Model Theoretic Methods for Fragments of FO and Special Classes of (Finite) Structures

Martin Otto
Technische Universität Darmstadt
otto@mathematik.tu-darmstadt.de

October 31, 2008

Abstract

Some prominent fragments of first-order logic are discussed from a game-oriented and modal point of view, with an emphasis on model theoretic techniques for the non-classical context of finite model theory or of other natural non-elementary classes of structures. We stress the modularity and compositionality of the games as a key ingredient in the exploration of the expressive power of logics over specific classes of structures. The leading model theoretic theme is expressive completeness – or the characterisation of fragments of first-order logic as expressively complete over some class of (finite) structures for first-order properties with some prescribed semantic preservation behaviour. In contrast with classical expressive completeness arguments, the emphasis here is on explicit model constructions and transformations, which are guided by the game analysis of both first-order logic and of the imposed semantic constraints.

keywords: finite model theory, model theoretic games, bisimulation, modal and guarded logic, expressive completeness, preservation and characterisation theorems

Contents

1 Introduction

1.1 Expressiveness over restricted classes of structures

The purpose of this survey is to highlight game-oriented methods and explicit model constructions for the analysis of fragments of first-order logic, in particular in restriction to non-elementary classes of structures. The following is meant to highlight and preview some key points in terms of both the material to be covered and the perspective that we want to adopt in its presentation. All these points will be addressed in a more self-contained manner in the technical sections; an outline of the structure of the technical sections concludes this preview.

Varying the class of structures The class of all finite structures is one prominent non-elementary class of interest, but recent developments in finite model theory have broadened the perspective. While the first tier of results in finite model theory, which set the stage and clarified much of the specifics of finite model theory, brought predominantly negative results ('failures' in comparison to classical model theory, the first and foremost being the 'failure of compactness' in finite model theory), a much more positive picture has emerged with a focus on specific classes of well-behaved finite structures rather than the class of all finite structures (cf. Weinstein's *tame fragments and tame classes* [48]). What good behaviour means for classes of structures, may of course depend on the model theoretic issue at hand. Nevertheless, there are some interesting recurring themes, revolving around tree-likeness on the one side and locality criteria on the other side, in delineating *well-behaved classes* of (finite) structures.

Expressiveness and expressive completeness Our leading model theoretic theme in terms of results is that of expressive completeness. We regard expressive completeness results as classical hallmarks in the study of expressiveness of fragments of first-order logic. Think of a classical example like the Łos–Tarski existential preservation theorem (cf. Theorem 5.2) that a first-order formula is preserved under extensions if, and only if, it is logically equivalent to an existential formula. The preservation claim in this statement – that existential formulae are preserved under extensions – is a trivial exercise in syntactic induction, and its truth carries over to any restricted class of structures. The expressive completeness statement – that within first-order, the existential fragment is *expressively complete* for properties preserved under extensions – requires real model theoretic proof. The classical proof in [10] uses elementary extensions, whose availability hinges on the use of the compactness theorem for first-order logic. So that proof does not relativise to arbitrary restricted classes, and in fact the relativisation to the class of all finite structures is a typical example of a 'failure in finite model theory' (due to Tait and Gurevich, see for instance [14]). Preservation of a first-order property under extensions among finite structures does not imply expressibility in existential first-order logic over finite structures. Some instances of classical preservation theorems, like Łos–Tarski, fail in retsriction to the class of all finite structures, but are true – with totally new proofs – in interesting restricted classes of finite structures (cf. Theorem 5.9 for results pertaining to extension preservation, from [3]). Other instances, like van Benthem's theorem concerning preservation under bisimulation (cf. Corollary 3.5), or, more classically, the Lyndon–Tarski theorem (cf. Theorem 5.3), which associates preservation under homomorphisms with the existential positive fragment, do have literal analogues in restriction to the class of all finite structures as well as to some other restricted classes of structures of interest (cf. sections 3.2 and 5.2.3) – with new proofs that do not draw on the classical proofs but shed interesting new light on the classical results as well. And in some few instances we know of expressive completness results over restricted classes of (finite) structures that require more expressive fragments than the classical analogue; a recent example concerning bisimulation preservation is discussed in section 3.2.

Explicit model constructions and transformations Compactness, and with it many of the typical model constructions prevalent in classical expressive completeness results, are typically not available over the restricted classes of structures under consideration. Where expressive completeness results can be obtained over non-elementary classes, the methods are very different from the classical ones. The technical crux of many expressive completeness results, classical or otherwise, consists in an *upgrading* of transfer or equivalence relations between structures. For instance, in the case of preservation under some equivalence relation $\leftrightharpoons$ like bisimulation associated with expressibility in the fragment L: here preservation under $\leftrightharpoons$ must be linked to preservation under finitary approximations $\leftrightharpoons^{\ell}$ to L-equivalence, finitary in the sense of finite index and in the sense that its classes are L-definable (think of approximations parameterised, e.g., by quantifier rank ℓ). As these finitary approximations $\leftrightharpoons^{\ell}$ are rougher than full $\leftrightharpoons$, the task of showing that every first-order property φ preserved under $\leftrightharpoons$ is even preserved under some $\leftrightharpoons^{\ell}$, involves model theoretic transformations that allow us to boost $\leftrightharpoons^{\ell}$ either to $\leftrightharpoons$ or to some other equivalence under which φ is preserved (e.g., on account of being first-order of a certain quantifier rank). The classical treatment of the Łos–Tarski theorem, for instance, can similarly be viewed as an upgrading of a transfer relationship $\mathfrak{A} \Rrightarrow_{\exists} \mathfrak{B}$ (existential sentences true in $\mathfrak{A}$ are also true in $\mathfrak{B}$), or of its finitary approximations, to a substructure relationship between elementarily equivalent companion structures of $\mathfrak{A}$ and $\mathfrak{B}$. (In this case, $\mathfrak{B}$ admits an elementary extension that embeds $\mathfrak{A}$ as a substructure, by compactness.) It follows that any first-order φ preserved under extensions is preserved under $\Rrightarrow_{\exists}$, and – by another compactness argument – therefore also under some finite quantifier rank approximation $\Rrightarrow_{\exists}^{\ell}$ to $\Rrightarrow_{\exists}$.

As will be discussed in section 3.2, such upgrading arguments tend to proceed in orthogonal directions of entirely different character, depending on whether they are based on classical compactness arguments (often involving elementary chains and saturation) or on explicit and finitary model transformations, which may also be carried out within some restricted, non-elementary class of structures like the class of just all finite structures. Explicit model constructions and transformations can thus sometimes replace the sweeping classical compactness arguments that guarantee the existence of nice and smooth (but typically infinite) representatives of the structures at hand, in which crucial technicalities (e.g., back-and-forth arguments) can be dealt with more elegantly. But there is also something to be gained, even from the classical point of view, from the more explicit, more controlled and more constructive nature of the alternative model transformations: in key examples of expressive completeness results to be discussed below, for instance, bounds on the quantifier rank of the target formulae are an integral part of the proofs based on explicit model constructions and transformations. In this sense, the alternative approach, which is necessitated by the loss of compactness in finite model theory, can offer a new perspective and sometimes extra information on classical results.

Model theoretic games The equivalences and transfer relations between structures underlying semantic preservation properties on the one hand, and logical equivalences or transfer relations induced by fragments of first-order logic on the other hand, are closely linked to model theoretic games or back-and-forth systems. As pointed out above, upgrading arguments between these equivalences and suitable finitary approximations, which are themselves naturally cast as game equivalences, play a crucial role in expressive completeness proofs. The methodological importance of model theoretic games, both to understand the semantics and expressive power of logics and to guide the desired explicit model constructions or transformations (over restricted classes of structures), is being put at the centre of this presentation. We shall here especially discuss variants of the classical Ehrenfeucht–Fraïssé game and the first-order model checking game for several fragments of first-order logic. A prominent place among these variants is given to the modal Ehrenfeucht–Fraïssé game, or bisimulation game. In section 3, bisimulation games and model transformations that respect bisimulation feature prominently in the discussion of expressive completeness results for modal logics over various classes of Kripke structures. Also locality of first-order logic in the sense of Gaifman's theorem (cf. Theorem 2.13) is presented in terms of the modularity of the first-order Ehrenfeucht–Fraïssé game w.r.t. locality in the Gaifman graph. Locality-based approximations to first-order equivalence also play a role in some of the expressive completeness results for modal logics, or in the upgrading between approximate levels of bisimulation and first-order equivalence. Structurally, the concept of locality will also be important in connection with classes of structures defined in terms of wideness criteria in section 5.

Bisimulation as the game of games Putting games – model checking games that define the semantics of a logic and Ehrenfeucht–Fraïssé model comparison games – at the centre of the analysis of fragments of first-order logic, it becomes very natural to adopt a modal perspective [8, 9] and to relate other fragments and their games to the bisimulation game. We thus draw on bisimulation games and bisimulation equivalence not just in the study of modal fragments but also on its role as an equivalence between game graphs that encapsulate the semantics of other fragments. The connection is made by looking at the natural game graphs associated with model checking games or Ehrenfeucht–Fraïssé games as Kripke structures. The elements of these Kripke structures are formed by the *observable configurations* in the underlying structures, their accessibility relations reflect the transitions between game positions, which in turn reflect the available quantification patterns of the fragment at hand. For the modal fragment itself, the structure (Kripke structure, transition system) *is* its own game graph, in which the elements can be navigated along the edges (of the given accessibility or transition relation). Richer fragments have access to more complex types of configurations within structures and possibly more complex rules for navigation between configurations. For instance, in the k-variable fragment $\mathrm{FO}^k \subseteq \mathrm{FO}$ we deal with arbitrary configurations consisting of up to k elements, while in the guarded fragment $\mathsf{GF} \subseteq \mathrm{FO}$ the configurations need to be covered by some relational ground atom. This view may not directly offer new technical insights, but has the advantage of making explicit a unifying and, I think, intuitive framework whose specialisations to individual fragments are of course very well understood.

Structure of the paper The overall structure of the paper is as follows. In section 2 we review model checking and model comparison games for FO and some of its fragments from a modal perspective; we also discuss Gaifman locality in relation to the FO Ehrenfeucht–Fraïssé game. Section 3 deals with expressive completeness issues for modal logics over specific classes of transition systems. The extension of the concept of bisimulation from graphs to hypergraphs, its relationship with the guarded fragment and a connection with extension properties for partial automorphisms is discussed in section 4. In section 5 we turn to locality based techniques for special classes of relational structures, and to expressive completeness for preservation under extensions and homomorphisms.

Sections 2 and 3 are meant to be fairly expository, and may serve either as a brief introduction to the fragments and methods discussed, or as an invitation to re-discover some rather familiar concepts in a slightly different light. Sections 4 and 5 are more technical and also less self-contained. To a large extent they may, on the other hand, also be considered independently of the first part. The intention is to give at least some high-level account of some more recent results and developments in the framework of this survey.

1.2 Basic terminology and notational conventions

Structures and assignments Throughout we only consider relational structures. Typically τ will be a finite relational signature, and we refer to the maximal arity of relations in τ as its *width*. A τ-structure with universe A will usually be denoted as $\mathfrak{A} = (A, (R^{\mathfrak{A}})_{R\in\tau})$, but we often omit superscripts where the structure is clear from context.

Within a τ-structure $\mathfrak{A}$, we look at (partial) assignments (to an official set of first-order variables $x_1, x_2, \ldots$), described by partial functions $\beta\colon (x_i) \to A$. Assignments to finite tuples of variables are often regarded as momentarily fixed parameter tuples, like $\mathbf{a} = (a_1, \ldots, a_k) \in A^k$ as an assignment $\beta\colon (x_i \mapsto a_i)_{i=1,\ldots,k}$. Such (finite) assignments will also play a role in games as *configurations* (tuples of marked elements) within a structure, often directly associated also with the substructure induced on the subset $[\mathbf{a}] := \{a_1, \ldots, a_k\} \subseteq A$. Because we do not want to clutter terminology with a fine distinction between tuples and assignments, we also think of assignments (which officially are assignments to variables x_i) as partial functions $\beta\colon i \mapsto \beta(i)$ over a domain of positive natural numbers. Notation for modifications of assignments is as in $\beta\frac{a}{i}$, for the assignment obtained by changing (or extending) β at i (at x_i) to take the value a. For the semantics of formulae $\varphi(\mathbf{x})$ with free variables among those listed in the tuple $\mathbf{x}$, notations $\mathfrak{A}, \mathbf{a} \models \varphi$, $\mathfrak{A} \models \varphi[\mathbf{a}]$, and $\mathfrak{A}, \beta \models \varphi$ are used interchangeably, if β is an assignment to (at least) the free variables of φ and assigns $\mathbf{a}$ to $\mathbf{x}$.

