Lindström theorems for fragments of first-order logic
(full version with appendices)

Balder ten Cate*
ISLA, Univ. of Amsterdam
balder.tencate@uva.nl

Johan van Benthem
ILLC, Univ. of Amsterdam
johan@science.uva.nl

Jouko Väänänen
ILLC, Univ. of Amsterdam
vaananen@science.uva.nl

Abstract

Lindström theorems characterize logics in terms of model-theoretic conditions such as Compactness and the Löwenheim-Skolem property. Most existing Lindström theorems concern extensions of first-order logic. On the other hand, many logics relevant to computer science are fragments or extensions of fragments of first-order logic, e.g., k-variable logics and various modal logics. Finding Lindström theorems for these languages can be challenging, as most known techniques rely on coding arguments that seem to require the full expressive power of first-order logic.

In this paper, we provide Lindström characterizations for a number of fragments of first-order logic. These include the k-variable fragments for $k > 2$, Tarski's relation algebra, graded modal logic, and the binary guarded fragment. We use two different proof techniques. One is a modification of the original Lindström proof. The other involves the modal concepts of bisimulation, tree unraveling, and finite depth. Our results also imply semantic preservation theorems.

Characterizing the 2-variable fragment or the full guarded fragment remain open problems.

1 Introduction

There are many ways to capture the expressive power of a logical language $\mathcal{L}$. For instance, one can characterize $\mathcal{L}$ as being a model-theoretically well behaved fragment of a richer language $\mathcal{L}'$ (a *preservation theorem*), or as being maximally expressive while satisfying certain model-theoretic properties (a *Lindström theorem*). The main contribution of this paper is a series of Lindström theorems for fragments of first-order logic. We also show connections between our Lindström theorems and preservation theorems.

The original Lindström theorem for first-order logic, in one of its formulations, says the following:

The (first-order) Lindström Theorem [8] *An extension of first-order logic satisfies Compactness and the Löwenheim-Skolem property iff it is no more expressive than first-order logic.*

*The first author is supported by NWO research grant 639.021.508.

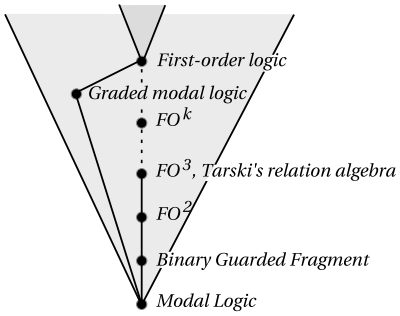

Figure 1. Fragments of first-order logic

There are several other versions of the theorem, characterizing first-order logic for instance in terms of Compactness and invariance for potential isomorphisms. Analogues of this result have been obtained for various extensions of first-order logic. On the other hand, few Lindström theorems are known for *fragments* of first-order logic. One notable example is Van Benthem's recent Lindström theorem for modal logic:

The modal Lindström theorem [14] *An extension of basic modal logic satisfies Compactness and bisimulation invariance iff it is no more expressive than basic modal logic.*[1]

Our general motivation for considering fragments comes from computer science logic. Many logics relevant to computer science are fragments (or extensions of fragments) of first-order logic, for example k-variable logics and various modal logics. Finding Lindström theorems for such languages can be a challenging problem, since most techniques used in the past to prove Lindström theorems rely on coding arguments that seem to require the full expressive power of first-order logic. For a recent survey of Lindström theorems in a general setting, see [6].

We follow two global lines of attack. First, we take the original Lindström theorem for first-order logic and generalize the proof as much as possible. In this way, we obtain Lindström theorems for the finite variable fragments FO^k

[1] By *extensions of basic modal logic* we mean language extensions, not axiomatic extensions.

with $k > 2$ and Tarski's relation algebra. Next, we take the modal Lindström theorem as a starting point, and try to generalize it to richer languages. In this way, we obtain Lindström theorems for graded modal logic (on arbitrary Kripke structures and on trees) and the binary guarded fragment.

Many open question remain. For example, we have not been able to find Lindström theorems for the two-variable fragment or the full guarded fragment.

2 From first-order logic downwards

In this first part, we take the classic Lindström theorem as a starting point, and we generalize the argument to obtain characterizations for some fragments of first-order logic.

2.1 A strengthening of the Lindström theorem for first-order logic over binary vocabularies

The first-order Lindström theorem uniquely characterizes first-order logic in terms of Compactness and the Löwenheim-Skolem property *within the class of all its extensions*. As we will show in this section, this result can be improved: first-order logic can be characterized in terms of Compactness and the Löwenheim-Skolem property *within the class of all extensions of the three-variable fragment* FO^3, if we consider vocabularies consisting only of unary and binary relation symbols. The proof is not substantially more difficult than that of the original Lindström theorem, but this strengthening will allow us to obtain new results on Tarski's relation algebra and finite variable fragments.

To keep things simple, we will work with a fixed relational signature consisting of a set of unary relation symbols and a set of binary relation symbols, both countably infinite.

By an *abstract logic* we will mean a pair $\mathcal{L} = (\mathsf{Fml}_{\mathcal{L}}, \models_{\mathcal{L}})$, where $\mathsf{Fml}_{\mathcal{L}}$ is the set of sentences of $\mathcal{L}$ and $\models_{\mathcal{L}}$ is a binary relation between $\mathcal{L}$-sentences and models, indicating which sentences are true in which models. If no confusion arises, we will sometimes write $\mathcal{L}$ for $\mathsf{Fml}_{\mathcal{L}}$ and $\models$ for $\models_{\mathcal{L}}$. We assume that $\mathcal{L}$-sentences are preserved under isomorphisms, and that $\mathcal{L}$ has the following properties:

► *closure under Boolean connectives*: for every $\phi \in \mathcal{L}$ there is a $\psi \in \mathcal{L}$ defining its negation (i.e., for all models M, $M \models \psi$ iff $M \not\models \phi$), and for every $\phi, \psi \in \mathcal{L}$ there is $\chi \in \mathcal{L}$ defining the conjunction of ϕ and ψ.

► *closure under renamings*: for every mapping ρ sending relation symbols to relation symbols of the same arity, and for every sentence $\phi \in \mathcal{L}$, there is a sentence $\psi \in \mathcal{L}$ such that for all models M, $M \models \psi$ iff $\rho(M) \models \phi$.

► *closure under relativisation by unary predicates*: for every sentence $\phi \in \mathcal{L}$ and unary relation symbol P, there is a sentence $\psi \in \mathcal{L}$ such that for all models M, $M \models \psi$ iff $M^P \models \phi$, with M^P the submodel of M induced by P.

Examples of abstract logics include first-order logic (FO) and its k-variable fragments (FO^k), with $k \geq 1$.

Given two abstract logics, $\mathcal{L}$ and $\mathcal{L}'$, we say that $\mathcal{L}$ extends $\mathcal{L}'$ (or, $\mathcal{L}'$ is contained in $\mathcal{L}$, denoted by $\mathcal{L}' \subseteq \mathcal{L}$), if there is a map $f : \mathsf{Fml}_{\mathcal{L}'} \to \mathsf{Fml}_{\mathcal{L}}$ preserving truth in the sense that, for all models M and sentences $\phi \in \mathcal{L}'$, $M \models_{\mathcal{L}'} \phi$ iff $M \models_{\mathcal{L}} f(\phi)$.

An abstract logic $\mathcal{L}$ has Compactness if for every set of $\mathcal{L}$-formulas Σ, if every finite subset of Σ is satisfiable then the entire set Σ is satisfiable. An abstract logic $\mathcal{L}$ has the Löwenheim-Skolem property if every satisfiable set of $\mathcal{L}$-formulas has a countable model.

First, we show that each compact extension of FO (in fact already of FO^2) has the "finite occurrence property".

Lemma 2.1 (Finite occurrence property) *Let $\mathcal{L}$ be any abstract logic extending FO^2 that has Compactness. Then for any $\phi \in \mathcal{L}$ there is a finite set of relation symbols $REL(\phi)$ such that the truth of ϕ in any model is independent of the denotation of relation symbols outside $REL(\phi)$.*

The proof (a standard argument) is given in Appendix A.

Lemma 2.2 *Let $\mathcal{L}$ be any abstract logic with the Löwenheim-Skolem property, such that $\mathcal{L}$ extends FO^3 and is not contained in FO. Then "$\mathcal{L}$ can relatively projectively define finiteness": there is a formula $\psi \in \mathcal{L}$ containing a unary predicate N, such that, for each $n \in \mathbb{N}$, there is a model of ψ in which exactly n elements satisfy N, while no model of ψ has infinitely many elements satisfying N.*

Proof: The basic idea is the same as in traditional proofs of the Lindström theorem (e.g., [5]). Our main contribution is to show that, in the case of binary vocabularies, the coding argument requires only three variables.

Take any $\phi \in \mathcal{L}$ not belonging to FO. Then for each $k \in \mathbb{N}$, there are models $\mathfrak{A}_k \models \phi$ and $\mathfrak{B}_k \not\models \phi$ that are potentially isomorphic up to back-and-forth depth k, while, at the same time, no potentially isomorphic models disagree on ϕ. We can describe this situation *inside* $\mathcal{L}$. The construction is outlined in Figure 2.

