi), qr(\theta))$. Recall that, by Lemma 1 or 6, there are only finitely many $\sim^{\overline{P}}_n$ equivalence classes. We claim that it suffices to show that for any $\sim^{\overline{P}}_n$

class $\mathcal{C}$, if there is an $A \in \mathcal{C}$ such that $A \models \exists\overline{Q}\varphi(\overline{P},\overline{Q})$, then for all $B \in \mathcal{C}$, $B \models \forall\overline{R}\theta(\overline{P},\overline{R})$. If this is true, for each $\sim_n^{\overline{P}}$ class $\mathcal{C}$ containing an A that satisfies $\exists\overline{Q}\varphi(\overline{P},\overline{Q})$, let ψ_i be a formula with signature $\overline{P}$, $qr(\psi_i) \leq n$, that defines the class. Then $\psi = \bigvee \psi_i$ is an interpolant.

Suppose, towards a contradiction, that there are A and B such that $A \sim_n^{\overline{P}} B$, $A \models \exists\overline{Q}\varphi(\overline{P},\overline{Q})$ and $B \models \exists\overline{R}\neg\theta(\overline{P},\overline{R})$. Let A' and B' be expansions of A and B such that $A' \models \varphi(\overline{P},\overline{Q})$ and $B' \models \neg\theta(\overline{P},\overline{R})$. By Lemma 6, there are n-trees A'' and B'' that are $\sim_n$-equivalent to A' and B', respectively. Finally, let A_1 and B_1 be the σ-reducts of A'' and B''. It is clear that $A_1 \models \exists\overline{Q}\varphi(\overline{P},\overline{Q})$ and $B_1 \models \exists\overline{R}\neg\theta(\overline{P},\overline{R})$. We now want to find a D such that $D \models \exists\overline{Q}\varphi(\overline{P},\overline{Q}) \wedge \exists\overline{R}\neg\theta(\overline{P},\overline{R})$. This will establish the contradiction.

D is constructed by extending A_1 and B_1 "simultaneously" by iteratively adding copies of families of elements. First we show that for any model M, if M' is obtained from M by adding a copy of a family F^m, for any $m \in M$, then every Σ_1^1 formula satisfied in M is also satisfied in M'. Suppose that $M \models \exists\overline{P}\psi(\overline{P},\overline{Q})$. Let N be an expansion of M that verifies the (first-order) modal formula $\psi(\overline{P},\overline{Q})$; and let N' be obtained from N by adding a copy of the family of m. It is clear that $N \sim N'$; thus $N' \models \psi(\overline{P},\overline{Q})$. Since N' is an expansion of M', $M' \models \exists\overline{P}\psi(\overline{P},\overline{Q})$, as desired.

We now describe the construction of D. As in the proof of Lemma 3, $\sim_{n-1}$ induces an equivalence relation on the set of children of c^{A_1} and c^{B_1} such that every equivalence class has at least one member in each model. Let A_2 and B_2 be obtained from A_1 and B_1 by adding enough copies of families of these children so that there is a bijection $g_1(x)$ from the children of c^{A_2} to those of c^{B_2} such that for all a_i, $F^{a_i} \sim_{n-1} F^{g_1(a_i)}$. Observe that $N_1(c^{A_2}) \cong N_1(c^{B_2})$. Repeat this procedure at each level $m \leq n$ of the trees, on pairs of subtrees in A_m and B_m determined by the bijection $g_{m-1}(x)$ at the previous level. By the argument of the preceding paragraph, for all m, $A_m \models \exists\overline{Q}\varphi(\overline{P},\overline{Q})$ and $B_m \models \exists\overline{R}\neg\theta(\overline{P},\overline{R})$. Furthermore, $N_m(c^{A_{m+1}}) \cong N_m(c^{B_{m+1}})$ This construction yields trees A_{n+1} and B_{n+1} such that $A_1 \sim A_{n+1}$, $B_1 \sim B_{n+1}$, and $A_{n+1} \cong B_{n+1}$. Let $D = A_{n+1}$. $\square$

4. Conclusion

In this paper, we have begun investigating the finite model theory of modal logic. Our results indicate that modal logic remains "well behaved" over the class of finite structures. In contrast, it is well known that most results from classical model theory, including various preservation theorems, become false when relativized to the class of finite structures. One way to extend this work would be to prove that other theorems of modal logic remain true over $\mathcal{F}$. Another line of research involves investigating the behavior, over $\mathcal{F}$, of somewhat stronger fragments of FO, e.g. the bounded quantifier fragments from (Andréka et al., 1995).

Acknowledgements

I would like to thank Prof. Johan van Benthem for introducing me to this subject and to his work, and for many helpful suggestions. I am also extremely grateful to Scott Weinstein for his insight, encouragement, and guidance, and to Marco Hollenberg, Steven Lindell, and an anonymous referee for detailed comments on earlier drafts of this paper. This paper constitutes a revised version of a chapter of my dissertation.

References

Andréka, H., van Benthem, J., and Németi, I., 1995, "Back and forth between modal logic and classical logic," *Journal of the IGPL* **3**(5), 685–720.

Barwise, J., 1977, "On Moschovakis closure ordinals," *Journal of Symbolic Logic* **42**, 292–296.

van Benthem, J., 1976, "Modal correspondence theory," Ph.D. Thesis, University of Amsterdam.

van Benthem, J., 1985, *Modal Logic and Classical Logic*, Naples: Bibliopolis.

Fagin, R., Stockmeyer, L., and Vardi, M.Y., 1995, "On monadic NP vs monadic co-NP," *Information and Computation* **120**, 78–92.

Gurevich, Y., 1984, "Toward logic tailored for computational complexity," pp. 175–216 in *Computation and Proof Theory*, M. Richter et al., eds., Berlin: Springer-Verlag.

Hennessy, M. and Milner, R., 1985, "Algebraic laws for nondeterminism and concurrency," *Journal of the Association for Computing Machinery* **32**, 137–161.

Hodges, W., 1993, *Model Theory*, Cambridge: Cambridge University Press.

Sahlqvist, H., 1975, "Completeness and correspondence in the first and second order semantics for modal logic," pp. 110–143 in *Proceedings of the Third Scandinavian Logic Symposium*, S. Kanger, ed., Amsterdam: North-Holland.