Among important specific types of structures we mention the following to clarify terminology. Other more specific classes of structures will be introduced at appropriate places.

Directed and undirected graphs are structures over relational vocabularies of width 2, i.e., we admit several binary relations (edge-labelled directed graphs) and unary predicates (vertex colours). More traditional plain directed graphs are a special case, with just a single binary edge relation. We also view directed edge-labelled and vertex-coloured graphs as *transition systems*, with several transition relations and atomic state predicates. Such transition systems are just a terminological variant of *Kripke structures*, as the structures for modal logics. Undirected graphs are graphs with a single edge relation that is symmetric and irreflexive, viewed as a special case of directed graphs.

A *(directed) tree* is a directed graph that has a *root* w.r.t. the union of its binary relations such that every other element is reachable on a unique edge-labelled directed path from this node. Note that this implies irreflexivity (no loops), antisymmetry (no edges in opposite directions, not even with different labels) and that there are no multiple edges (with different labels). More generally, a directed graph or transition system is called *simple* if it has no loops and no multiple edges (not even in opposite directions).[1]

Hypergraphs, which are at the centre of section 4, are not regarded as relational structures but as second-order structures of the format $H = (A, S)$ with a universe A and a subset of the power set $S \subseteq \mathcal{P}(A)$ as the set of hyperedges. We shall encounter hypergraphs as auxiliary combinatorial structure, induced by relational structures, but will not look at logics over hypergraphs.

Gaifman graph and distance With any structure in a finite relational vocabulary τ we associate an undirected graph, its Gaifman graph.

Definition 1.1. The *Gaifman graph* of the τ-structure $\mathfrak{A}$ is the undirected graph $G(\mathfrak{A}) = (A, E^{G(\mathfrak{A})})$ with the same universe A and an edge $(a, b) \in E^{G(\mathfrak{A})}$ for $a \neq b$ if a and b occur together in some tuple within some relation $R^{\mathfrak{A}}$, $R \in \tau$.

The associated notion of *Gaifman distance* is just ordinary graph distance (minimal length of a connecting path, or infinity) between elements in $G(\mathfrak{A})$. We denote this distance as $d(\cdot\,,\cdot)$. Finite distance relations like $d(x, y) \leqslant k$ are clearly FO-definable in $\mathfrak{A}$. In graphs (τ finite and of width 2), $d(x, y) \leqslant 1$ is quantifier free definable, while in general the required quantifier rank is the width of τ minus 2. An easy induction shows that $d(x, y) \leqslant 2^q$ is definable by a first-order formula $\varphi(x, y)$ for any finite τ.

Definition 1.2. The *Gaifman neighbourhood of radius ℓ*, or *ℓ-neighbourhood* for short, of an element a in $\mathfrak{A}$ is the subset $N^\ell(a) = \{b \in A : d(a, b) \leqslant \ell\} \subseteq A$. By extension, the ℓ-neighbourhood of a tuple $\mathbf{a} = (a_1, \ldots, a_k)$ in $\mathfrak{A}$ is the union of the $N^\ell(a_i)$.

A subset (or tuple) in $\mathfrak{A}$ is *ℓ-scattered* if its elements (or components) have pairwise distance greater than 2ℓ (i.e., if their ℓ-neighbourhoods are disjoint).

By the above considerations, ℓ-neighbourhoods of tuples, or the property of a tuple to be ℓ-scattered, are all first-order definable, for every $\ell \in \mathbb{N}$ and for any fixed finite τ.

A relational structure is called *acyclic* if its Gaifman graph is acyclic; for directed graphs as relational structures, this is different from the usual notion which only forbids directed cycles.

A directed graph or transition system is *ℓ-acyclic* if its Gaifman graph is acyclic in every ℓ-neighbourhood (this rules out *undirected* cycles of lengths up to $2\ell + 1$).

[1] In section 3.2.3 we also discuss transitive tree structures, which are trees in the partial order sense, not in the graph sense, but that will be highlighted there.

Logics We write FO for first-order logic, or more specifically FO[τ] for the set of first-order formulae over vocabulary τ. The set of free variables of a first-order formula φ is denoted free(φ). Notation as in $\varphi = \varphi(\mathbf{x})$ indicates that $\mathrm{free}(\varphi) \subseteq [\mathbf{x}]$ (the set of variables listed as components of the tuple $\mathbf{x}$).

Quantifier-rank is defined as usual for first-order formulae, and denoted qr(φ). Atomic and quantifier-free types of tuples $\mathbf{a}$ in a τ-structure $\mathfrak{A}$ provide full descriptions of $\mathbf{a}$ at the quantifier-free level. Formally we may define the atomic type of $\mathbf{a}$ (in a matching tuple of variables, so that $\beta\colon \mathbf{x} \mapsto \mathbf{a}$ is appropriate as an assignment) as the set of all atomic and negated atomic formulae $\alpha(\mathbf{x})$ in variables $\mathbf{x}$ for which $\mathfrak{A} \models \alpha[\mathbf{a}]$. It is clear that the correspondingly defined quantifier-free type is fully determined by the atomic type, and that both can be summarised by a single quantifier-free formula in case τ is finite. The atomic or quantifier-free type of $\mathbf{a}$ in $\mathfrak{A}$ fully determines the isomorphism type of $\mathfrak{A}{\restriction}[\mathbf{a}]$ (of configuration $\mathbf{a}$ in $\mathfrak{A}$).

FO^k stands for the k-variable fragment of FO, which uses only the variable symbols $x_1, \ldots, x_k$. The finite variable fragments have played a very prominent role in the development of finite model theory as witnessed for instance in [14, 32]; we shall not focus on these fragments very much here, but treat the associated k-pebble games as a typical and natural example in the exposition of section 2.

Apart from fragments of FO, we occasionally look at its infinitary extension FO_∞ (classically denoted $L_{\infty\omega}$), which extends the syntactic framework of FO by allowing disjunctions and conjunctions over arbitrary sets of formulae. Connectedness of graphs, for instance, becomes definable in FO_∞ with the use of an infinite disjunction to express "$d(x, y) < \infty$" as "$\bigvee\{d(x, y) \leqslant n\colon n \in \omega\}$". Formulae in FO_∞ have ordinal quantifier rank, defined by the usual inductive clauses extended by taking suprema for infinite disjunctions or conjunctions. The quantifier-rank of the formula "$d(x, y) < \infty$" would thus be ω, that of the natural sentence defining connectivity $\omega + 2$. Similar infinitary extensions naturally arise, e.g., for the modal fragment to be discussed next.

Basic modal logic is denoted ML, or $\mathsf{ML}[\tau]$ for a given vocabulary of width 2 appropriate for transition systems (Kripke structures). We typically use a τ with binary transition relations E_α (regarding the indices α as edge labels) and unary predicates P_j (associated to atomic state properties or atomic propositions p_j). The formulae of $\mathsf{ML}[\tau]$ are generated from the atomic propositions p_j by means of boolean connectives and modal quantifications with $\Diamond_\alpha$ or $\Box_\alpha$. The defining clause for the semantics of $\varphi = \Diamond_\alpha\psi$, say at a state a in a τ-structure $\mathfrak{A}$, is

$$\mathfrak{A}, a \models \varphi \qquad \text{iff} \quad \mathfrak{A}, b \models \psi \text{ for some } b \text{ such that } (a, b) \in E_\alpha^{\mathfrak{A}},$$

and dually for $\Box_\alpha\psi$, which is equivalent to $\neg\Diamond_\alpha\neg\psi$. We also view $\mathsf{ML}[\tau]$ as a fragment of FO[τ], having only formulae in one free variable, via the standard translation that associates p_j with $P_j x$ and $\Diamond_\alpha\psi$ with $\exists y(R_\alpha xy \wedge \psi(y))$ so that, dually, $\Box_\alpha\psi$ is associated with $\forall y(R_\alpha xy \rightarrow \psi(y))$. This is briefly reviewed in connection with the model checking game for modal logic in section 2.3.3.

The extension of basic modal logic with modal quantification backward along E_α (*inverse modalities*) is denoted ML^{-}; the extension by a *global modality*, corresponding to the introduction of modal quantification associated with the full binary relation, is denoted $\mathsf{ML}^{\forall}$; the combined extension with both these additions is $\mathsf{ML}^{-\forall}$. For background in connection with our treatment of modal logics and much more material on the model theory of modal logics see in particular [16].

The *guarded fragment* GF is defined to be a syntactic fragment of FO consisting of formulae in which all quantifications are relativised as in

$$\begin{aligned} \varphi(\mathbf{x}) &= \exists \mathbf{y}(\alpha(\mathbf{x}') \wedge \psi(\mathbf{x}')), \text{ or} \\ \varphi(\mathbf{x}) &= \forall \mathbf{y}(\alpha(\mathbf{x}') \rightarrow \psi(\mathbf{x}')), \end{aligned}$$

where $\alpha(\mathbf{x}')$ is an atomic τ-formula (a relational atom, or an equality: the *guard atom*) such that $\text{free}(\psi) \subseteq \text{var}(\alpha)$ (and $\mathbf{y}$ is a sub-tuple of $\mathbf{x}'$ such that $[\mathbf{x}'] \setminus [\mathbf{y}] \subseteq [\mathbf{x}]$).

The quantification pattern of guarded logic extends that of modal logic. For a modal vocabulary τ, $\mathsf{GF}[\tau]$ properly contains (the standard first-order translations of) $\mathsf{ML}[\tau]$ and even $\mathsf{ML}^{-\forall}[\tau]$. One motivation for the study of the guarded fragment stems from the analogy with modal logic, and the extension of modal quantification patterns from Kripke structures to more general relational structures. Guarded fragments were proposed in [2] with a view to explaining the good algorithmic and model theoretic properties of modal logics in a richer fragment of first-order logic and other than the 2-variable fragment [23]; see [20]. In many ways the guarded fragment has been shown to be a rather well-behaved intermediary between first-order and modal logic, in terms of its model theoretic and algorithmic properties. For instance (like modal logic and unlike FO^k for $k \geqslant 3$), GF has the finite model property and is decidable: the satisfiability problem for $\mathsf{GF}[\tau]$ is complete for deterministic exponential time if τ is fixed (more precisely, for any fixed bound on the width of τ), and complete for doubly exponential time without this constraint [20]. Similarly to the tree model property of modal logic (which is a consequence of bisimulation invariance and the model transformation of tree unfolding, see in particular section 3.1.1), GF has a generalised tree model property, which similarly stems from invariance under guarded bisimulation and the availability of guarded tree unfoldings. For these considerations we refer to the discussion in section 4.2, where we interpret these phenomena in the light of a generalisation of bisimulations from graphs to hypergraphs. For further results concering the model theory of GF and some of its generalisations see [20, 30, 22, 24, 7, 31] among many others.

The semantics of the above-mentioned fragments, though assumed familiar, will be reviewed again in section 2.3 when we discuss the associated model checking games. There we shall proceed in the order of increasing specialisation, from FO to FO^k to GF to (variants of) ML.

2 Model theoretic games and bisimulation

As mentioned above, we adopt a non-standard perspective of looking at first-order logic (and some of its fragments) through modal eyes. Connections are made through games, at two levels: at the level of *model checking games*, which capture the semantics, and at the level of *model comparison games*, which capture degrees of logical indistinguishability between structures.

No technical knowledge of model checking games and Ehrenfeucht-Fraïssé games is assumed. The reader who has some familiarity with model checking games and the

Ehrenfeucht-Fraïssé technique for various fragments and extensions of FO on the other hand, will recognise the familiar notions in a slightly different prespective.

2.1 The semantic game: verifier vs. falsifier

We take a look at the first-order model checking game from a modal point of view. We shall then want to present some fragments of first-order logic in terms of restricted game boards; the same view will uniformly be applied to the Ehrenfeucht–Fraïssé model comparison games in the next section.