The model depicted in Figure 2 describes two models, connected via a collection of partial isomorphisms, that disagree on ϕ. The most important feature is that, if N is an infinite set, then the collection of partial isomorphisms constitutes a potential isomorphism, whereas if N is finite (say, of size k), the collection of partial isomorphisms constitutes a potential isomorphism *up to back-and-forth depth* k.

More precisely, A and B are unary predicates defining the domains of two (sub)models, P is a unary predicate whose elements denote pairs from $A \times B$, and the elements of F represent partial isomorphisms (i.e., sets of pairs constituting structure preserving bijections). The arrows represent a binary relation R. For instance, in the given example,

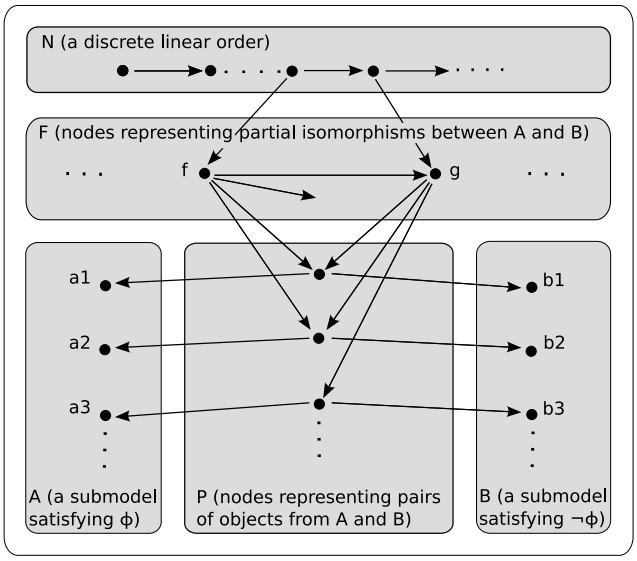

Figure 2. Model from the proof of Lemma 2.2.

f represents the partial isomorphism $\{(a_1, b_1), (a_2, b_2)\}$, and g represents the partial isomorphism that extends f with the pair (a_3, b_3). The elements of the linearly ordered set N are used as index of the partial isomorphisms.

Claim: Each of the following properties of this model can be expressed by a sentence of $\mathcal{L}$:

1. Every $p \in P$ is associated to a pair from $A \times B$.
2. Every $f \in F$ is associated to a set of elements of P that form a partial *bijection* between A and B.
3. Each such partial bijection preserves structure on the submodels defined by A and B, as far as the (finitely many) relations occurring in ϕ are concerned.
4. Every $f \in F$ has a unique associated 'index' from N.
5. N is linearly ordered by R, such that there is a minimal element, and each non-maximal element has an immediate successor (in particular, if N is infinite then it contains an infinite ascending chain).
6. If fRg for $f, g \in F$, this means that g extends f (as a partial bijection), and that the index of g is the successor of the index of f.
7. The *back-and-forth* properties hold for partial isomorphisms whose index is not the maximal element of N.
8. Some $f \in F$ has as index the minimal element of N.
9. The submodels defined by A and B disagree on ϕ. (Recall that $\mathcal{L}$ is closed under the Boolean connectives and relativisation by unary predicates).

Proof of claim: The first eight properties can already be expressed in FO^3 by a clever re-use of variables, and the ninth property can be expressed in $\mathcal{L}$ by closure under the Boolean connectives and relativisation by unary predicates.

For instance, the third property is expressed as the conjunction of all FO^3-formulas of the following forms, for $S \in REL(\phi)$ a binary relation symbol, and $Q \in REL(\phi)$ a unary relation symbol.

$$\forall xy\Big(Px \wedge Py \wedge \exists z(Fz \wedge Rzx \wedge Rzy) \rightarrow \big(\exists z(Rxz \wedge Az \wedge \exists x(Ryx \wedge Ax \wedge Szx)) \leftrightarrow \exists z(Rxz \wedge Bz \wedge \exists x(Ryx \wedge Bx \wedge Szx))\big)\Big) \quad \text{and}$$

$$\forall x\Big(Px \rightarrow \big(\exists z(Rxz \wedge Az \wedge Qz) \leftrightarrow \exists z(Rxz \wedge Bz \wedge Qz)\big)\Big)$$

Note that we crucially use the fact that the signature consists of unary and binary relations only. **End of proof of claim.**

Let χ be the conjunction of all these $\mathcal{L}$-sentences. By construction, χ has models in which N has arbitrarily large finite cardinality (this follows from the existence, for each $k \in \mathbb{N}$, of models disagreeing on ϕ that are potentially isomorphic up to back-and-forth depth k). However, there is no model of χ in which N is an infinite set (if there were, then, by the Löwenheim-Skolem property, there would be a *countable* such model, and in countable models, being potentially isomorphic means being isomorphic; thus, there would be isomorphic models disagreeing on ϕ, which contradicts $\mathcal{L}$'s invariance for isomorphisms). In other words, χ relatively projectively defines finiteness. $\square$

Theorem 2.3 *An abstract logic extending FO^3 is contained in FO iff it satisfies both Compactness and the Löwenheim-Skolem property.*

Proof: If an abstract logic is contained in FO, then, clearly, it satisfies Compactness and the Löwenheim-Skolem property. If, on the other hand, an abstract logic $\mathcal{L}$ extends FO^3 but is not contained in FO, then it must lack either the Löwenheim-Skolem property or Compactness. For, suppose $\mathcal{L}$ satisfies the Löwenheim-Skolem property, let $\psi(N)$ be any $\mathcal{L}$-sentence projectively defining finiteness (cf. Lemma 2.2), and for each $k \in \mathbb{N}$, let χ_k be an FO^3-formula expressing that there are at least k distinct N-elements (by reusing bound variables as in $\exists x(Nx \wedge \exists y(Ny \wedge Rxy \wedge \exists x(Nx \wedge Ryx \wedge \cdots))))$. Then every finite subset of $\{\chi_k \mid k \in \mathbb{N}\} \cup \{\psi(N)\}$ has a model while the entire set has no model. Thus, $\mathcal{L}$ lacks compactness. $\square$

Note that this result relies on our restriction to at most binary relation symbols. In the case with at most k-ary relations ($k \geq 1$) an analogous result holds for FO^{k+1} (the case for $k = 1$ uses a different, easier argument).

The following results can be proved in a similar fashion (relying again on the restriction to unary and binary relation symbols). We omit the details.

Theorem 2.4 *An abstract logic extending FO^3 is contained in FO iff it satisfies Compactness and invariance for potential isomorphisms.*

Theorem 2.5 *A "concrete" abstract logic extending FO^3 is contained in FO iff it satisfies the Löwenheim-Skolem property and is recursively enumerable for validity.*

Here, by "concrete" we mean that formulas can be coded as finite strings over some alphabet, in such a way that negation, conjunction, and relativisation are computable operations, and there is a computable translation from FO^3-formulas to formulas of the logic. The proof of Theorem 2.5 uses the fact that satisfiability of FO^3 formulas on finite models is undecidable [3].

2.2 First application: Tarski's relation algebra

Tarski's relation algebra $\mathcal{RA}$ [11] is an algebraic language in which the terms denote binary relations. It has atomic terms $R, S, \ldots$ ranging over binary relations (over some domain), constants δ and $\top$ denoting the identity relation and the total relation, and operators $\cap, -, \bullet, (\cdot)^{\smile}$ for taking the intersection, complement, composition and converse of relations. Thus, the syntax of $\mathcal{RA}$ is as follows:

$$\alpha ::= R \mid \delta \mid \top \mid \alpha \cap \beta \mid -\alpha \mid \alpha \bullet \beta \mid \alpha^{\smile}$$

with R an element from some countably infinite set of variables standing for binary relations. An interpretation for this language is a set X together with an assignment of binary relations over X to the atomic terms. In other words, it is a first-order structure for the vocabulary that contains the atomic terms as binary relation symbols. We write $\alpha \equiv \beta$ if, in each interpretation, α and β denote the same binary relation, and we write $\alpha \subseteq \beta$ if, in each interpretation, α denotes a subrelation of the relation denoted by β.

In this section, we provide a Lindström-theorem for extensions of relation algebra. By an *extended relation algebra* we will mean any language obtained by extending the syntax of relation algebra with zero or more additional logical operations, where a *logical operation* is any operation that takes as input a fixed finite number of binary relations $R_1, \ldots, R_n$ (over some set X), and that produces a new binary relation S over the same set. We also require logical operations to respect isomorphisms, and to be *domain independent* in the following sense (familiar from database theory): the output relation S does not depend on elements of the domain X that do not participate in any pair belonging to any input relation R_i. This excludes for instance *complementation* as a logical operation, but, in terms of expressivity, it is not an essential restriction: one can always *relativise* such operations, by introducing an additional argument. For instance, absolute complementation as in $-R$ may be replaced by relative complementation as in $\top - R$.

One example of an extended relation algebra, $\mathcal{RA}_{FO}$, is the extension of relation algebra with *all first-order definable logical operations* (see e.g. [15]). Another example is $\mathcal{RAT}$, the extension of relation algebra with the *transitive closure* operation [10].