A transition system of observable configurations With the relational vocabulary τ associate the vocabulary τ^* consisting of binary transition relations E_i for $i \geqslant 1$ and unary predicates P_θ for atomic τ-types $\theta = \theta(\mathbf{x})$ in finite tuples of variables from $(x_i)_{i \geqslant 1}$. With a τ-structure $\mathfrak{A}$ associate the following τ^* transition system $\mathcal{O}(\mathfrak{A})$ of *observable configurations* over $\mathfrak{A}$:

– the universe of $\mathcal{O}(\mathfrak{A})$ is the set of partial assignments to variables $(x_i)_{i \geqslant 1}$;

– E_i is interpreted as $\{(\beta, \beta\frac{a}{i})\colon a \in A\}$ (modifications of assignments at x_i);

– P_θ as the set of assignments β satisfying θ (in particular $\mathrm{var}(\theta) \subseteq \mathrm{dom}(\beta)$).

In a straightforward manner one obtains a uniform translation from $\mathrm{FO}[\tau]$ over $\mathfrak{A}$ to $\mathsf{ML}[\tau^*]$ over the associated $\mathcal{O}(\mathfrak{A})$. This translation,

$$\begin{array}{ccc} \mathsf{FO}[\tau] & \longrightarrow & \mathsf{ML}[\tau^*] \\ \varphi(\mathbf{x}) & \longmapsto & \varphi^*, \end{array}$$

is such that for all β with $\mathrm{free}(\varphi) \subseteq \mathrm{dom}(\beta)$:

$$\mathfrak{A}, \beta \models \varphi \quad \Leftrightarrow \quad \mathcal{O}(\mathfrak{A}), \beta \models \varphi^*.$$

At the quantifier-free level, $\varphi = \varphi(\mathbf{x})$ translates into

$$\varphi^* := \bigvee \{P_\theta \colon \varphi \in \theta, \mathrm{var}(\theta) = \mathrm{var}(\varphi)\};$$

the translation is compatible with boolean connectives; and existential quantification translates into a modal diamond in a natural manner, as in

$$\varphi = \exists x_i \psi(\mathbf{x}) \quad \longmapsto \quad \varphi^* = \Diamond_i \psi^*.$$

Note that the modal vocabularies involved are a priori infinite; this can be avoided if we restrict attention to the k-variable fragment $\mathrm{FO}^k[\tau]$ for fixed k and fixed finite relational vocabulary τ. In this case, there are only finitely many P_θ corresponding to atomic τ-types in variables $\mathbf{x} = (x_1, \ldots, x_k)$; we may restrict attention to full assignments to all the variables $\{x_1, \ldots, x_k\}$, which can be identified with A^n; and we just retain k transition relations E_i for $1 \leqslant i \leqslant k$. Further natural restrictions to be discussed in section 2.3 lead to modal and guarded logics.

The model checking game The idea to associate a two-person game with the semantics of first-order logic goes back at least to Lorenz' and Lorenzen's dialogue games [39, 40] between a proponent and an opponent of some assertion. The current interest in these games stems not from foundational issues but from their algorithmic content, or more precisely from their conceptual strengths towards the design of efficient model checking algorithms, see, e.g., [21, 47].

With formulae φ and τ-structures $\mathfrak{A}$ with partial assignments β we associate a game played by two players, V (verifier) and F (falsifier) such that the winning positions in the game determine whether or not $\mathfrak{A}, \beta \models \varphi$.

We present this basic and simple idea in a modular fashion that uses the transition system of observable configurations as one constituent of the game (representing the structure input to the model checking problem). The other constituent is essentially the syntax tree of the formulae to be checked (representing the formula input to the model checking problem). For a transparent account of the algorithmic content of this game, and its complexity analysis, compare [21].

Let $\Phi \subseteq \mathrm{FO}[\tau]$ be a set of negation normal form formulae that is closed under subformulae (negation normal form restricts the occurrence of negations to negated atoms). Let $\mathcal{S}(\Phi)$ be the transition system whose universe is Φ, with transition relations $E_\vee$, $E_\wedge$, $E_{\exists x_i}$ and $E_{\forall x_i}$ ($i \geqslant 1$) interpreted as follows.

$E_\vee$ contains the pairs (φ, φ_1) and (φ, φ_2) for $\varphi = \varphi_1 \vee \varphi_2 \in \Phi$; similarly for $E_\wedge$;
$E_{\exists x_i}$ consists of all pairs (φ, ψ) for $\varphi = \exists x_i \psi \in \Phi$; similarly for $E_{\forall x_i}$.

The game graph $\mathbb{G} := \mathbb{G}(\mathfrak{A}, \Phi)$ for the Φ model checking game over $\mathfrak{A}$ may then be interpreted in a subsystem of the product system

$$\mathcal{O}(\mathfrak{A}) \times \mathcal{S}(\Phi).$$

More specifically, the universe of $\mathbb{G}(\mathfrak{A}, \Phi)$ is the set of all syntactically appropriate assignment/formula pairs, $\{(\beta, \varphi) \colon \mathrm{free}(\varphi) \subseteq \mathrm{dom}(\beta)\}$. The relevant transition relations of $\mathbb{G}(\mathfrak{A}, \Phi)$ are

in $\mathbb{G}(\mathfrak{A}, \Phi)$		in $\mathcal{O}(\mathfrak{A})$		in $\mathcal{S}(\Phi)$	
$E_\vee$	$:=$	id	$\times$	$E_\vee$	(disjunctive moves)
$E_\wedge$	$:=$	id	$\times$	$E_\wedge$	(conjunctive moves)
$E_{i,\exists}$	$:=$	E_i	$\times$	$E_{\exists x_i}$	(existential moves)
$E_{i,\forall}$	$:=$	E_i	$\times$	$E_{\forall x_i}$	(universal moves)

As atomic predicates we use P_V and P_F, which partition the universe of $\mathbb{G}(\mathfrak{A}, \Phi)$ according to:

$$\begin{aligned} P_F^{\mathbb{G}} = \; & \big\{(\beta, \varphi) \colon \varphi = \varphi_1 \wedge \varphi_2 \text{ or } \varphi = \forall x_i \psi\big\} \\ & \cup \big\{(\beta, \varphi) \colon \; \varphi \text{ atomic or negated atomic, } \mathfrak{A}, \beta \models \varphi\big\}, \\ P_V^{\mathbb{G}} = \; & \big\{(\beta, \varphi) \colon \varphi = \varphi_1 \vee \varphi_2 \text{ or } \varphi = \exists x_i \psi\big\} \\ & \cup \big\{(\beta, \varphi) \colon \; \varphi \text{ atomic or negated atomic, } \mathfrak{A}, \beta \not\models \varphi\big\}. \end{aligned}$$

The rules of the game are then simply the following, according to which the players move a pebble in the game graph $\mathbb{G}$:

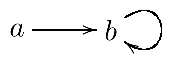
– Positions in P_V require a move by V:
V moves along any $E_\vee$- or $E_{i,\exists}$-edge (as available in current position);
V loses when stuck for a move.

– Positions in P_F require a move by F:
F moves along any $E_\wedge$- or $E_{i,\forall}$-edge (as available in current position);
F loses when stuck for a move.

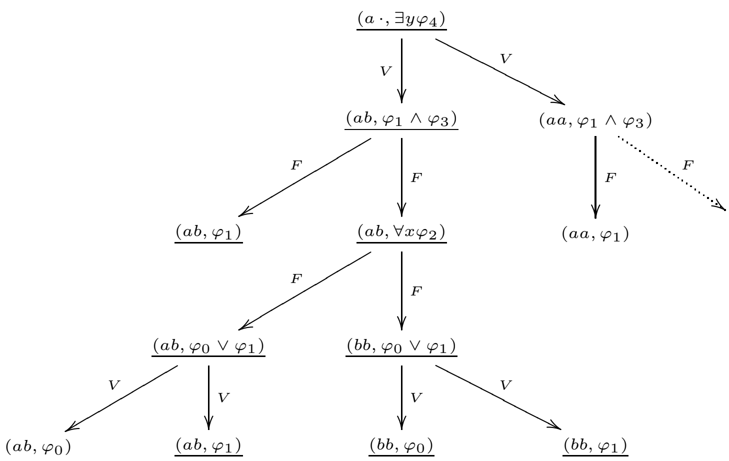
As formula complexity is strictly reduced in each move, all plays are finite. Positions in which neither player can move are terminal positions for the game and the player who ought to move has lost. This happens exactly in positions associated with atomic or negated atomic formulae, and here the attribution of these nodes to V and F is such that V wins (because F ought to move) if $\mathfrak{A}, \beta \models \varphi$, and vice versa. Clearly the game is positionally determined, and the following is proved by an easy induction on the structure of the formula (or on the length of the remaining game).

Lemma 2.1. *The verifier V has a winning strategy in the model checking game on $\mathfrak{A}$ precisely in those positions (β, φ) for which $\mathfrak{A}, \beta \models \varphi$.*

Let us sketch part of the game graph in one tiny example. For a binary relation R consider the formula $\varphi(x) = \exists y\big(Rxy \wedge \forall x(Rxx \vee Rxy)\big)$ over the R-structure $\mathfrak{A}$ with two elements a and b and with R-edges as indicated by arrows:



The model checking game to determine whether $\mathfrak{A} \models \varphi[a]$ has positions (β, ψ) where ψ is one of the subfurmulae of φ and β a (partial) assignment to variables x, y. We may represent β by an $\{a, b, \cdot\}$-word of length 2 and enumerate the subformulae ψ as $\varphi_0 := Rxx$, $\varphi_1 := Rxy$, $\varphi_2 = \varphi_0 \vee \varphi_1$, $\varphi_3 := \forall x\varphi_2$, $\varphi_4 := \varphi_1 \wedge \varphi_3$ such that $\varphi = \exists y\varphi_4$.



In this partial sketch of the game tree, winning positions for V are underlined.

There is a natural variant of the model checking game that does not restrict formulae to negation normal form. The transition corresponding to the elimination of a negation, say from $\neg\varphi$ to φ, corresponds to a swap of players' roles. Let us therefore call the players neutrally player 1 and player 0. Positions in the game graph are extended by an extra component $\wp \in \{0, 1\}$ to indicate which of the two players acts as verifier; the opponent, $\bar{\wp}$, correspondingly acts as falsifier. The two component games, $\mathbb{G}(\mathfrak{A}, \Phi) \times \{0\}$ and $\mathbb{G}(\mathfrak{A}, \Phi) \times \{1\}$ are each as before (but not insisting on negation normal form formulae, and with player $\wp$ in the role of V), and linked by $E_\neg$-edges from $(\beta, \neg\varphi, \wp)$ to $(\beta, \varphi, \bar{\wp})$. $E_\neg$-edges prescribe forced moves (for player $\wp$ say, but it does not matter) from configurations in which the leading connective of φ is a negation. Then the winning positions of player 1 are those $(\beta, \varphi, \wp)$ in which either $\wp = 1$ and $\mathfrak{A}, \beta \models \varphi$ or $\wp = 0$ and $\mathfrak{A}, \beta \not\models \varphi$.

It is also straightforward to adapt the model checking game to deal with FO_∞ rather than FO. $E_\vee$ and $E_\wedge$ can have infinite out-degree reflecting the syntax of infinitary disjunctions and conjunctions; everything else remains just the same; in particular plays are still finite, albeit not necessarily with a uniform finite bound.

2.2 The comparison game: back and forth

The familiar Ehrenfeucht–Fraïssé style model comparison games are two player games played over two structures. A game configuration in these games may be seen as a pairing between two observable configurations, one from each structure. The game is such that the winning positions determine whether or not (or to which degree) these two observable configurations are logically indistinguishable. We present the basic idea in the slightly non-standard terminology of (pairings between) observable configurations in order to highlight the connection between the comparison games and the model checking games. This point of view will contribute to a rather uniform presentation of fragments via restrictions imposed at the level of observable configurations.

The first-order Ehrenfeucht–Fraïssé game Consider two τ-structures $\mathfrak{A}$ and $\mathfrak{A}'$ over the same finite relational vocabulary τ. For partial assignments β, β' to the same (finite) subset of variables $(x_i)_{i \geqslant 1}$ in $\mathfrak{A}$ and $\mathfrak{A}'$, respectively, we write

$$\mathfrak{A}, \beta \equiv_q \mathfrak{A}', \beta'$$

for FO-equivalence up to quantifier-rank q, i.e., $\mathfrak{A}, \beta \models \varphi \;\Leftrightarrow\; \mathfrak{A}', \beta' \models \varphi$ for all $\varphi \in \mathrm{FO}[\tau]$ such that $\mathrm{free}(\varphi) \subseteq \mathrm{dom}(\beta) = \mathrm{dom}(\beta')$ and $\mathrm{qr}(\varphi) \leqslant q$. If $\mathfrak{A}$ and $\mathfrak{A}'$ are clear from the context, we also write just

$$\beta \equiv_q \beta'.$$

The coarsest of these equivalences, $\mathfrak{A}, \beta \equiv_0 \mathfrak{A}', \beta'$ corresponds to a local isomorphism: $\pi\colon \beta(i) \mapsto \beta'(i)$ for $i \in \mathrm{dom}(\beta) = \mathrm{dom}(\beta')$ being an isomorphism between the induced substructures $\mathfrak{A}{\restriction}\mathrm{image}(\beta)$ and $\mathfrak{A}'{\restriction}\mathrm{image}(\beta')$, which is the same as equality of quantifier-free types.

Elementary equivalence, $\mathfrak{A}, \beta \equiv \mathfrak{A}', \beta'$, without the restriction on quantifier-rank, is similarly defined. Note that $\equiv$ is the limit (coarsest common refinement) of the approximations $(\equiv_q)_{q \in \omega}$.

Further, $\mathfrak{A}, \beta \equiv_\infty \mathfrak{A}', \beta'$ stands for equivalence w.r.t. infinitary logic FO_∞.[2]

The first-order Ehrenfeucht–Fraïssé game over $\mathfrak{A}$ and $\mathfrak{A}'$ is played by two players, whom we call player **I** and player **II**. We describe the game protocol in terms of rounds, each round consisting of an exchange of moves: challenge by **I**/response by **II**.

The game board: positions. Positions between rounds are pairs (β, β') of assignments to the same finite subset of variables $(x_i)_{i \geqslant 1}$. Only locally isomorphic assignments will be admissible for player **II**; we speak of *sound positions*:

Sound positions. Position (β, β') is sound if $\mathfrak{A}, \beta \equiv_0 \mathfrak{A}', \beta'$, i.e., if the correspondence $\beta(i) \mapsto \beta'(i)$ describes a local isomorphism. In terms of $\mathcal{O}(\mathfrak{A})$ and $\mathcal{O}(\mathfrak{A}')$: $\beta \in P_\theta \Leftrightarrow \beta' \in P_\theta$ for all atomic θ.

Single round and overall protocol. A single round consists of a challenge/response exchange of moves as follows. In position (β, β'),

– **I** chooses $i \geqslant 1$ and makes a move $\begin{cases} \text{either along an } E_i\text{-edge in } \mathcal{O}(\mathfrak{A}) \text{ from } \beta, \\ \quad \text{or along an } E_i\text{-edge in } \mathcal{O}(\mathfrak{A}') \text{ from } \beta'. \end{cases}$

– **II** must make a move along an E_i-edge in the opposite structure.

This exchange of moves results in an overall transition from position (β, β') to some successor position (γ, γ'), where $\gamma = \beta\frac{a}{i}$ for some $a \in A$ and $\gamma' = \beta'\frac{a'}{i}$ for some $a' \in A'$.

We distinguish different levels of the game according to how many rounds are played.

The q-round game $\mathbb{G}_q(\mathfrak{A}; \mathfrak{A}')$ *(for fixed* $q \in \omega$*)*: play continues from an initial position through q rounds (or until a position is reached that is not sound).

The finite-round game $\mathbb{G}_\omega(\mathfrak{A}; \mathfrak{A}')$: in the initial position, player **I** first selects some $q \in \omega$, then play continues in $\mathbb{G}_q(\mathfrak{A}; \mathfrak{A}')$ from the initial position.

The infinite game $\mathbb{G}_\infty(\mathfrak{A}; \mathfrak{A}')$: play continues through an infinite number of rounds (or until a position is reached that is not sound).

In each variant, **II** loses as soon as the position is not sound. Maintaining soundness of the evolving position is in fact the only commitment for **II**: **II** wins the q-round game $\mathbb{G}_q$ after completion of round q if this final position is sound; similarly **II** wins the finite-round game $\mathbb{G}_\omega$ if she wins $\mathbb{G}_q$ for the q initially selected by **I**; and she wins the infinite game $\mathbb{G}_\infty$ if play continues indefinitely without violation of soundness.[3]

In all of these games we typically also specify the initial position as in $\mathbb{G}_q(\mathfrak{A}, \beta; \mathfrak{A}', \beta')$. For instance, we say that **II** has a winning strategy in $\mathbb{G}_q(\mathfrak{A}, \beta; \mathfrak{A}', \beta')$ if (β, β') is a winning position for player **II** in $\mathbb{G}_q(\mathfrak{A}; \mathfrak{A}')$ (or in $\mathbb{G}_q(\mathfrak{A}, \beta; \mathfrak{A}', \beta')$).

It is obvious that plays of $\mathbb{G}_q$ and $\mathbb{G}_\omega$ are finite and end in a position in which one of the players has won; hence $\mathbb{G}_q$ and $\mathbb{G}_\omega$ are positionally determined. But also $\mathbb{G}_\infty$ is

[2] Equivalence up to quantifier-rank α in FO_∞ can be defined, for every ordinal α. For finite relational vocabularies, $\equiv$ coincides with $\equiv_\omega$, equivalence up to quantifier-rank ω in FO_∞. Note, however, that finitary and infinitary first-order equivalences do not coincide even at quantifier-rank 1 for infinite relational vocabularies.

[3] Clearly a variant formulation to essentially the same effect would restrict the game board to sound positions right away, making **II** lose when she is stuck for a response. This formulation, however, has the slight disadvantage of restricting us to sound initial positions, too.

rather easily shown to be positionally determined, without recourse to deeper results from game theory, as part of the model theoretic analysis underpinning the following theorem. The core of this well-known analysis can be summarised as follows.

Theorem 2.2 (Ehrenfeucht–Fraïssé and Karp). *For all structures of the same* finite *relational vocabulary, $\mathfrak{A}$ and $\mathfrak{A}'$, winning positions in games characterise levels of first-order equivalence in the sense of the following equivalences.*

(a) (β, β') is a winning position for **II** *in $\mathbb{G}_q(\mathfrak{A}; \mathfrak{A}')$ if, and only if, $\mathfrak{A}, \beta \equiv_q \mathfrak{A}', \beta'$.*

(b) (β, β') is a winning position for **II** *in $\mathbb{G}_\omega(\mathfrak{A}; \mathfrak{A}')$ if, and only if, $\mathfrak{A}, \beta \equiv \mathfrak{A}', \beta'$.*

(c) (β, β') is a winning position for **II** *in $\mathbb{G}_\infty(\mathfrak{A}; \mathfrak{A}')$ if, and only if, $\mathfrak{A}, \beta \equiv_\infty \mathfrak{A}', \beta'$.*

We sketch the game-oriented skeleton of the underlying arguments in their most rudimentary form to highlight this aspect (and deliberately ignoring some of the logical niceties, like characteristic formulae, which the more thorough analysis presented in textbooks typically yields).

(i) For the direction from left to right, one shows that logical *in*equivalence yields a winning strategy for player **I**. This follows from the observation that **I** can choose his challenge in a single round from a sound position such that, no matter what response **II** chooses, the resulting position is logically inequivalent at a lower quantifier-rank.

Why is that? A glance at the model checking game helps to illustrate the point. For instance, if $\beta \not\equiv_{m+1} \beta'$ (but $\beta \equiv_0 \beta'$), then this inequivalence manifests itself in some formula $\exists x_i \psi$ with ψ of quantifier-rank at most m. Suppose w.l.o.g. that $\mathfrak{A}, \beta \models \exists x_i \psi$ while $\mathfrak{A}', \beta' \models \forall x_i \neg\psi$. Then a good move for the verifier in position $(\beta, \exists x_i \psi)$ in the model checking game over $\mathfrak{A}$ obviously makes a good move for **I** in this game.[4]

(ii) In the opposite direction, player **II** always has a strategy, for her response to **I**'s challenge in a single round, to maintain the required level of logical equivalence. For instance towards (a) or (b), for a challenge $\gamma = \beta\frac{a}{i}$ in a position (β, β') such that $\beta \equiv_{m+1} \beta'$, **II** can find $a' \in A$ such that $\beta\frac{a}{i} \equiv_m \beta'\frac{a'}{i}$. Otherwise, there would have to be a distinguishing formula $\psi_{a'}$ of quantifier-rank m for every choice of $a' \in A'$, such that $\mathfrak{A}, \beta\frac{a}{i} \models \psi_{a'}$ while $\mathfrak{A}', \beta'\frac{a'}{i} \not\models \psi_{a'}$. But then the formula $\exists x_i \bigwedge_{a'} \psi_{a'}$ would distinguish β and β' at quantifier-rank $m + 1$.

If the underlying structures (and hence the branching degree of the transition systems of observable configurations) are infinite, this argument crucially uses the fact that, for a fixed tuple of free variables there are only finitely many formulae of quantifier-rank m over a fixed finite relational vocabulary, up to logical equivalence – this is what brings $\exists x_i \bigwedge_{a'} \psi_{a'}$ into first-order, even if A' is infinite.[5] We note that the corresponding claims in (a) and (b) of the theorem actually fail for infinite relational vocabularies, even over finite structures. For (c) on the other hand, to which the above argument is readily adapted, finiteness (of the conjunction or of the vocabulary) is not essential.

The equally familiar description in terms of *back-and-forth systems* corresponds to a delineation of a winnning region for **II** with the appropriate closure conditions (the *back-and-forth conditions*) that guarantee that player **II** has responses to keep the game

[4]Entirely analogous reasoning applies towards (c) and for inequivalence in FO_∞, w.r.t. its ordinal-valued quantifier-rank.

[5]While this is easily proved by induction on quantifier-rank, these preparatory considerations are clearly not even required for the argument if we deal just with finite models.

within the prescribed region, against all challenges by **I**. The essential difference between the finite and the infinite game is that, in the finite games, winning regions are stratified according to how many rounds are still to be survived. The winning region for the infinite game, on the other hand, is static, corresponding to an invariant that needs to be maintained indefinitely (this is the classical notion of back-and-forth equivalence or *partial isomorphism* in model theory, see for instance [29]).

Example 1: finite linear orderings The first example illustrating the usefulness of the first-order Ehrenfeucht–Fraïssé game in almost any textbook presentation concerns the limitations of FO in expressing properties of finite linear orderings (or discrete linear orderings more generally). We just state the following well-known result in order to stress its technical affinity with simple locality based arguments to be considered later.

Lemma 2.3. *Consider two finite linear orderings* $\mathfrak{A} = (\mathbb{N}, <)\restriction[0, m]$ *and* $\mathfrak{A}' = (\mathbb{N}, <)\restriction[0, m']$ *with assignments to tuples*

$$\begin{array}{l} \beta = \mathbf{n} \;\; = (n_0, \dots, n_k) \textit{ where } 0 = n_0 < n_1 < \cdots < n_{k-1} < n_k = m \textit{ and} \\ \beta' = \mathbf{n}' = (n'_0, \dots, n'_k) \textit{ where } 0 = n'_0 < n'_1 < \cdots < n'_{k-1} < n'_k = m'. \end{array}$$

We write $d_i := n_{i+1} - n_i$ *and* $d'_i := n'_{i+1} - n'_i$ *for distances between consecutive points in these assignments. Then the following are equivalent for any* $q \geqslant 1$*:*

(i) $\mathfrak{A}, \beta \equiv_q \mathfrak{A}', \beta'$

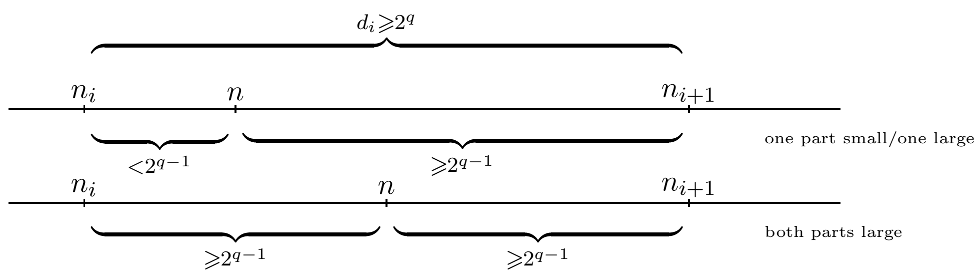
(ii) for $0 \leqslant i < k$*:* $d_i = d'_i$ *or both* $d_i, d'_i \geqslant 2^q$*.*

For the naked finite linear orderings one obtains that

$$\mathfrak{A} \equiv_q \mathfrak{A}' \quad \Longleftrightarrow \quad |A| = |A'| \text{ or } |A|, |A'| \geqslant 2^q - 1.$$

For (ii) ⇒ (i) in the lemma, consider the first round in a game played from a position satisfying the distance constraints (ii) with critical distance 2^q. It suffices to exhibit a strategy for player **II** to respond to any challenge by player **I** in such a manner that the resulting position satisfies the analogous distance constraints (ii), but now with critical distance 2^{q-1} instead of 2^q. W.l.o.g. we may assume that **I** extends the configuration β by some new element $n \in I_i = (n_i, n_{i+1})$. The case that **I** plays in $\mathfrak{A}'$ instead is symmetric. In case $d_i = d'_i$ (the pair of intervals concerned have exactly the same length), **II** may select an element $n' \in I'_i = (n'_i, n'_{i+1})$ at precisely the same distances from the end points in I'_i as n has in I_i; the resulting position even satisfies the distance constraints with critical distance 2^q again.