The compactness and Löwenheim-Skolem properties can be defined for extended relation algebras as usual. For instance, we say that an extended relation algebra $\mathcal{L}$ has the Löwenheim-Skolem property if every set of $\mathcal{L}$-expressions Φ, if there is an interpretation under which $\bigcap \Phi$ is a non-empty relation, then there is such an interpretation over a countable domain. As is not hard to see, $\mathcal{RA}$ and $\mathcal{RA}_{FO}$ satisfy both Compactness and the Löwenheim-Skolem property, whereas $\mathcal{RAT}$ satisfies the Löwenheim-Skolem property but lacks Compactness.

The following Lindström-style theorem shows that all extended relation algebras containing non-first-order definable operations lack either Compactness or the Löwenheim-Skolem property.

Theorem 2.6 *Let $\mathcal{L}$ be any extended relation algebra with the Compactness and Löwenheim-Skolem properties. Then every logical operation of $\mathcal{L}$ is first-order definable.*

Proof: Lemma 2.2 can be adapted to the relation algebra setting, allowing us to show that every extended relation algebra containing a non-elementary logical operation and having the Löwenheim-Skolem property can projectively define finiteness, and hence lacks Compactness. We will not spell out the details, but merely mention the following key points of the proof:

▶ For every first-order sentence ϕ containing only three variables, in a signature consisting only of binary relations, there is a relation algebra expression α such that for every model M, $M \models \phi$ iff $\alpha \equiv \top$ holds in M [11].

▶ Unary relations can be mimicked by binary relations by systematically intersecting them with the identity relation.

▶ In this way, every extended relation algebra gives rise to an abstract logic extending FO^3. Closure under relativisation of the logic is guaranteed by the domain independence of the logical operations of $\mathcal{L}$. □

In other words, the extension of relation algebra with *all* elementary operations is the greatest extension that satisfies Löwenheim-Skolem and Compactness. The same holds if we replace the Löwenheim-Skolem property by invariance for potential isomorphisms, or if we replace Compactness by recursive enumerability.

Theorem 2.6 nicely complements a known result: *every extended relation algebra with Craig interpolation can define all first-order definable operations* [12]. Together, these results show that $\mathcal{RA}_{FO}$ is the unique (up to expressive equivalence) extension of $\mathcal{RA}$ satisfying Compactness, Löwenheim-Skolem, and Craig Interpolation.

2.3 Second application: finite variable fragments

In this section, we provide Lindström theorems for the finite variable fragments FO^k with $k > 3$, over vocabularies consisting of unary and binary relation symbols only. It is well known that the finite variable fragments can be characterized as fragments of first-order logic using potential isomorphisms with a restricted number of pebbles:

Definition 2.7 (k-pebble potential isomorphisms) *A k-pebble potential isomophism between M and N is a non-empty family F of finite partial isomorphisms f between M and N with $|dom(f)| \leq k$ with the following properties:*

► ***Forth:*** *for all $f \in F$, $(w_1, v_1), \ldots, (w_n, v_n) \in f$ with $n < k$, and $w \in M$, there is an $v \in N$ such that $\{(w_1, v_1), \ldots, (w_n, v_n), (w, v)\} \in F$*
► ***Back:*** *for all $f \in F$, $(w_1, v_1), \ldots, (w_n, v_n) \in f$ with $n < k$, and $v \in N$, there is an $w \in M$ such that $\{(w_1, v_1), \ldots, (w_n, v_n), (w, v)\} \in F$*

Fact 2.8 *FO^k is (up to logical equivalence) the fragment of FO invariant for k-pebble potential isomorphisms ($k \geq 1$).*

For a proof of Fact 2.8 see for instance [1].

Using Lemma 2.2, we can turn this into the following Lindström characterization (remember that we only consider unary and binary relation symbols):

Theorem 2.9 (Lindström theorem for FO^k) *Let $k \geq 3$. An abstract logic extending FO^k satisfies Compactness and invariance for k-pebble potential isomorphisms iff it is no more expressive than FO^k.*

Proof: Consider any abstract logic $\mathcal{L}$ extending FO^3 that has Compactness and is invariant for k-pebble potential isomorphisms. Then in particular it is invariant for potential isomorphisms, and therefore by Theorem 2.4 it must be contained in FO. But then, by Fact 2.8, it must be contained in FO^k. □

Theorem 2.9 can be seen as a strengthening of Fact 2.8 (for $k \geq 3$, and on binary vocabularies). Indeed, it implies the following "generalized preservation theorem" (again, with only unary and binary relation symbols):

Corollary 2.10 *Let $k \geq 3$. Let $\mathcal{L}$ be any abstract logic extending FO^k with Compactness. Then FO^k is the fragment of $\mathcal{L}$ invariant for k-pebble potential isomorphisms (up to logical equivalence). In particular, FO^k is the fragment of FO invariant for k-pebble potential isomorphisms.*

Proof: Let $\mathcal{L}$ be any compact abstract logic extending FO^k, and let $\mathcal{L}'$ be the fragment of $\mathcal{L}$ invariant for k-pebble potential isomorphisms. Then $\mathcal{L}'$ satisfies all the requirements of abstract logics. For instance, it is closed under relativisation: if $\phi \in \mathcal{L}$ is invariant for k-pebble potential isomorphisms, then so is its relativisation by a unary predicate. Thus, $\mathcal{L}'$ is an abstract logic extending FO^k that is compact and invariant for k-pebble potential isomorphisms. Hence, by Theorem 2.9, it is contained in FO^3. □

3 From modal logic upwards

The approach of generalizing the classic Lindström theorem only got us so far. It enabled us to characterize FO^k for $k \geq 3$ but is unlikely to reach much further down. We will now take a different approach, by considering the modal Lindström theorem, and trying to generalize it to richer languages. In particular, we obtain two Lindström theorems for the graded modal logic.

3.1 The modal Lindström theorem revisited

We recall the proof of the modal Lindström theorem of [14] (which improves on an earlier result in [4]). First, we need to define "abstract modal logics". As before, we assume a fixed vocabulary, now consisting of a single binary relation symbol R and a countably infinite set of unary relation symbols, also called *proposition letters*. Structures for this vocabulary are usually called *Kripke models* (the restriction to a single binary relation symbol is not essential but is convenient for presentation). We associate to each formula a class of pairs (M, w), where M is a Kripke model and w is an element of the domain of M. This is because modal formulas are always evaluated *at a point in a model*. We call such pairs (M, w) *pointed Kripke models*. Thus, an *abstract modal logic* is a pair $\mathcal{L} = (Fml_{\mathcal{L}}, \models)$, where $Fml_{\mathcal{L}}$ is the set of formulas of $\mathcal{L}$ and $\models$ is a binary relation between $\mathcal{L}$-formulas and pointed Kripke models. As before, when no confusion arises we will write $\mathcal{L}$ for $Fml_{\mathcal{L}}$. Also, as before, we assume that $\mathcal{L}$-formulas are invariant for isomorphisms, and that $\mathcal{L}$ is closed under the Boolean operations, renaming, and relativisation by unary predicates.

Examples of abstract modal logics include *basic modal logic*, its extension with counting modalities called *graded modal logic* (GML), first-order logic (by which we mean the collection of all first-order formulas in one free variable, over the appropriate signature), and the modal μ-calculus. For the syntax and semantics of basic modal logic, the reader may consult any modal logic textbook (e.g., [2]).

The modal Lindström theorem characterizes basic modal logic in terms of Compactness and invariance for bisimulations. A bisimulation between Kripke models M and N is a binary relation Z between the domains of M and N satisfying the following three conditions:

► **Atomic harmony:** if wZv then w and v agree on all proposition letters (i.e., unary predicates).

► **Zig:** if wZv and wR^Mw', there is a v' such that vR^Nv' and $w'Zv'$.
► **Zag:** if wZv and vR^Nv', there is a w' such that wR^Mw' and $w'Zv'$.

Two pointed Kripke models (M, w) and (N, v) are *bisimilar* if there is a bisimulation Z between M and N such that wZv. A formula is *bisimulation invariant* if it does not distinguish bisimilar pointed Kripke models, and an abstract modal logic is bisimulation invariant if all its formulas are.

Given a pointed Kripke model (M, w), we denote by M_w the submodel of M containing all points that are reachable in finitely many steps from w along the binary relation. Likewise, for $k \in \mathbb{N}$, M_w^k is the submodel of M containing all points reachable from w in at most k steps along the binary relation. We say that a formula ϕ is *invariant for generated submodels* if, for all models M with worlds w, $(M, w) \models \phi$ iff $(M_w, w) \models \phi$. We say that ϕ *has the finite depth property* if there is a $k \in \mathbb{N}$ such that $(M, w) \models \phi$ iff $(M_w^k, w) \models \phi$, for all models M with worlds w. Clearly, the latter implies the former. Also, bisimulation invariance implies invariance for generated submodels, because the natural inclusion map is a bisimulation. An abstract modal logic $\mathcal{L}$ is invariant for generated submodels (or, has the finite depth property), if every $\phi \in \mathcal{L}$ is invariant for generated submodels (respectively, has the finite depth property).