In the more interesting case, we have $d_i \neq d'_i$ but $d_i, d'_i \geqslant 2^q$. We consider cases, as to the sub-division of the interval $I_i = (n_i, n_{i+1})$ by n:



As the distances of n from the end points of I_i add up to d_i, at most one of these distances can be less than 2^{q-1}; if one distance is 'small' in this sense, **II** may copy this distance exactly to find a matching $n' \in I'_i$ (the other distance will automatically be 'large', i.e., $\geqslant 2^{q-1}$ just as on the side of I_i); if both distances are at least 2^{q-1}, then **II** similarly finds $n' \in I'_i$ which is at least that far from both end points of I'_i.

It is a nice exercise to formalise sentences in quantifier rank q that, over finite linear orderings, require at least $2^q - 1$ elements, thus showing that the given bounds are tight.

It is also useful to draw on the compositionality of strategies for **II** w.r.t. concatenation of linearly ordered intervals (slightly more generally, strategies for player **II** are compatible with ordered sums of linearly ordered structures in an otherwise monadic vocabulary; or with concatenation of word structures). The implicit decomposition of the game into subgames on intervals in the above strategy considerations reflects this.

These game arguments illustrate the well-known fact that, for instance, no FO sentence can distinguish even length from odd length finite linear orderings. Any sentence φ proposed for the purpose is defeated by the example of linear orderings of lengths 2^q and $2^q - 1$ for $q := \mathrm{qr}(\varphi)$.

Remark Maybe somewhat unexpectedly (and disturbing only from a didactic point of view), this particular finite model theory assertion can also be shown by classical means. Suppose there were a sentence $\varphi \in \mathrm{FO}[<]$ such that a *finite* linear ordering satisfies φ if, and only if, it is of even length. Let $[\varphi]^{\leqslant x}$ be the relativisation of φ to the initial segment formed by x. Let $\psi_0 \in \mathrm{FO}[<]$ be the usual characterisation of discrete linear orderings with first and without last element; $\psi_1 \in \mathrm{FO}[<]$ the assertion that precisely every other element x satisfies $[\varphi]^{\leqslant x}$. Then $\psi_0 \wedge \psi_1$ would characterise the order type of $(\omega, <)$, which is impossible by compactness. Consider any non-standard model $(A, <)$ of ψ_0 as in the sketch. Since the non-standard part of $(A, <)$ consists of an ordered sum of parts ordered like $(\mathbb{Z}, <)$, the successor operation induces an automorphism of the non-standard part. Therefore $[\varphi]^{\leqslant x}$ cannot distinguish next neighbours within the non-standard part, and $(A, <) \not\models \psi_0 \wedge \psi_1$.

Example 2: a simple locality argument (also compare section 2.5) Let τ be a finite relational vocabulary. A formula $\varphi(\mathbf{x}) \in \mathrm{FO}[\tau]$ is called *ℓ-local* if, in any τ-structure $\mathfrak{A}$, whether $\mathfrak{A} \models \varphi[\mathbf{a}]$ is fully determined by $\mathfrak{A}{\restriction}N^\ell[\mathbf{a}]$ (the ℓ-neighbourhood of $\mathbf{a}$):

$$\mathfrak{A} \models \varphi[\mathbf{a}] \quad \Leftrightarrow \quad \mathfrak{A}{\restriction}N^\ell[\mathbf{a}] \models \varphi[\mathbf{a}].$$

Similarly $\varphi(\mathbf{x})$ is *invariant under disjoint unions* if for all $\mathfrak{A}, \mathbf{a}$ and $\mathfrak{B}$,

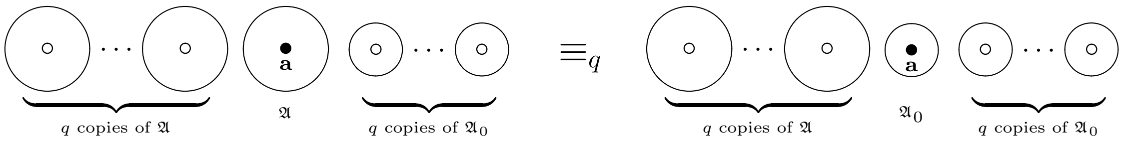
$$\mathfrak{A} \models \varphi[\mathbf{a}] \quad \Leftrightarrow \quad \mathfrak{A} \oplus \mathfrak{B} \models \varphi[\mathbf{a}],$$

where $\mathfrak{A} \oplus \mathfrak{B}$ is the disjoint union of $\mathfrak{A}$ and $\mathfrak{B}$.

Lemma 2.4. *If $\varphi \in \mathrm{FO}[\tau]$ is invariant under disjoint unions, then φ is ℓ-local for $\ell = 2^{\mathrm{qr}(\varphi)} - 1$.*

Remark: the bound on ℓ is optimal, since there is, for every q, a quantifier-rank q formula $\varphi_q(x) \in \mathrm{FO}[E, P]$ asserting that $N^{2^q-1}(x) \cap P \neq \emptyset$.[6]

Proof. Let φ be invariant under disjoint unions, $q := \mathrm{qr}(\varphi)$ and $\ell := 2^q - 1$. For $\mathbf{a} \in \mathfrak{A}$ and $\mathfrak{A}_0 := \mathfrak{A} \restriction N^\ell(\mathbf{a})$ it suffices to show that $\mathfrak{A} \models \varphi[\mathbf{a}]$ iff $\mathfrak{A}_0 \models \varphi[\mathbf{a}]$. By invariance under disjoint unions, moreover, it suffices to establish an equivalence of the form $\mathfrak{A}, \mathbf{a} \oplus \mathfrak{C} \equiv_q \mathfrak{A}_0, \mathbf{a} \oplus \mathfrak{C}$ for a suitable structure $\mathfrak{C}$. Taking $\mathfrak{C}$ to be the disjoint union of q further disjoint isomorphic copies each of $\mathfrak{A}$ and of $\mathfrak{A}_0$, we argue this equivalence:



In the game on these structures, **II** wins the q-round game as follows. We use $d_m := 2^{q-m}$ as a critical distance to be observed in round m. **II** is to play such that the configurations resulting from round m are linked by a component-wise trivial isomorphism between their $(d_m - 1)$-neighbourhoods. This condition is satisfied at the start, for $m = 0$; for $m = q$ it still guarantees a local isomorphism between the final configurations, hence a win for **II**.

Here is how to maintain the condition through round m of the game for $m \geqslant 1$:

(i) if **I**'s challenge goes to some element at distance greater than d_m from the current configuration, then **II** responds with the same element in a new isomorphic component on the opposite side (new in the sense of not yet involved in the current configuration; such are always left).

(ii) if **I**'s challenge goes to an element within distance d_m of the current configuration, then **II** finds a response via the trivial local isomorphism between the $(d_{m-1} - 1)$-neighbourhoods of the current configurations. We note that $d(x, y) \leqslant d_m$ implies $N^{d_m - 1}(y) \subseteq N^{d_{m-1}-1}(x)$, as $d(x, z) \leqslant d(x, y) + d(y, z) \leqslant d_m + d_m - 1 = d_{m-1} - 1$. $\square$

2.3 Natural restrictions/variations

Several of the most natural fragments of FO can be presented in terms of restrictions or modifications of the system $\mathcal{O}(\mathfrak{A})$ of observable configurations associated with structure $\mathfrak{A}$. The k-variable fragment FO^k of FO, for instance, exactly corresponds to the restriction that only up to k elements of $\mathfrak{A}$ are "simultaneously observable" – we just need to restrict the assignments to size k. While this is a uniform, purely quantitative restriction, the modal and guarded fragments of first-order logic are based on structural, qualitative restrictions. In the guarded fragment GF, access to observable configurations is restricted by the requirement that the target configuration be guarded, i.e., covered by some relational ground atom (which is explicitly reflected in the syntax of guarded quantification). In the more basic modal fragment ML of FO, $\mathfrak{A}$ itself *is* the system of observable configurations – in this case the transition relations between the (trivial, one-point) configurations are the key to the restrictions imposed in modal quantification.

[6]One obtains $\varphi_q(x)$ inductively, based on $\varphi_{q+1}(x) := \exists y(d(x, y) \leqslant 2^q \wedge \varphi_q(y))$.

Below we treat k-variable logic, guarded logic and modal logic in this order of increasing specialisation, with particular emphasis on modal logic and its comparison game, the bisimulation game.

2.3.1 The k-variable fragment and the k-pebble game

The k-variable fragment $\mathrm{FO}^k \subseteq \mathrm{FO}$ in a relational vocabulary τ consists of all first-order formulae in which only the variable symbols $x_1, \ldots, x_k$ occur, bound or free. Assignments over τ-structures $\mathfrak{A}$ can thus be restricted and normalised to be full assignments to these k variables. We therefore identify assignments with k-tuples.

Correspondingly, let $\mathcal{O}^k(\mathfrak{A})$ be the restriction of $\mathcal{O}(\mathfrak{A})$ to $\{\beta \colon |\beta| = k\} = A^k$.

It is easy to see that the restriction of both the model checking game and of the comparison game that ensues if $\mathcal{O}(\mathfrak{A})$ is consistently replaced with $\mathcal{O}^k(\mathfrak{A})$ are adequate for the semantics of FO^k and for the induced notions of k-variable equivalence. The k-variable Ehrenfeucht–Fraïssé game is just the k-pebble game: moves along E_i-edges in $\mathcal{O}^k(\mathfrak{A})$ correspond to the re-positioning of the i-th pebble on $\mathfrak{A}$. The correspondence between the different levels of the game and of k-variable equivalence are the following, for finite relational vocabularies τ:

$\equiv^k_q$: FO^k-equivalence up to quantifier-rank q;
captured by the q-round k-pebble game $\mathbb{G}^k_q$.

$\equiv^k$: FO^k-equivalence;
captured by the finite-round k-pebble game $\mathbb{G}^k_\omega$.

$\equiv^k_\infty$: FO^k_∞-equivalence;
captured by the infinite k-pebble game $\mathbb{G}^k_\infty$.

It is important to note that $\mathcal{O}^k(\mathfrak{A})$ is of finite type for each finite τ, and of polynomial size in the size of $\mathfrak{A}$, for finite $\mathfrak{A}$. For the model checking implications see [23, 21] and also some related remarks in section 5.1. In particular, the *combined model checking complexity* for FO^k is complete for Ptime, while for FO it is complete for Pspace.

2.3.2 The guarded fragment and the guarded bisimulation game

The characteristic feature of the guarded fragment GF of first-order logic [2] is the relativisation of first-order quantification to guarded tuples – similar to the restriction along accessibility edges in modal logic. Also compare the remarks in section 1.2 where GF was introduced as a fragment of FO.

We start out with a discussion of a very liberal setting for the guarded fragment that most naturally reflects the syntactic freedom allowed in the standard formalisation of GF as given in section 1.2. Afterwards we also indicate some more succinct alternative formulations that correspond to certain syntactic normalisations (e.g., regarding the number of variables used) to which GF can be subject without impairing its expressive power; such less liberal formalisations can be of technical advantage in the model theoretic analysis of GF and its relatives.

Recall that a subset $s \subseteq A$ is *guarded* in the τ-structure $\mathfrak{A}$ if it is a singleton set or, if for one of the relations $R \in \tau$, there is some tuple $\mathbf{a} = (a_1, \ldots, a_r) \in R^{\mathfrak{A}}$ for which

$s = [\mathbf{a}]$.[7] In particular, the cardinality of guarded subsets is bounded by the width of the vocabulary τ. A *tuple* $\mathbf{b}$ in $\mathfrak{A}$ is called guarded if $[\mathbf{b}] \subseteq s$ for some guarded subset s. The same terminology applies to assignments β in $\mathfrak{A}$.

Call a tuple $\mathbf{b}$ or an assignment β in $\mathfrak{A}$ *strictly guarded* if $[\beta]$ is itself a guarded subset. More specifically, for an atomic τ-formula α, we say the assignment β is *strictly guarded by* α if $\mathrm{var}(\alpha) = \mathrm{dom}(\beta)$ and $\mathfrak{A}, \beta \models \alpha$, which implies that $[\beta]$ is indeed a guarded subset. (In order to capture also guarded singleton sets, we allow α to be an equality atom.)

A system of observable configurations for GF We work with the following system of observable configurations $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$ over the set of all finite (partial) assignments over $\mathfrak{A}$ with new binary transitions relations $E_{\alpha,\rho}$ (see below) and unary predicates P_θ (as before). The universe of $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$ is the same as in $\mathcal{O}(\mathfrak{A})$ for FO (this is for the liberal, redundant formalisation).

The transition relations of $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$ describe passages from some assignment β to a new assignment β' where the target assignment β' is required to be strictly guarded by some atomic formula α. Each transition relation specifies both the atomic formula α and a set of identities between components of the old and the new assignment. As both β and β' are finite partial functions on the positive integers, a set of identities between components can be specified as a finite set ρ of pairs of positive integers. We write $\beta \overset{\rho}{=} \beta'$ if $\beta(i) = \beta'(j)$ for all $(i,j) \in \rho$. Then for every ρ and α, let $E_{\alpha,\rho}$ be interpreted as the following transition relation on $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$:

$$E_{\alpha,\rho} = \big\{ (\beta, \beta') \colon \beta \overset{\rho}{=} \beta', \beta' \text{ strictly guarded by } \alpha \ \big\}.$$

Unary predicates P_θ for atomic types $\theta(\mathbf{x})$ are as in the basic system $\mathcal{O}(\mathfrak{A})$.

Guarded model checking The game graph for the model checking of formulae in GF is obtained from $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$ and a suitable formalisation of the syntax of guarded quantification in close analogy to the basic case. With the formation rule of existential guarded quantification, for instance,

$$\varphi(\mathbf{x}) = \exists \mathbf{y}(\alpha(\mathbf{x}') \wedge \psi(\mathbf{x}')),$$

where $\mathbf{y}$ is a subtuple of $\mathbf{x}'$, associate an $E_{\alpha,\rho,\exists}$-edge in the syntax tree from $\varphi(\mathbf{x})$ to $\psi(\mathbf{x}')$, where $\rho = \{(i,j) \colon x_i = x'_j\}$. In the game graph $\mathbb{G}^{\mathsf{G}}(\mathfrak{A}, \Phi)$, correspondingly, there are $E_{\alpha,\rho,\exists}$-edges from positions $(\beta, \varphi(\mathbf{x}))$ to positions $(\beta', \psi(\mathbf{x}'))$ such that $\mathbf{x} \subseteq \mathrm{dom}(\beta)$, $\mathbf{x}' \subseteq \mathrm{dom}(\beta')$, $(\beta, \beta') \in E_{\alpha,\rho}$ in $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$. Similarly, universal guarded quantifications $\varphi(\mathbf{x}) = \exists \mathbf{y}(\alpha(\mathbf{x}') \wedge \psi(\mathbf{x}'))$ give rise to edges in $E_{\alpha,\rho,\forall}$ in the syntax tree, and induce transition relations $E^{\mathbb{G}}_{\alpha,\rho,\forall}$ in $\mathbb{G}^{\mathsf{G}}(\mathfrak{A}, \Phi)$.

Note that existential and universal quantification of variables in GF proceeds in batches (so as to cover a guarded successor set fully in one step) rather than element-wise. Correspondingly, first-order quantifier-rank is replaced by the nesting depth of guarded quantification steps for an appropriate analysis of quantifier complexity. This is important for the induced levels of GF equivalence, which are considered in connection with the comparison game of guarded bisimulation below.

[7] Recall that we denote as $[\mathbf{b}]$ the set of components of a tuple $\mathbf{b}$, and similarly write $[\beta]$ for the image set of an assignment β.

Guarded bisimulation In line with the general idea, positions between rounds in the guarded Ehrenfeucht-Fraïssé game $\mathbb{G}^{\mathsf{G}}(\mathfrak{A}, \mathfrak{A}')$ are matching pairs of assignments (β, β') in $\mathfrak{A}$ and $\mathfrak{A}'$. With the possible exception of the initial position of the game, which we choose to ignore in the following, we may restrict attention to positions in which both β and β' are strictly guarded (this is guaranteed for successor positions after the first round, by the rules below).

Soundness means that the induced correspondence $\beta(i) \mapsto \beta'(i)$ for $i \in \mathrm{dom}(\beta) = \mathrm{dom}(\beta')$ is a local isomorphism; insofar as the assignments are strictly guarded in their structures, the correspondence is a bijection between guarded subsets and thus a local isomorphism between induced substructures on guarded subsets $s = [\beta]$ and $s' = [\beta']$. Challenge/response pairs of moves responsible for taking the game through a single round are governed by **I**'s selection of an $E_{\alpha,\rho}$ and an $E_{\alpha,\rho}$ successor γ of β in $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$ or an $E_{\alpha,\rho}$ successor γ' of β' in $\mathcal{O}^{\mathsf{G}}(\mathfrak{A}')$, and thus, together with **II**'s response, to a new local isomorphism between substructures induced on a new pair of guarded subsets $t = [\gamma]$ and $t' = [\gamma']$ (insofar as the successor position is sound again, i.e., unless **II** has lost).

A conceptually smoother, equivalent formulation therefore is the following, which we take as the preferred description of the *guarded bisimulation game*. Positions in the game are local bijections $\sigma\colon s \to s'$ between guarded subsets $s \subseteq A$ and $s' \subseteq A'$. In a single round played from position $\sigma\colon s \to s'$, **I** proposes either a guarded subset $t \subseteq A$ or a guarded subset $t' \subseteq A'$; **II** has to respond with a guarded subset in the opposite structure (call this other subset $t' \subseteq A'$ or $t \subseteq A$, as the case may be) and a bijection $\rho\colon t \to t'$ that is compatible with σ. Compatibility of ρ with σ means that ρ needs to agree with σ on $s \cap t$ if **I** chose t; and on $s' \cap t'$ if **I** chose t'. **II** loses if there is no such ρ or if ρ is not a local isomorphism.

Either formulation of the game supports the usual analysis, which, as expected, establishes correspondences between winning positions for **II** in the different levels of the game and equivalence in GF. For finite relational vocabularies τ these are:

- $\equiv^{\mathsf{G}}_q$: GF-equivalence up to guarded nesting depth q;
  captured by the q-round guarded bisimulation game $\mathbb{G}^{\mathsf{G}}_q$.
- $\equiv^{\mathsf{G}}$: GF-equivalence;
  captured by the finite-round guarded bisimulation game $\mathbb{G}^{\mathsf{G}}_\omega$.
- $\equiv^{\mathsf{G}}_\infty$: GF_∞-equivalence;
  captured by the infinite guarded bisimulation game $\mathbb{G}^{\mathsf{G}}_\infty$.

More succinct representations Another, much more succinct view on the observable configurations can be based on the use of more restricted assignments: it essentially suffices to admit strictly guarded assignments with domain $\{1 \ldots, k\}$ where k is the width of τ. This second aspect corresponds to the normalisation of variables to $x_1, \ldots, x_k$ as in FO^k.[8] Here we use strictly guarded assignments to variables $x_1, \ldots, x_k$, or surjective partial maps from $\{1, \ldots, k\}$ onto guarded subsets of $\mathfrak{A}$.

[8]Even more restrictively, [22] for technical convenience uses a format with only injective assignments, there called *guarded lists*.

The type of the resulting system of guarded observable configurations is finite for finite τ. The model checking game obtained in analogy with the above, by making the obvious changes and restrictions regarding the syntax of formulae, then really is for (a specific syntactic variant of) $\mathsf{GF}^k := \mathsf{GF} \cap \mathrm{FO}^k$.

A closer analysis of the Ehrenfeucht–Fraïssé games and notions of guarded equivalence resulting from the two different formalisations would show that there is no loss of expressiveness as far as properties of (strictly) guarded tuples are concerned. The only real restriction concerns expressiveness at the quantifier-free level and in boolean combinations, and this is inessential for many purposes. The difference arises, trivially, because GF does not impose any restrictions on boolean combinations. Analysis of the game shows, however, that any formula of GF (in the liberal format) is logically equivalent to a boolean combination of quantifier-free formulae and strictly guarded formulae (each of which can, up to a necesssary renaming of variables, be formalised in the above fragment GF^k).

Corollary 2.5. *Any formula in $\mathsf{GF}[\tau]$ with explicitly guarded free variables is equivalent to a formula in $\mathsf{GF} \cap \mathrm{FO}^k$ where k is the width of τ.*

2.3.3 The modal fragment and the bisimulation game

Modal logic is naturally interpreted over transition systems (Kripke structures in traditional terminology). Having chosen a modal perspective for our analysis of fragments, we may now choose the transition system $\mathfrak{A}$ itself – as a relational structure in a given vocabulary τ with binary relations E_α and unary predicates P_j – as the system of modally observable configurations, putting $\mathcal{O}^{\mathsf{M}}(\mathfrak{A}) = \mathfrak{A}$. To keep in line with the general framework we may want to replace the individual P_j in $\mathfrak{A}$ by P_θ that are complete propositional types in the p_j/P_j (in first-order terms: atomic P_j-types in single variables x, containing for each P_j either the atomic formula $P_j x$ or its negation $\neg P_j x$).

Modal model checking The modal model checking game over structure $\mathfrak{A}$ is played in a game graph based on $\mathfrak{A}$ and the syntax tree of the modal formulae under consideration. With the formation rule of existential modal quantification

$$\varphi = \Diamond_\alpha \psi$$

we associate an $E_{\Diamond_\alpha}$ edge in the syntax tree from φ to ψ. In the game graph $\mathbb{G}^{\mathsf{M}}(\mathfrak{A}, \Phi)$, this induces $E_{\alpha,\rho,\exists}$ edges from positions (a, φ) to positions (b, ψ) for $(a, b) \in E_\alpha^{\mathfrak{A}}$. Analogously for $\Box_\alpha$ quantification: edges in $E_{\Box_\alpha}$ from $\varphi = \Box_\alpha \psi$ to ψ in the syntax tree give rise to transitions in $\mathbb{G}^{\mathsf{M}}(\mathfrak{A}, \Phi)$ from (a, φ) to (b, ψ) for every $(a, b) \in E_\alpha^{\mathfrak{A}}$.

It is clear that the model checking game for FO^2 emulates the modal model checking game, via the standard translation of ML into FO^2:

$$\begin{array}{rcl} (\Diamond_\alpha \psi)_x & = & \exists y (E_\alpha xy \wedge \psi_y), \\ (\Box_\alpha \psi)_x & = & \forall y (E_\alpha xy \rightarrow \psi_y), \end{array}$$

where $\{x, y\} = \{x_1, x_2\}$. In terms of this translation, a move along an E_α edge (a, b) in the $\mathfrak{A}$ component of $\mathbb{G}^{\mathsf{M}}(\mathfrak{A}, \Phi)$ is simulated by an E_2 move from any position of the form $(a,*)$ to (a,b) or by an E_1 move from any $(*,a)$ to (b,a) in the $\mathcal{O}^2(\mathfrak{A})$ component of $\mathbb{G}^2(\mathfrak{A},\mathrm{FO}(\Phi))$. At the same time this emulation can be interpreted in $\mathbb{G}^{\mathsf{G}}(\mathfrak{A},\mathrm{FO}(\Phi))$, since $\{a,b\}$ is a strictly guarded assignment and (a,a) is linked to (a,b), for instance, by an $E_{\alpha,\rho}$ edge in $\mathcal{O}^2(\mathfrak{A})$ for $\rho = \{(1,1)\}$.