We are now ready to proceed with the proof of Theorem 3.3. We first prove a finite occurrence property for compact extensions of basic modal logic that satisfy invariance for bisimulation (in fact, we only need to assume invariance for generated submodels).

Lemma 3.1 (Finite occurrence property) *Let $\mathcal{L}$ be any abstract modal logic extending basic modal logic that is Compact and invariant for generated submodels. Then for each $\phi \in \mathcal{L}$ there is a finite set of proposition letters $PROP(\phi)$ such that the truth of ϕ in any pointed Kripke model is independent of the denotation of proposition letters outside $PROP(\phi)$.*

The proof is given in Appendix A.

Lemma 3.2 *Let $\mathcal{L}$ be any abstract modal logic extending basic modal logic that is Compact and invariant for generated submodels. Then $\mathcal{L}$ has the finite depth property.*

Proof: Let $\phi \in \mathcal{L}$, let p be a proposition letter not occurring in ϕ, and let ϕ^p be the relativisation of ϕ by p. By the generated submodel-invariance of $\mathcal{L}$, $\{p, \Box p, \Box\Box p, \ldots\} \models \phi \leftrightarrow \phi^p$. By the compactness of $\mathcal{L}$, there is an $n \in \mathbb{N}$ such that $\{p, \Box p, \ldots, \Box^n p\} \models \phi \leftrightarrow \phi^p$. But this expresses exactly that ϕ has the finite depth property, for depth n. □

Theorem 3.3 ([14]) *An abstract modal logic extending basic modal logic satisfies Compactness and bisimulation invariance iff it is no more expressive than basic modal logic.*

Proof: Let $\mathcal{L}$ be any abstract modal logic extending basic modal logic and satisfying Compactness and bisimulation invariance. Since bisimulation invariance implies invariance for generated submodels, $\mathcal{L}$ is also invariant for generated submodels. Hence, by Lemma 3.1 and Lemma 3.2, $\mathcal{L}$ has the finite occurrence property and the finite depth property. Next, we use the following well known fact [2]:

> Assuming a finite vocabulary, every bisimulation-invariant class of pointed models with the finite depth property is definable by a formula of basic modal logic.

We conclude that $\mathcal{L}$ is contained in basic modal logic. □

Theorem 3.3 can be seen as a strengthening of the more familiar characterization of basic modal logic as the bisimulation invariant fragment of first-order logic. Indeed, it implies the following "generalized preservation theorem":

Corollary 3.4 *Let $\mathcal{L}$ be any abstract modal logic extending basic modal logic that has Compactness. Then basic modal logic is the bisimulation invariant fragment of $\mathcal{L}$ (up to logical equivalence). In particular, basic modal logic is the bisimulation invariant fragment of first-order logic.*

Proof: Let $\mathcal{L}'$ be the bisimulation invariant fragment of $\mathcal{L}$. Then $\mathcal{L}'$ satisfies all criteria for being an abstract modal logic. For instance it is closed under relativisation: whenever $\phi \in \mathcal{L}$ is invariant for bisimulations then the relativisation of ϕ by a unary predicate is also invariant for bisimulations. Likewise for the Boolean connectives.

Hence, $\mathcal{L}'$ is an abstract modal logic extending basic modal logic and it is bisimulation invariant and Compact. Hence, it is no more expressive than basic modal logic. □

Corollary 3.4 strengthens the traditional bisimulation preservation theorem, as there are compact extensions of basic modal logic not contained in first-order logic. Indeed:

Theorem 3.5 *There is an abstract modal logic extending basic modal logic that is not contained in first-order logic, but still satisfies Compactness, the Löwenheim-Skolem property, invariance for potential isomorphisms, and Craig interpolation. Moreover, it has a finite axiomatization and a* PSPACE*-complete satisfiability problem.*

The proof is given in Appendix B. The logic in question extends basic modal logic with an extra operator $\bullet$, with the following semantics: $M, w \models \bullet\phi$ *iff* w *has infinitely many reflexive successors satisfying* ϕ.

Theorem 3.5 is also interesting for another reason: it shows that Theorem 2.3 no longer holds when FO^3 is replaced by basic modal logic.

3.2 Graded modal logic

Graded modal logic (GML) extends basic modal logic with counting modalities: for each formula ϕ and natural number k, $\Diamond_k \phi$ is admitted as a formula, and it says that at least k successors of the current node satisfy ϕ.

GML-formulas are in general not invariant for bisimulations. Still, an important (weaker) invariance property does hold: GML formulas are invariant for *tree unraveling*. A *tree model* is a Kripke model whose underlying frame is a tree (possibly infinite, but well-founded, and with a unique root). We will denote tree models by $T, T', \ldots$, and we will use $\mathsf{root}(T)$ to denote the root of the tree model T. Every pointed Kripke model (M, w) can be unravelled into a tree model, by the following standard construction:

Definition 3.6 (Tree unraveling) *Given a Kripke model $M = (W, R, V)$ and $w \in W$, the tree unraveling $\mathsf{unr}(M, w)$ is defined as (W', R', V'), where*

- ▶ *W' consists of all finite paths $\langle w_1, w_2, \ldots, w_n \rangle$ satisfying $w_1 = w$ and $w_i R w_{i+1}$.*
- ▶ *R' contains all pairs of sequences of the form $(\langle w_1, \ldots, w_n \rangle, \langle w_1, \ldots, w_n, w_{n+1} \rangle) \in W' \times W'$*
- ▶ *$\langle w_1, \ldots, w_n \rangle \in V'(p)$ iff $w_n \in V(p)$.*

It is easily seen that, for any pointed Kripke model (M, w), $\mathsf{unr}(M, w)$ is indeed a tree model, and that $\langle w \rangle$ is its root. GML-formulas are invariant for this operation:

Fact 3.7 (GML is invariant for tree unravelings) *For all pointed Kripke models (M, w) and GML-formulas ϕ, $M, w \models \phi$ iff $\mathsf{unr}(M, w), \langle w \rangle \models \phi$.*

We will prove two Lindström theorem for GML. The first characterizes GML on arbitrary structures in terms of Compactness, the Löwenheim-Skolem property and invariance for tree unravelings. It can be seen as a natural generalization of Theorem 3.3. The second theorem, which will be proved in the following section, considers GML as a language for describing (nodes in) tree models, and it characterizes GML as being maximal with respect to Compactness and the Löwenheim-Skolem property on such structures.

Recall Theorem 3.3, which characterizes modal logic in terms of Compactness and bisimulation invariance. One might wonder if, likewise, GML can be characterized in terms of Compactness and invariance for tree unraveling. The answer is negative: the extension of GML with the modal operator $\Diamond_{\aleph_1}$ ("uncountably many successors ...") still has these properties. Instead, we prove the following:

Theorem 3.8 *An abstract modal logic extending GML satisfies invariance for tree unravelings, Compactness, and the Löwenheim-Skolem property iff it is no more expressive than GML.*

As in the case of modal logic, we obtain the following "generalized preservation theorem" as a corollary (the proof is analogous to that of Corollary 3.4):

Corollary 3.9 *Let $\mathcal{L}$ be any abstract modal logic extending GML and satisfying Compactness and the Löwenheim-Skolem property. Then GML is the tree unraveling invariant fragment of $\mathcal{L}$ (up to logical equivalence). In particular, GML is the tree unraveling invariant fragment of FO.*

The rest of this section is dedicated to the proof of Theorem 3.8. Two facts about GML will be used in the proof:

Fact 3.10 (GML has the finite tree model property) *Every satisfiable GML formula is satisfied at the root of a finite tree model.*

Fact 3.11 (GML can describe finite tree models up to isomorphism) *Assuming a finite vocabulary, for every finite tree model T there is a GML-formula ψ_T such that for every tree model T', $(T', \mathsf{root}(T')) \models \psi_T$ iff $T' \cong T$.*

Now for the main argument. Fix an abstract modal logic $\mathcal{L}$ extending GML and satisfying Compactness and Löwenheim-Skolem, and invariance for tree unravelings. By Lemma 3.1 and 3.2, $\mathcal{L}$ has the finite occurrence property and the finite depth property (note that invariance for tree unraveling implies generated submodel invariance). The following Lemma shows a kind of "finite width property".