Bisimulation The bisimulation game is the Ehrenfeucht–Fraïssé game for modal logic. It also has a special status because of its fundamental nature as the quintessential back-and-forth game – game equivalence of game graphs – to be discussed in the following section.

In line with the general approach, the positions (between rounds) in $\mathbb{G}^{\mathsf{M}}(\mathfrak{A},\mathfrak{A}')$ are pairs of observable configurations in $\mathcal{O}(\mathfrak{A}) = \mathfrak{A}$ and $\mathcal{O}(\mathfrak{A}') = \mathfrak{A}'$, i.e., pairs $(a,a') \in A \times A'$. The challenge/response exchange that constitutes a single round is as follows:

– **I** selects a transition relation E_α, and $\begin{cases} \text{either some } E_\alpha \text{ successor } b \text{ of } a \text{ in } \mathfrak{A}, \\ \text{or some } E_\alpha \text{ successor } b' \text{ of } a' \text{ in } \mathfrak{A}'. \end{cases}$

– **II** has to respond by selecting an E_α successor in the opposite structure.

Overall this results in a successor position (b,b') for which $(a,b) \in E_\alpha^{\mathfrak{A}}$ and $(a',b') \in E_\alpha^{\mathfrak{A}'}$. A position (a,a') is sound if a and a' satisfy exactly the same predicates P_j (atomic propositions p_j in modal terminology), which clearly corresponds to quantifier-free indistinguishability in $\mathsf{ML}[\tau]$.

Because of their immediate importance we introduce the usual dedicated notation for the levels of equivalence that are defined in terms of winning positions for player **II** in the different levels of this bisimulation game. As above, the q-round, finite-round, and infinite bisimulation game on $\mathfrak{A}$ and $\mathfrak{A}'$ are denoted $\mathbb{G}^{\mathsf{M}}_q(\mathfrak{A},\mathfrak{A}')$, $\mathbb{G}^{\mathsf{M}}_\omega(\mathfrak{A},\mathfrak{A}')$, and $\mathbb{G}^{\mathsf{M}}_\infty(\mathfrak{A},\mathfrak{A}')$. We then define

$$\begin{array}{rcl} \mathfrak{A},a \sim_q \mathfrak{A}',a' & \text{iff} & (a,a') \text{ is a winning position for } \mathbf{II} \text{ in } \mathbb{G}^{\mathsf{M}}_q(\mathfrak{A},\mathfrak{A}'); \\ \mathfrak{A},a \sim_\omega \mathfrak{A}',a' & \text{iff} & (a,a') \text{ is a winning position for } \mathbf{II} \text{ in } \mathbb{G}^{\mathsf{M}}_\omega(\mathfrak{A},\mathfrak{A}'); \\ \mathfrak{A},a \sim \mathfrak{A}',a' & \text{iff} & (a,a') \text{ is a winning position for } \mathbf{II} \text{ in } \mathbb{G}^{\mathsf{M}}_\infty(\mathfrak{A},\mathfrak{A}'). \end{array}$$

Note that $\sim$ is the classical notion of bisimulation equivalence – equivalence w.r.t. the infinite bisimulation game, and as such the modal counterpart of partial isomorphism.

We denote the relevant levels of equivalence in modal logic as $\equiv^{\mathsf{M}}_q$ (up to modal nesting depth q), $\equiv^{\mathsf{M}}$ (full equivalence in finitary ML), and $\equiv^{\mathsf{M}}_\infty$ (equivalence in the infinitary extension ML_∞). The associated Ehrenfeucht–Fraïssé and Karp theorems then state, for finite modal vocabularies τ, the following equivalences:

$$\begin{array}{rcl} \mathfrak{A},a \sim_q \mathfrak{A}',a' & \Leftrightarrow & \mathfrak{A},a \equiv^{\mathsf{M}}_q \mathfrak{A}',a'. \\ \mathfrak{A},a \sim_\omega \mathfrak{A}',a' & \Leftrightarrow & \mathfrak{A},a \equiv^{\mathsf{M}}_\omega \mathfrak{A}',a'. \\ \mathfrak{A},a \sim \mathfrak{A}',a' & \Leftrightarrow & \mathfrak{A},a \equiv^{\mathsf{M}}_\infty \mathfrak{A}',a'. \end{array}$$

Modal variations The simple extensions of basic modal logic by inverse modalities and/or global modality, ML^-, $\mathsf{ML}^\forall$ and $\mathsf{ML}^{-\forall}$, are matched by corresponding variations in $\mathcal{O}(\mathfrak{A})$ and $\mathbb{G}(\mathfrak{A},\mathfrak{A}')$. To deal with inverse modalities, $\mathcal{O}(\mathfrak{A})$ is enriched with the converse relations to the E_α, $(E_\alpha^-)^{\mathfrak{A}} = \{(b,a) \colon (a,b) \in E_\alpha^{\mathfrak{A}}\}$; to deal with the global modality, $\mathcal{O}(\mathfrak{A})$ is expanded by the full binary relation $U^{\mathfrak{A}} = A \times A$. Everything else, including associated Ehrenfeucht–Fraïssé and Karp theorems, is then set up by straightforward analogy and we leave the details as an exercise. For later use, we denote the levels of *two-way global bisimulation equivalence* corresponding to the combined extension by inverse modalities and the global modality by $\approx_q$, $\approx_\omega$ and $\approx$.

Bisimulations as relations and back-and-forth systems We also want to use the notational variants corresponding to back-and-forth systems for bisimulation games. Infinitary bisimulation equivalence (the modal counterpart of partial isomorphism) between the nodes of two structures $\mathfrak{A}$ and $\mathfrak{A}'$, in particular, is captured by the relation $Z \subseteq A \times A'$ comprising exactly the winning positions for **II** in $\mathbb{G}^{\mathsf{M}}_\infty(\mathfrak{A}, \mathfrak{A}')$ (known as the *largest bisimulation relation* between $\mathfrak{A}$ and $\mathfrak{A}'$, cf. [8, 16]). Any other relation $Z \subseteq A \times A'$ that delineates an appropriately closed winning region for **II** is also a bismulation relation, and necessarily a subset of the largest such. Corresponding finite bisimulation levels are described by stratified back-and-forth systems in the usual manner. Again, natural and straightforward adaptations for, e.g., two-way global bisimulations are obtained. The difference lies in the closure conditions (back-and-forth conditions), which reflect the nature of the challenges that **I** is allowed, since **II** must have resposes to all of them within the prescribed collection of positions.

A particular variant of bisimulation relationships is realised by homomorphisms whose graphs are bisimulation relations (*bounded morphisms* in cassical modal terminology, cf. [8, 16]). For instance, in the case of the two-way global bisimulation relation $\approx$, we write

$$\pi \colon \mathfrak{A}, a \xrightarrow{\approx} \mathfrak{A}', a'$$

to indicate that $\pi \colon A \to A'$ is a map sending a to a' and such that its graph is a bisimulation relation with the back-and-forth closure conditions approriate for global two-way bisimulation game (in particular π needs to be a surjective homomorphism).

Saturation and Hennessy–Milner properties We shall later look at the relationship between equivalence w.r.t. the infinite game $\mathbb{G}_\infty$ and the finite approximations to the finite-round game $\mathbb{G}_\omega$ induced by the q-round games $(\mathbb{G}_q)_{q\in\omega}$ also for games other than bisimulation. It is therefore interesting to understand under which conditions there is no gap between the limit of the finite approximations and full infinitary equivalence. In the modal situation, or for the bisimulation game, this situation is particularly transparent, and at the same time holds the key to the general situation for other fragments in the game-oriented analysis.

Definition 2.6. Let $\mathfrak{A}$ be a τ transition system with transition relations E_α.

(i) $\Phi \subseteq \mathsf{ML}[\tau]$ is called a $\Diamond_\alpha$*-type* at $a \in \mathfrak{A}$ if $\mathfrak{A}, a \models \Diamond_\alpha \bigwedge \Phi_0$ for every finite $\Phi_0 \subseteq \Phi$; it is *realised* at $a \in \mathfrak{A}$ if there is some b such that $(a, b) \in E_\alpha^{\mathfrak{A}}$ and $\mathfrak{A}, b \models \Phi$.

(ii) $\mathfrak{A}$ is called *modally saturated* if, for all α and all $a \in \mathfrak{A}$, every $\Diamond_\alpha$-type at a is realised at a.

It is not hard to see that ω-saturated transition systems, and in particular finite transition systems are modally saturated. But a very simple argument also shows that even all finitely branching transition systems are modally saturated. In the case of a structure $\mathfrak{A}$ that is finitely branching (w.r.t. E_α) at a, consider some $\Diamond_\alpha$-type Φ at a. Suppose Φ were not realised at a. This means that, for every E_α successor b of a

there must be some $\varphi_b \in \Phi$ not satisfied at b. But then the finite subset Φ_0 of these φ_b would violate the defining condition for a $\Diamond_\alpha$-type at a: $\mathfrak{A}, a \models \Box_\alpha \bigvee_b \neg\varphi_b$, whence $\mathfrak{A}, a \not\models \Diamond_\alpha \bigwedge \Phi_0$.

For this and also for the reasoning behind the lemma below, compare part (ii) of the argument indicated in connection with Theorem 2.2.

Definition 2.7. A class of transition systems has the *Hennessy–Milner property* if over this class, modal equivalence $\equiv^{\mathsf{M}}$ coincides with full bisimulation $\sim$.

Note that, since even for not necessarily finite vocabularies τ, $\sim_\omega$ implies $\equiv^{\mathsf{M}}$, the Hennessy–Milner property implies that in particular also finite bisimulation equivalence coincides with full bisimulation equivalence. The following lemma also implies that for modally saturated transition systems, modal equivalence, finite and full bisimulation equivalence all fall into one, even for infinite vocabularies.

Lemma 2.8. *The class of modally saturated transition systems has the Hennessy–Milner property.*

The straightforward game argument for this is again suggested by the reasoning underlying Theorem 2.2, part (ii), but finiteness of τ is not required. Playing over modally saturated structures, **II** can maintain modal equivalence between configurations. Consider a position (a, a') in the game $\mathbb{G}^{\mathsf{M}}_\infty(\mathfrak{A}, \mathfrak{A}')$ for which $\mathfrak{A}, a \equiv^{\mathsf{M}} \mathfrak{A}', a'$, and think of a challenge played by **I**, with a move along $(a, b) \in E^{\mathfrak{A}}_\alpha$ say. In general (and even for finite vocabulary) modal equivalence $\mathfrak{A}, a \equiv^{\mathsf{M}} \mathfrak{A}', a'$ (or even $\mathfrak{A}, a \sim_\omega \mathfrak{A}', a'$) would only provide **II** with responses b' that are good for surviving q further rounds, where this could be a separate response for each individual q. Now, however, the full modal theory of b in $\mathfrak{A}$ constitues a $\Diamond_\alpha$-type at a in $\mathfrak{A}$, and modal equivalence $\mathfrak{A}, a \equiv^{\mathsf{M}} \mathfrak{A}', a'$ is good enough to ensure that it therefore also is a $\Diamond_\alpha$-type at a' in $\mathfrak{A}'$. By modal saturation, therefore, this $\Diamond_\alpha$-type is realised at a' in $\mathfrak{A}'$, and any such realisation gives **II** a valid response in the game which maintains $\equiv^{\mathsf{M}}$. But maintaining $\equiv^{\mathsf{M}}$ equivalence throughout the game, **II** cannot lose; so this gives her a strategy in $\mathbb{G}^{\mathsf{M}}_\infty$.

2.4 Bisimulation as the master game

An analysis of whole families of fragments of FO w.r.t. their notions of finite and infinitary equivalence can very nicely be based on the analysis of the bisimulation game over the transition systems of observable configurations associated with the particular fragment.

The possible advantage of this perspective lies in the conceptual separation of the game theoretic commonality, which is here uniformly described in terms of bisimulation, and the particular constraints of the fragment under consideration, which enters the picture through the right formalisation of the observable configurations. The natural criterion for the *right* formalisation lies in the adequacy of the induced model checking game for the semantics of the given fragment.