Lemma 3.12 ($\mathcal{L}$-formulas can only count successors up to a finite number) *For each formula $\phi \in \mathcal{L}$ and finite tree model T, there is a natural number k such that "ϕ can only count T-successors up to k": whenever a tree model contains a node v that has k successor subtrees isomorphic to T, then adding more copies of T will not affect the truth value of ϕ at the root.*

Proof: Since T is a finite tree model, it can be described up to isomorphism by a single GML-formula ψ_T. Let Σ be the following set of formulas, where p is a proposition letter not occurring in ϕ, and $\Box^n$ stands for a sequence of n boxes:

$\{p,\ \Box^n(\neg p \to \Box \neg p) \mid n \in \mathbb{N}\}$
"p defines an initial subtree"
$\cup\ \{\Box^n(p \to \Box(\neg p \to \psi_T)) \mid n \in \mathbb{N}\}$
"the root of every $\neg p$-subtree satisfies ψ_T"
$\cup\ \{\Box^n(p \wedge \Diamond \neg p \to \Diamond_m(p \wedge \psi_T)) \mid n, m \in \mathbb{N}\}$
"every p-node with a $\neg p$-successor has infinitely many p-successors satisfying ψ_T"

Whenever a *countable tree model* satisfies Σ at the root, the submodel defined by p is isomorphic to the whole model — isomorphic in the language without p, to be precise (see Figure 3). Since $\mathcal{L}$ satisfies the Löwenheim-Skolem property, invariance for tree unraveling and invariance for isomorphisms, we can conclude that $\Sigma \models \phi \leftrightarrow \phi^p$. But then,

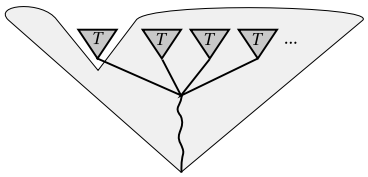

Figure 3. Model from the proof of Lemma 3.12

by compactness, there is a $k \in \mathbb{N}$ such that $\Sigma_k \models \phi \leftrightarrow \phi^p$, where Σ_k is the following subset of Σ:

$$\{p,\ \Box^n(\neg p \to \Box\neg p) \mid n \in \mathbb{N}\} \cup$$
$$\{\Box^n(p \to \Box(\neg p \to \psi_T)) \mid n \in \mathbb{N}\} \cup$$
$$\{\Box^n(p \wedge \Diamond\neg p \to \Diamond_m(p \wedge \psi_T)) \mid n, m \in \mathbb{N} \textit{ with } m \leq k\}$$

This shows that the Lemma holds. □

Proof of Theorem 3.8: Let $\mathcal{L}$ be any abstract modal logic extending GML, satisfying Compactness, the Löwenheim-Skolem property, and tree unraveling invariance. Observe that $\mathcal{L}$ still satisfies Compactness and the Löwenheim-Skolem property if we restrict attention to trees (note that the tree unraveling of a countable model is countable).

Consider any formula $\phi \in \mathcal{L}$. We will construct a set of equivalence relations $\sim_i^\phi$ for tree models (with $i \geq 0$), satisfying the following two properties:

1. $T \sim_i^\phi T'$ implies that the truth value of ϕ at the root of a tree model is not affected if subtrees isomorphic to T at depth i are replaced by copies of T' (or vice versa).
2. Each $\sim_i^\phi$ has only finitely many equivalence classes, and each is definable by a GML-formula.

This then implies that ϕ is equivalent to a GML formula (take the disjunction of the GML-formulas defining the $\sim_0^\phi$-equivalence classes that satisfy ϕ).

The claim holds trivially for $i > depth(\phi)$. Next, assume that the claim holds for $i+1$. We will show that it also holds for i. Let $K_1, \ldots, K_n$ be the (finitely many) $\sim_{(i+1)}^\phi$-equivalence classes, and for each $\ell \leq n$, pick a finite representative $T_\ell \in K_\ell$ (using Proposition 3.10). It follows from Lemma 3.12, there is a $k \in \mathbb{N}$ such that, for all $\ell \leq n$ even: "ϕ can only count T_ℓ-successors up to k", and hence, by $\sim_{i+1}^\phi$-equivalence, "ϕ can only count K_ℓ-successors up to k, at depth i". But then, it follows that there are at most $k^n \cdot 2^{PROP(\phi)}$ many $\sim_i^\phi$-equivalence classes. Moreover, they are definable by GML-formulas (in fact, by Boolean combinations of proposition letters and formulas of the form $\Diamond_m\psi$ with $m \leq k$ and ψ a GML-formula defining some $\sim_{i+1}^\phi$-equivalence class).

Thus, $\mathcal{L}$ is not more expressive than GML on tree models. It follows that $\mathcal{L}$ is not more expressive than GML on arbitrary Kripke models: consider any $\mathcal{L}$-formula ϕ, and let ψ be any GML-formula equivalent to ϕ on tree models. If $\phi \leftrightarrow \psi$ were falsifiable, then, by unraveling, it could be falsified on a tree, which, by assumption, is not the case. Thus, ϕ and ψ are equivalent on all Kripke models. □

3.3 Graded modal logic on trees

In this section, rather than assuming tree unraveling invariance, we consider only tree models from the start. That is, we view GML as a language for describing nodes of tree models. From this perspective, GML has three distinctive limiting features: *(i)* when evaluated in a node, formulas can only see the subtree starting from that node; *(ii)* when evaluated at a node, each formula can only look finitely deeply into the subtree starting at that node; *(iii)* each formula can only count successors up to a finite number (viz. the largest natural number occurring as an index in the formula).

Fact 3.13 *On trees, GML has the Compactness and Löwenheim-Skolem properties.*

Proof: Follows from the fact that GML has these properties on arbitrary structures (recall that the tree unraveling of a countable model is still countable). □

We will turn this into a Lindström characterization for GML on trees. However, before we proceed, two technical issues need to be discussed. The first concerns *closure under relativisation*. Recall from Sect. 3.1 that all abstract modal logics are assumed to satisfy this property. But what does it mean for a logic to be closed under relativisation if we consider only trees? Note that a submodel of a tree is not necessarily a tree. We solve this problem as follows. Given a tree model T containing a node n, and a unary predicate p true at n, we define $Subtree(T, n, p)$ to be the largest subtree of T that contains n and contains only nodes satisfying p. Note that n is not necessarily the root of $Subtree(T, n, p)$. We say that an abstract modal logic $\mathcal{L}$ is closed under relativisation on trees, if for every formula $\phi \in \mathcal{L}$ and proposition letter p, there is a formula $\psi \in \mathcal{L}$ such that for all pointed tree models (T, n), $(T, n) \models \psi$ iff $(T, n) \models p$ and $(Subtree(T, n, p), n) \models \phi$. In the case of GML, we can simply pick ψ to be the syntactic relativisation of ϕ by p, i.e., the formula obtained from $\phi \wedge p$ by replacing all subformulas of the form $\Diamond_n\psi$ by $\Diamond_n(p \wedge \psi)$,

Secondly, we need to make an extra assumption, namely that the extensions $\mathcal{L}$ we consider are *closed under substitution*. Intuitively, this means that $\mathcal{L}$ allows us to uniformly substitute formulas for proposition letters. More precisely, $\mathcal{L}$ is closed under substitution if for all formulas $\phi, \psi \in \mathcal{L}$ and proposition letters p, there is a formula χ such that for all pointed (tree) models (M, w), $(M, w) \models \chi$ iff $(M^{[p \mapsto \{v \mid (M,v) \models \psi\}]}, w) \models \phi$, where $M^{[p \mapsto \{v \mid (M,v) \models \psi\}]}$ is obtained from M by changing the valuation of p to $\{v \mid (M, v) \models \psi\}$. If we would not assume closure under

substitution, there would be proper extensions of GML with Compactness and Löwenheim-Skolem. Indeed, the extension of GML with formulas of the form $\Diamond^- p$ ("the current node has a parent that satisfies the proposition letter p") is an example. This logic is not closed under substitution.

Theorem 3.14 *An abstract modal logic closed under substitution and extending GML on trees satisfies Compactness and the Löwenheim-Skolem property on trees iff it is no more expressive than GML on trees.*

This is remarkable, since Compactness and Löwenheim-Skolem are also the characterizing features of first-order logic in the classic Lindström theorem. Note that first-order logic lacks Compactness on trees (due to well foundedness).

Now for the proof. Let $\mathcal{L}$ be an abstract modal logic satisfying the conditions of Theorem 3.14. Lemma 3.16 below shows that $\mathcal{L}$-formulas can only look downwards in the tree.

Lemma 3.15 *If $\mathcal{L}$ is not invariant for generated submodels, then there is a formula $\chi \in \mathcal{L}$ containing a unary predicate p such that the following two conditions hold:*

1. *for all pointed trees (T, n), $(T, n) \models \chi$ implies that n has a parent satisfying p*
2. *there is a pointed tree (T, n) satisfying χ in which p is true only at the parent of n.*

Proof: Since $\mathcal{L}$ is not invariant for generated submodels, there is a formula $\phi \in \mathcal{L}$ and a pointed tree (T, n) such that $(T, n) \models \phi$ and $(T_n, n) \not\models \phi$, or vice versa. Since $\mathcal{L}$ is closed under negation, we may assume w.l.o.g. that the former applies. Moreover, since $\mathcal{L}$ is closed under renamings and the set of proposition letters is infinite, we may assume w.l.o.g. that there are infinitely many proposition letters not occurring in ϕ, in the sense that their interpretation does not influence truth of ϕ at any state. We will refer to these proposition letters as being "fresh".

Pick two distinct fresh proposition letters p, q, let ϕ^p and ϕ^q be the relativisations of ϕ to p and to q, and let Σ be the following set of $\mathcal{L}$-formulas:

$$\Sigma = \{\phi^p, \neg\phi^q, \Box^k(p \wedge q) \mid k \geq 0\}$$

Observe that Σ is satisfiable: it is true at (T, n) when we make p true at all nodes, and q only at n and its descendants. We claim that truth of Σ at a node n in a tree implies that n has a parent and it satisfies either p or q. For, if not then the submodels $Subtree(T, n, p)$ and $Subtree(T, n, q)$ would coincide, and hence ϕ^p and ϕ^q would have to have the same truth value at n.

Next, we will use Compactness to obtain a finite subset of Σ that implies that the current node has a parent satisfying $p \vee q$. First, we 'redescribe' the situation encoded by Σ from the perspective of the parent node. Let Σ' be the following set of $\mathcal{L}$-formulas, where r is another fresh proposition letter:

$$\Sigma' = \{\Diamond(\phi^p \wedge \neg\phi^q \wedge r), \Box(r \to \Box^k(p \wedge q)) \mid k \geq 0\}$$

By the previous observations, $\Sigma' \models p \vee q$. Hence, by Compactness, there is a $\ell \in \mathbb{N}$ such that $\Diamond(\phi^p \wedge \neg\phi^q \wedge r) \wedge \bigwedge_{k \leq \ell} \Box(r \to \Box^k(p \wedge q)) \models p \vee q$. Going back to the perspective of the node n, if we define ψ to be the formula $\{\phi^p \wedge \neg\phi^q \wedge \bigwedge_{k \leq \ell} \Box^k(p \wedge q)$ then ψ is satisfiable and implies the existence of a parent node satisfying $p \vee q$.

Finally, we take two more fresh proposition letters, s and t, and we use the fact that $\mathcal{L}$ is closed under substitution: we define χ to be $s \wedge \psi[p/(p \wedge (\Diamond s \to t)), q/(q \wedge (\Diamond s \to t))]$.

On the one hand, truth of χ at a node implies that it has a parent satisfying either $(p \wedge (\Diamond s \to t))$ or $(q \wedge (\Diamond s \to t))$, and hence t. On the other hand, there is a pointed tree satisfying χ in which t is only true at the parent node: take (T, n) and extend the valuation by making s true only at n, and t at its parent. □

Lemma 3.16 *$\mathcal{L}$ is invariant for generated submodels*

Proof: Suppose not. Let $\chi(p) \in \mathcal{L}$ be as described by Lemma 3.15. By a "fresh renaming" of χ we will mean a copy in which all proposition letters have been renamed to fresh ones, and which has been relativised by an additional fresh proposition letter. For the reasons explained in the proof of Lemma 3.15, we may assume that $\chi(p)$ has infinitely many fresh renamings $(\chi_i(p_i))_{i \in \mathbb{N}}$.

Finally, we define Σ to be the set of $\mathcal{L}$-formulas $\{\chi_1(p_1), \chi_1(\chi_2(p_2)), \chi_1(\chi_2(\chi_3(p_3))), \ldots\}$. Every finite subset of Σ is satisfiable. Indeed, a satisfying model may be constructed by "overlaying" different copies of the model (T, n) from Lemma 3.15(2). On the other hand, if a node would satisfy all formulas in Σ at once, its ancestors would form an infinite ascending chain, which contradicts the well-foundedness property of trees. □

The remainder of the proof of Theorem 3.14 is along the same lines as in Sect. 3.2: first we prove that $\mathcal{L}$ has the finite occurrence, finite depth and finite width properties on trees (using the fact that it is invariant for generated submodels), and then we derive the Lindström theorem by the same argument used in the proof of Theorem 3.8.

3.4 The binary guarded fragment

The *guarded fragment* GF forms a second extension of modal logic, incomparable to graded modal logic. It allows for arbitrary quantifications of the form $\exists\vec{y}(G(\vec{x}, \vec{y}) \wedge \phi(\vec{x}, \vec{y}))$, where $\vec{x}$ and $\vec{y}$ are tuples of variables, and the *guard* G is an atomic formula containing the variables in $\vec{x}$ and $\vec{y}$. The guarded fragment is decidable and has many

'modal' meta-properties, due to its invariance for *guarded bisimulations* [1, 13], see also Appendix C.

Because of its modal character, GF seems an obvious case for a Lindström-style analysis like the one we have given for modal logic and graded modal logic. However, there are some technical difficulties, and we have not been able to prove a Lindström theorem for this language yet. In this section, we focus on a special case, in fact the same special case as in the first half of the paper, namely for vocabularies with only unary and binary predicates. The *Binary Guarded Fragment* (GF_{bin}) has the following syntax:

$$\phi ::= R\vec{x} \mid x = y \mid \neg\phi \mid \phi \vee \psi \mid \exists y(G(x,y) \wedge \psi(x,y))$$

where R is a unary or binary atomic predicate, x and y are distinct variables, the *guard* G is an atomic formula containing both x and y (in any order) and ψ contains no free variables besides (possibly) x and y. 'Unary' guards such as Py, Ryy and $y = y$ are not allowed. GF_{bin} is easily seen to be an abstract logic contained in FO^2. When interpreted over Kripke structures (and considering only formulas with one free variable), it constitutes an abstract modal logic extending basic modal logic.

Guarded bisimulations admit a natural adaptation to this restricted version of the guarded fragment, which we call GF_{bin}-bisimulations, cf. Appendix C.

Theorem 3.17 *An abstract logic extending GF_{bin} satisfies Compactness and Invariance for GF_{bin}-bisimulation iff it is no more expressive than GF_{bin}.*

The proof is along the same lines as for modal logic and graded modal logic: using Compactness, we prove a finite occurrence property and a finite depth property (where depth is now measured as distance in the Gaifman graph). We then use a tree-unraveling argument to show that GF_{bin} can express all properties invariant for guarded bisimulations and having the finite depth property. It is exactly in this last step that the restriction to unary and binary relation symbols is crucial (roughly, it allows us to relate distance in the unravelled tree to Gaifman distance in the original structure). More details are in Appendix C.

4 Open questions

To conclude, we identify three lines of future research.

The scope of Lindström's theorem. Within how broad a class of languages can we characterize first-order logic? Theorem 2.3 gives only a partial answer. For instance, *are there extensions of FO^2 not contained in FO that satisfy Compactness and the Löwenheim-Skolem property?* Note that there *are* extensions of *modal logic* not contained in first-order logic with these properties (cf. Theorem 3.5).

Characterizing more languages. Another line of open problems is to find Lindström characterizations for other fragments, such as FO^2 and the guarded fragment. Also, we can consider non-elementary extensions of modal logic such as the modal μ-calculus. For instance, *can one characterize the modal μ-calculus in terms of bisimulation invariance and the finite model property?*

Characterizing logics on specific classes of structures. No Lindström characterizations are known for first-order logic on finite structures, or on trees. Compactness fails for first-order logic on such structures, and, on finite structures, the Löwenheim-Skolem property becomes meaningless. In this paper, we proved one positive result: we showed that GML behaves on trees as first-order logic does on arbitrary structures: it is maximal with respect to Compactness and the Löwenheim-Skolem property. In general, however, this area remains underexplored [7].

References

[1] H. Andréka, J. van Benthem, and I. Németi. Modal logics and bounded fragments of predicate logic. *Journal of Philosophical Logic*, 27(3):217–274, 1998.

[2] P. Blackburn, M. Rijke, and Y. Venema. *Modal logic*. Cambridge University Press, Cambridge, UK, 2001.

[3] E. Börger, E. Grädel, and Y. Gurevich. *The Classical Decision Problem*. Springer-Verlag, 1997.

[4] M. de Rijke. A Lindström theorem for modal logic. In A. Ponse, M. de Rijke, and Y. Venema, editors, *Modal Logic and Process Algebra: A Bisimulation Perspective*, pages 217–230. CSLI Publications, 1995.

[5] J. Flum. Characterizing logics. In *Model-Theoretic Logics*. Springer, 1985.

[6] M. García-Matos and J. Väänänen. Abstract model theory as a framework for universal logic. In J.-Y. Beziau, editor, *Logica Universalis*. Birkhäuser, 2005.

[7] P. G. Kolaitis and J. A. Väänänen. Generalized quantifiers and pebble games on finite structures. *Annals of Pure and Applied Logic*, 74(1):23–75, 1995.

[8] P. Lindström. On extensions of elementary logic. *Theoria*, 35:1–11, 1969.

[9] M. Marx and Y. Venema. *Multi-dimensional Modal Logic*. Applied Logic Series. Kluwer Academic Publishers, 1997.

[10] K. C. Ng and A. Tarski. Relation algebras with transitive closure. *Notices of the AMS*, 24:A29–A30, 1977.

[11] A. Tarski and S. Givant. *A Formalization of Set Theory without Variables*, volume 41 of *Colloquium Publications*. American Mathematical Society, 1987.

[12] B. ten Cate. Interpolation for extended modal languages. *Journal of Symbolic Logic*, 70(1):223–234, 2005.

[13] J. van Benthem. Guards, bounds, and generalized semantics. *Journal of Logic, Language, and Information*, 14:263–279, 2005.

[14] J. van Benthem. A new modal Lindström theorem. *Logica Universalis*, 1:125–148, 2007.

[15] Y. Venema. Relational games. In H. Andréka, D. Monk, and I. Németi, editors, *Algebraic Logic*, 1991.

A Some missing proofs

Proof of Lemma 2.1: Cf. [5]: since our vocabulary contains infinitely many unary and binary relation symbols, there are renamings ρ_1, ρ_2 whose range is disjoint. Take any $\phi \in \mathcal{L}$, and let ϕ_1, ϕ_2 be its renamings according to ρ_1 and ρ_2. Let $\Sigma = \{\forall x_1 \dots x_k(\rho_1(R)(x_1 \dots x_k) \leftrightarrow \rho_2(R)(x_1 \dots x_k)) \mid R \text{ a } k\text{-ary relation symbol}\}$. Then $\Sigma \models \phi_1 \leftrightarrow \phi_2$, and hence, by Compactness, a finite subset $\Sigma' \subseteq \Sigma$ implies $\phi_1 \leftrightarrow \phi_2$. We can pick for $REL(\phi)$ the relation symbols occurring in Σ'. Note that, in case of binary vocabularies, all formulas in Σ belong to FO^2. □

Proof of Lemma 3.1: Like that of Lemma 2.1. Since the set of proposition letters (unary predicates) is infinite, we can find renamings ρ_1, ρ_2 for them whose range is disjoint. Now, take any $\phi \in \mathcal{L}$, and let ϕ_1, ϕ_2 be its renamings according to ρ_1 and ρ_2. Let $\Sigma = \{\Box^n(\rho_1(p) \leftrightarrow \rho_2(p)) \mid n \in \mathbb{N} \text{ and } p \text{ a proposition letter}\}$, where $\Box^n$ stands for a sequence of n boxes. It follows from generated submodel invariance that $\Sigma \models \phi_1 \leftrightarrow \phi_2$, hence, by Compactness, a finite subset $\Sigma' \subseteq \Sigma$ implies $\phi_1 \leftrightarrow \phi_2$. We can pick for $PROP(\phi)$ the set of proposition letters occurring in Σ'. □

B A well-behaved non-elementary extension of modal logic

In this appendix, we introduce an extension of modal logic, $ML^\bullet$, that is non-elementary but at the same time is very well behaved (e.g., has Compactness and Löwenheim-Skolem). $ML^\bullet$ extends basic modal logic with an extra operator $\bullet$. Thus, the formulas of $ML^\bullet$ are given by:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \psi \mid \Diamond\phi \mid \bullet\phi$$

The semantics of the newly added operator is as follows: $M, w \models \bullet\phi$ *iff* w *has infinitely many reflexive successors satisfying* ϕ (a reflexive world is one related to itself).

The reader may verify that $\bullet$ behaves as any normal modal operator: $\bullet(\phi \vee \psi)$ is equivalent to $\bullet\phi \vee \bullet\psi$, and $\bullet\bot$ is equivalent to $\bot$. The dual of $\bullet$ is denoted by $\bar{\bullet}$.

Proposition B.1 *$ML^\bullet$ is a non-elementary abstract modal logic extending basic modal logic.*

Proof: The reader may verify that this is the case (e.g., the language is closed under relativisation). □

Proposition B.2 *$ML^\bullet$ has the Löwenheim-Skolem property and is invariant for potential isomorphisms.*

Proof: By the containment of $ML^\bullet$ in $L_{\omega_1\omega}$. □

Before we proceed to prove the remaining properties of $ML^\bullet$, we will first provide an alternative semantics for the logic. Let K be the class of all bi-modal Kripke frames $\mathfrak{F} = (W, R_\Diamond, R_\bullet)$ such that $R_\bullet \subseteq R_\Diamond$ and for all $(w, v) \in R_\bullet$, $(v, v) \in R_\Diamond$. Kripke models based on such frames can be thought of as quasi-models for $ML^\bullet$. The next lemma shows that they are adequate as such.

Lemma B.3 *For all $\phi \in ML^\bullet$, ϕ is satisfiable according to the intended semantics iff ϕ is satisfiable on the class K.*

Proof: In one direction, suppose that $M, w \models \phi$ according to the intended semantics, where $M = (W, R, V)$ is a uni-modal Kripke model. Define an equivalence relation $\sim$ on W by letting $v \sim v'$ iff v and v' assign the same truth value to all subformulas of ϕ. Let $R_\Diamond := R$ and $R_\bullet := \{(v, v') \in R \mid (v', v') \in R \text{ and there are infinitely many } v'' \sim v' \text{ such that } (w, v'') \in R \text{ and } (v'', v'') \in R\}$. Then the underlying Kripke frame of M' belongs to K, and a straightforward inductive argument shows that, for all subformulas ψ of ϕ, and for all $v \in W$, $M, v \models \psi$ according to the intended semantics iff $M', v \models \psi$. In particular, $M', w \models \phi$.

Conversely, let $M, w \models \phi$, where $M = (W, R_\Diamond, R_\bullet, V)$ is a bi-modal Kripke model based on a frame in K. We will constuct a uni-modal model M' with a world w' such that $M', w' \models \phi$ according the intended semantics. Roughly, the construction of M' is based on the following ideas: (i) make sure that all $R_\Diamond \setminus R_\bullet$-successors of a world are irreflexive, by unraveling, and (ii) taking infinitely many copies of each $R_\bullet$ successor of each world.

More precisely, as the domain of M', we choose $W \times \mathbb{N}$. The accessibility relation R is the set of all pairs $((v, n), (v', m))$ satisfying one of the following conditions:

- $m \leq 1$, $m \neq n$, and $(v, v') \in R_\Diamond$; or
- $m \geq 2$, and $(v, v') \in R_\bullet$.

Observe that, for all $v \in W$, $(v, 0)$ and $(v, 1)$ are R-irreflexive by construction, whereas (v, n) for $n \geq 2$ is R-reflexive or R-irreflexive, depending on v.

Finally, the valuation function V' for the atomic propositions is defined as usual, by letting $V'(p) = \{(v, n) \mid v \in V(p)\}$. Let $M' = (W \times \mathbb{N}, R, V')$.

A straightforward inductive argument now shows that, for all formulas ψ, worlds $v \in W$ and natural numbers n, $M, v \models \psi$ iff $M', (v, n) \models \psi$ according to the intended semantics. In particular, $M', (w, 0) \models \phi$. □

In the remainder, we use this alternative semantics.

Proposition B.4 *A complete axiomatization for $ML^\bullet$ can be obtained by extending any standard axiomatization for basic (bi-)modal logic with the following axioms:*

$$\bullet\phi \to \Diamond\phi \qquad \textit{and} \qquad \bar{\bullet}(\phi \to \Diamond\phi)$$

In fact, this axiomatization is strongly complete: every consistent set of formulas is jointly satisfiable.

Proof: Both axioms are Sahlqvist formulas, and together, they precisely define the frame class K. It follows by the Sahlqvist completeness theorem that the axiomatization is strongly complete for K: every consistent set of formulas is jointly satisfiable in a Kripke model based on a frame in K.

It follows by Lemma B.3 that the axiomatization is also sound and strongly complete with respect to the intended semantics of $ML^\bullet$ (the proof of Lemma B.3 actually shows that every set of formulas jointly satisfiable on K is jointly satisfiable according to the intended semantics). □

It follows that $ML^\bullet$ also satisfies Compactness.

Lemma B.5 *The satisfiability problem for $ML^\bullet$ is* PSPACE-*complete.*

Proof: The lowerbound follows from the PSpace-hardness of basic modal logic. As for the upper bound, the standard PSpace-algorithm for basic modal logic due to Ladner can be extended straightforwardly to $ML^\bullet$ (using the alternative semantics provided by Lemma B.3).

Alternatively, one can also perform a direct reduction: for any $ML^\bullet$ formula ϕ, let ϕ^* be the uni-modal formula obtained by replacing each subformula of the form $\bullet\psi$ by $\Diamond(r \wedge \psi)$, for r a (fixed) fresh proposition letter. Let $Sub(\phi)$ be the set of subformulas of ϕ, and let $depth(\phi)$ be the modal nesting depth of ϕ.

It can be shown that, for any $\phi \in ML^\bullet$, ϕ is satisfiable on K iff

$$\phi^* \wedge \bigwedge_{\substack{k \leq depth(\phi) \\ \psi \in Sub(\phi)}} \Box^k(r \wedge \psi \rightarrow \Diamond\psi)$$

is a satisfiable basic modal formula. Roughly, the idea is that the additional conjuncts in the above formula guarantee that we can make every r-world reflexive in the model without affecting the truth of ϕ. Together with Lemma B.3 this gives us a polynomial reduction from $ML^\bullet$-satisfiability to satisfiability of basic modal formulas. □

Proposition B.6 *$ML^\bullet$ has Craig interpolation*

Proof: Follows from a result in [9], according to which every (multi-)modal logic axiomatized by Sahlqvist formulas with first-order universal Horn correspondents has Craig interpolation. The two axioms from Proposition B.4 are precisely of this form.

Note that, in this setting, we define Craig interpolation in terms of the proposition letters, not the modalities. □

Note that $ML^\bullet$ is not invariant for bisimulations.

C Proof of the Lindström theorem for the Binary Guarded Fragment

Before proving Theorem 3.17, we start with some modal features that hold for GF in general. First, there is a natural syntactic notion of formula depth, whose inductive definition counts the above polyadic quantifiers as single steps:

$$\begin{aligned}
&depth(Px) &&= 0\\
&depth(\neg\phi) &&= depth(\phi)\\
&depth(\phi \vee \psi) &&= \max(depth(\phi), depth(\psi))\\
&depth(\exists\vec{y}(G(\vec{x},\vec{y} \wedge \phi(\vec{x},\vec{y}))) &&= depth(\phi) + 1
\end{aligned}$$

Next, we define distance for points in models $(M, \vec{s})$, where $\vec{s}$ is a tuple of worlds:

> $dist(\vec{s}, s_i, 0)$ holds for $s_i \in \vec{s}$, and $dist(s, t, n{+}1)$ holds if there is a u with $dist(s, u, n)$ and $G(\vec{v})$ holds for some atomic predicate G and tuple of objects $\vec{v}$ containing t and u.

We write $Cut((M, \vec{s}), n)$ for the submodel $\{t \in (M, \vec{s}) \mid dist(\vec{s}, t, n)\}$ consisting of all points t in M lying at distance at most n from s. The following result shows that GF, like basic modal logic, satisfies a *finite depth property*, suitably defined:

Lemma C.1 (Distance-Depth Lemma) *Let ϕ be any guarded formula of depth n, and let $(N, \vec{s})$ be any submodel of $(M, \vec{s})$ containing all of $Cut((M, \vec{s}), n)$. Then $(M, \vec{s}) \models \phi$ iff $(N, \vec{s}) \models \phi$.*

Next comes a generalization of modal bisimulation. A guarded bisimulation is a non-empty set F of finite partial isomorphisms between two models M and N which has the following back-and-forth conditions. Call a set of objects '*guarded*' if some tuple containing these objects stands in some atomic relation. Now, given any function $f : X \rightarrow Y$ in F, (i) for any guarded $Z \subseteq M$, there is a $g \in F$ with domain Z such that g and f agree on the intersection $X \cap Z$, (ii) for any guarded $W \subseteq N$, there is a $g \in F$ with range W such that the inverses g^{-1} and f^{-1} agree on $Y \cap W$.

Also, '*rooted*' guarded bisimulations F run between models $(M, \vec{s})$ and $(N, \vec{t})$ with given initial objects, where one requires that some match between $\vec{s}, \vec{t}$ is already a partial isomorphism in F. By a simple inductive argument,

Fact C.2 *GF-formulas are invariant for rooted guarded bisimulations.*

Andréka, van Benthem and Németi [1] show that GF consists, up to logical equivalence, of just those first-order formulas which are invariant for guarded bisimulations.

Another 'modal' use of guarded bisimulation in the same paper is model unraveling. This is like standard modal unraveling, but the construction is a bit more delicate:

Definition C.3 (Tree unravelings for GF) *The tree unraveling* $\mathsf{unr}_{GF}(M)$ *of* M *has for its objects all pairs* (π, d) *where the 'path'* π *is a finite sequence of guarded sets in* M*, and the* M*-object* d *is 'new' in* π*: it occurs in the final set of* π *but not in the one before that. The interpretation of predicate symbols* Q *on these objects* (π, d) *is as follows.* $I(Q)$ *holds for a finite sequence of objects* $\langle(\pi_1, d_1), \ldots, (\pi_k, d_k)\rangle$ *iff* $Q^M(d_1, \ldots, d_k)$ *and there is some maximal path* π^* *among those listed of which all other* π_i *are initial segments in such a way that their new objects* d_i *remain present in each set until the end of* π^**.*

For a model $(M, \vec{s})$ *this generalizes as follows.* $\mathsf{unr}_{GF}(M, \vec{s})$ *has paths* π *all starting from the initial set* $\vec{s}$*, but then continuing with guarded sets only. The objects* (π, d) *are defined as before.*

The point here is that the set F of all restrictions of the finite maps sending (π_i, d_i) to d_i for all guarded finite domains in $\mathsf{unr}_{GF}(M, \vec{s})$ is a rooted guarded bisimulation between $(M, \vec{s})$ and $\mathsf{unr}_{GF}(M, \vec{s})$. Checking the zigzag conditions for the bisimulation will reveal the reason for the above definition of the predicate interpretations $I(Q)$.

Now we have the generalities in place for our Lindström Theorem, but it remains to make some adjustments. Firstly, the notion of guarded bisimulation needs to be slightly adjusted. GF_{bin}*-bisimulations* are defined like guarded bisimulations, except that the back-and-forth conditions are only required to hold when $X \cap Z \neq \emptyset$ and $Y \cap W \neq \emptyset$, respectively. It can be shown that GF_{bin} is precisely the GF_{bin}-bisimulation invariant fragment of FO (over vocabularies consisting of unary and binary relation symbols only).

Secondly, in the definition of tree unravelings, we make one simple change for the binary case:

> The finite paths of guarded sets always introduce one new object at each stage. At each step, one takes a new object related to that new object.

This allows paths starting with object a and then continuing with Rab, Qcb, $\ldots$, while ruling out paths like Rab, Qac. But the final atom is not omitted from the unraveled model, since one can have paths starting with a and then placing Qac immediately. Thus, even with these restricted paths, we still have a GF_{bin}-bisimulation between tree unravelings and their original models. The real point of this adjustment is the following.

The definition of predicates for path objects makes binary relations hold only between objects $(\pi_1, d_1), (\pi_2, d_2)$ where π_2 is a one-step continuation of the path π_1, or vice versa. But then, counting distance as before,

> The new object at the end of a path of length k lies at distance k from the initial object of the path.

Put in more vivid terms, '*tree distance is true distance*' in the original model. This is a non-obvious fact. E.g., with ternary guards $Rayz$, objects at the end of a path may keep links to the initial object a which might recur in the guarded sets along the path (this observation is due to Martin Otto).

We will write $\mathsf{unr}_{GF_{\mathsf{bin}}}(M)$ and $\mathsf{unr}_{GF_{\mathsf{bin}}}(M, \vec{s})$ for this new type of unravelling.

Proof of Theorem 3.17: Let $\mathcal{L}$ be any abstract logic extending GF and satisfying Compactness and invariance for guarded bisimulations. As before, it suffices to show that every formula $\phi \in \mathcal{L}$ is invariant for models that are equivalent for all GF_{bin}-formulas up to some finite depth n.

First, largely as in the earlier modal proof of Sect. 3.1, we use the Compactness of $\mathcal{L}$, together with its *Relativization closure*, to show that ϕ must have the *Finite Occurrence Property* and a *Finite Distance Property* for some level n. Before, universal prefix formulas $\Box^k p$ (for all finite k) made sure that p holds in the generated submodel at the current world. This time, one uses all nested sequences of universal guarded quantifiers up to depth k, requiring that some new predicate P holds for all objects reached at the end. The n thus found for the local depth of the formula ϕ is the same n as needed for the following semantic invariance:

Given the above unraveling construction and invariance for guarded bisimulation for $\mathcal{L}$, we may assume, without loss of generality, that we have the following situation:

(a) $\mathsf{unr}_{GF_{\mathsf{bin}}}(M, \vec{s}) \models \phi$

(b) $\mathsf{unr}_{GF_{\mathsf{bin}}}(M, \vec{s})$ and $\mathsf{unr}_{GF_{\mathsf{bin}}}(N, \vec{t})$ satisfy the same GF_{bin}-formulas up to depth n

Our aim is to show that $(\mathsf{unr}_{GF_{\mathsf{bin}}}(N, \vec{t})) \models \phi$.

We cut the tree models to depth n, as before in our modal argument, obtaining $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(M, \vec{s}), n)$, $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(N, \vec{t}), n)$. Since tree depth is true depth, this does not change truth values of ϕ in either model.

Finally, we define a GF_{bin}-bisimulation between $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(M, \vec{s}), n)$ and $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(N, \vec{t}), n)$. First of all, note that all elements in the first model are of the form $(\langle\{\vec{s}\}, \{d_1, d_2\}, \ldots, \{d_{k-1}, d_k\}\rangle, d_k)$ with $d \in \vec{s}$ and $0 \leq k \leq n$) where each set in the sequence is guarded. Likewise for the second model. Secondly, note that the only guarded subsets in these models are singleton sets and sets of the form $\{(\pi, d), (\langle\pi, \{d, e\}\rangle, e)\}$. Let Z be the set consisting of $(\{\vec{s}\}, \{\vec{t}\})$, as well as all sets of pairs

$$\{(\{(\pi, d), (\pi', d')), ((\langle\pi, \{d, e\}\rangle, e), (\langle\pi', \{d', e'\}\rangle, e'))\}$$

with π, π' sequences of some length $k \leq n$ for which it holds that, for all $i \leq k$, the guarded sets π_i and π'_i satisfy the same GF_{bin}-formulas of quantifier depth $n - i$.

It can be shown (the argument involves a case distinction) that Z is a GF_{bin}-bisimulation between $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(M, \vec{s}), n)$ and $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(N, \vec{t}), n)$. Hence, the truth of ϕ transfers from $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(M, \vec{s}), n)$ to $Cut(\mathsf{unr}_{GF_{\mathsf{bin}}}(N, \vec{t}), n)$, and hence to $(N, \vec{t})$. $\Box$