The treatment of FO and fragments like FO^k, GF and ML (and some of its simple variants) can be put in a uniform format as follows. Let $\mathsf{L} \subseteq \mathrm{FO}$ be a fragment associated with systems of observable configurations $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$ over relational structures $\mathfrak{A}$ in a finite relational vocabulary τ. Together with the overhead that links syntax of L with moves in the model checking game with structure inputs $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$, this model checking game can be taken as a specification of the semantics of L. The bisimulation game between $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$ and $\mathcal{O}^{\mathsf{L}}(\mathfrak{A}')$ then *is* a representation of the Ehrenfeucht-Fraïssé or model comparison game for L. This representation is adequate at a round-by-round level in terms of a syntactic notion of depth in L that corresponds to the number of quantification rounds required in model checking a formula in L. The specification of the model checking game is in turn reflected in the format of $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$. As an example for the latter point, consider GF as presented above: we deliberately chose transitions in $\mathcal{O}^{\mathsf{G}}(\mathfrak{A})$ to link any two strictly guarded patches in one transition rather than a sequence of transitions corresponding to one-new-element-at-a-time moves as in $\mathcal{O}(\mathfrak{A})$. The latter option would have turned FO quantifier-rank into our measure of semantic complexity in GF whereas the chosen stipulation relates to the coarser but more intuitive notion of guarded nesting depth. With the appropriate notion of depth that is implicit in the granularity of the model checking game based on $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$ come the notions of $\equiv^{\mathsf{L}}_q$ as finite approximations to $\equiv^{\mathsf{L}}$, and (for finite vocabulary) an Ehrenfeucht–Fraïssé theorem of the format

$$\begin{aligned} \mathfrak{A},\beta \equiv^{\mathsf{L}}_q \mathfrak{A}',\beta' \quad &\Leftrightarrow \quad \mathcal{O}^{\mathsf{L}}(\mathfrak{A}),\beta \sim_q \mathcal{O}^{\mathsf{L}}(\mathfrak{A}'),\beta', \text{ for } q \in \omega, \text{ and} \\ \mathfrak{A},\beta \equiv^{\mathsf{L}} \mathfrak{A}',\beta' \quad &\Leftrightarrow \quad \mathcal{O}^{\mathsf{L}}(\mathfrak{A}),\beta \sim_\omega \mathcal{O}^{\mathsf{L}}(\mathfrak{A}'),\beta'. \end{aligned}$$

At the same time, a notion of infinitary L-equivalence is induced by the full bisimulation relation, $\mathcal{O}(\mathfrak{A}),\beta \sim \mathcal{O}^{\mathsf{L}}(\mathfrak{A}'),\beta'$, supporting a Karp theorem of the format

$$\mathfrak{A},\beta \equiv^{\mathsf{L}}_\infty \mathfrak{A}',\beta' \quad \Leftrightarrow \quad \mathcal{O}^{\mathsf{L}}(\mathfrak{A}),\beta \sim \mathcal{O}^{\mathsf{L}}(\mathfrak{A}'),\beta',$$

which can now also be seen as a specification of what L_∞ (in terms of its model checking game) needs to be.

Beyond a uniform perspective on the games and equivalences themselves, the modal perspective on fragments of FO can also indicate what the right transfer of other game-related notions to fragments should be. As one example we state the following observation concerning ω-saturation (in the usual first-order context).

Observation 2.9. *$\mathfrak{A}$ is ω-saturated if, and only if, $\mathcal{O}(\mathfrak{A})$ is modally saturated.*

Similar correspondences can then be taken to define the appropriate notion of ω-saturation in the context of fragments $\mathsf{L} \subseteq \mathrm{FO}$ (e.g., for FO^k or GF), in terms of modal saturation of the corresponding $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$. This allows us to extrapolate to a range of in-between fragments from the Hennessy–Milner property of modal logic to other fragments with the appropriate notion of ω-saturation. In particular, the right types to be considered for this notion of saturation are derived from the modal $\Diamond$-types in the $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$.

On the other hand, for many natural fragments including FO^k, GF^k and all the modal fragments, classical first-order ω-saturation implies ω-saturation (and the Hennessy–Milner property) in the sense of L. This is due to the following.

Observation 2.10. *For any fragment $\mathsf{L} \subseteq \mathrm{FO}$ for which the system of observable configurations $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$ is uniformly first-order interpretable in $\mathfrak{A}$ itself, ω-saturation of $\mathfrak{A}$ implies ω-saturation of $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$, which (by the previous observation) implies modal saturation of $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$, and hence the analogue of the Hennessy–Milner property for L over the class of ω-saturated structures.*

Note that this modal view is based on imposing the modal picture and the bisimulation game on richer fragments of first-order logic, uniformly via the appropriate system of observable configurations and games. Alternatively, one may think of a specialisation of the classically well understood situation for first-order and its infinitary counterpart, their links with classical Ehrenfeucht-Fraïssé games and Karp's theorem (cf. Theorem 2.2). In connection with the last observation for instance, ω-saturation (in the classical sense, w.r.t. FO-types) implies ω-saturation in the sense of L for a fragment $\mathsf{L} \subseteq$ FO, since L-types are (partial) FO-types; a Hennessy–Milner property for ω-saturated structures then follows because player **II** has a strategy to maintain L-equivalence in the infinite L-game starting from L-equivalent configurations. But this, and how L-types are to be defined so that they can be transferred between L-equivalent configurations as required for this argument, may be best understood systematically in terms of the game and its observable configurations as discussed above.

2.5 Locality and modularity of the first-order game

Games and the Ehrenfeucht–Fraïssé method are well suited to the exploration of the expressive power of FO not just classically but equally well over restricted classes of structures, and also to understanding the nature of fragments within FO. Such explorations typically depend on the availability of suitable structures over which the game can be usefully analysed. In order to facilitate the analysis, and equally importantly also as an indication of where to look for the right candidate structures, one can often use the modularity of the game w.r.t. Gaifman locality. We saw a glimpse of that aspect in Lemma 2.4 above.

For Gaifman's theorem, we want to establish that position $(\mathbf{a}, \mathbf{a}')$ in $\mathbb{G}_q(\mathfrak{A}; \mathfrak{A}')$ is a winning position for **II**, i.e., that $\mathfrak{A}, \mathbf{a} \equiv_q \mathfrak{A}', \mathbf{a}'$, on the basis of

– suitable global conditions on $\mathfrak{A}$ and $\mathfrak{A}'$ (without reference to $\mathbf{a}$ and $\mathbf{a}'$), and

– purely local conditions on these parameters within their structures of the form

$$\mathfrak{A}{\restriction}N^\ell(\mathbf{a}), \mathbf{a} \equiv_r \mathfrak{A}'{\restriction}N^\ell(\mathbf{a}'), \mathbf{a}'$$

for values of ℓ and r that are recursively determined as functions of q.

Towards an understanding of the nature of the global requirement, and for a gradation of both the local and global equivalences involved, we need the following definition.

Definition 2.11. (i) For any $\varphi(\mathbf{x})$ we write $\varphi^\ell(\mathbf{x})$ for its relativisation to the (FO-definable) ℓ-neighbourhood of its free variables, $\varphi^\ell(\mathbf{x}) := [\varphi]^{N^\ell(\mathbf{x})}$.
If $q = \mathrm{qr}(\varphi)$, we refer to φ^ℓ as a *local formula* of Gaifman rank (ℓ, q).

(ii) A *basic ℓ-local sentence* is a sentence of the form

$$\exists x_1 \dots \exists x_m \bigwedge_{i<j} d(x_i, x_j) > 2\ell \;\wedge\; \bigwedge_i \psi^\ell(x_i),$$

asserting the existing of an ℓ-scattered m-tuple whose components satisfy the ℓ-local formula $\psi^\ell(x)$. If $q = \mathrm{qr}(\psi)$, we regard the above basic local sentence as one of Gaifman rank (ℓ, q, m).

Definition 2.12. The configurations $\mathfrak{A}, \mathbf{a}$ and $\mathfrak{A}', \mathbf{a}'$ are (ℓ, q, m)-Gaifman-equivalent, denoted as $\mathfrak{A}, \mathbf{a} \equiv^{\ell}_{q,m} \mathfrak{A}', \mathbf{a}'$, if:

(i) $\mathfrak{A}{\restriction}N^{\ell}(\mathbf{a}), \mathbf{a} \equiv_q \mathfrak{A}'{\restriction}N^{\ell}(\mathbf{a}'), \mathbf{a}'$, i.e., $\mathbf{a}$ and $\mathbf{a}'$ satisfy the same ℓ-local formulae φ^{ℓ} for $\mathrm{qr}(\varphi) \leqslant q$ (*local condition*).

(ii) $\mathfrak{A}$ and $\mathfrak{A}'$ satisfy the same basic local sentences of ranks (ℓ', q', m') for all $\ell' \leqslant \ell$, $q' \leqslant q$ and $m' \leqslant m$ (*global condition*).

For fixed finite relational vocabulary and fixed arity of the tuples $\mathbf{a}$, each $\equiv^{\ell}_{q,m}$ has finite index, and respects $\equiv$. Clearly also $\equiv^{\ell}_{q,m}$ is monotone w.r.t. the ranks (ℓ, q, m). Gaifman's theorem says that $\equiv^{\ell}_{q,m}$ approximates full first-order equivalence $\equiv$ well, in the sense that $\equiv$ is the common refinement or limit of all levels $\equiv^{\ell}_{q,m}$.

Theorem 2.13 (Gaifman)**.** *Any* FO*-formula is preserved under* $\equiv^{\ell}_{q,m}$ *for suitable* (ℓ, q, m)*. Equivalently: any formula of* FO *is logically equivalent to a boolean combination of local formulae and basic local sentences.*

Gaifman's original proof establishes the second statement by induction on the FO formula under consideration. The link with the modularity of the Ehrenfeucht–Fraïssé game, however, is brought out more clearly in an argument given in [14], which we adapt to give a brief sketch. To prove the first of the statements in the theorem, it inductively suffices to establish the following assertion about good responses for **II**.

Claim 2.14. *If* $\mathfrak{A}$ *and* $\mathfrak{A}'$ *are* (L, Q, m)*-Gaifman-equivalent* [9] *for values of* L *and* Q *that are sufficiently large in relation to* ℓ *and* q*, and if* $\mathbf{a}$ *and* $\mathbf{a}'$ *of arity less than* m *are such that*

$$\mathfrak{A} \restriction N^{L}(\mathbf{a}), \mathbf{a} \;\equiv_Q\; \mathfrak{A}' \restriction N^{L}(\mathbf{a}'), \mathbf{a}' \qquad \textit{local pre-condition}$$

then for any $b \in A$ *there is some* $b' \in A'$ *such that*

$$\mathfrak{A} \restriction N^{\ell}(\mathbf{a}b), \mathbf{a}b \;\equiv_q\; \mathfrak{A}' \restriction N^{\ell}(\mathbf{a}'b'), \mathbf{a}'b', \qquad \textit{local post-condition}$$

and, symmetrically, with the roles of b *and* b' *exchanged.*

The claim is established on the basis of a case distinction w.r.t. the distance of b from $\mathbf{a}$. Suitable conditions on the choices of L and Q are extracted along the way. Choosing $L \geqslant 3\ell + 1$ and $Q \geqslant q + 1$ at least, any $b \in N^{2\ell+1}(\mathbf{a})$ can be dealt with according to the local pre-condition. For b that are further away from $\mathbf{a}$, $\mathfrak{A}{\restriction}N^{\ell}(\mathbf{a}b)$ is the disjoint union of $\mathfrak{A}{\restriction}N^{\ell}(\mathbf{a})$ and $\mathfrak{A}{\restriction}N^{\ell}(b)$. Due to modularity of the game w.r.t. disjoint unions, it suffices to find $b' \in A'$ that is also far from $\mathbf{a}'$ and such that $\mathfrak{A}'{\restriction}N^{\ell}(b'), b' \equiv_q \mathfrak{A}{\restriction}N^{\ell}(b), b$. In this case we rely on the global condition on $\mathfrak{A}$ and $\mathfrak{A}'$ for a further case distinction. We use the global condition for scattered tuples w.r.t. a quantifier-rank q formula $\psi(x)$ that characterises $\mathfrak{A}{\restriction}N^{\ell}(b), b$ up to $\equiv_q$. We need to guarantee that $\mathfrak{A}'$ has a matching b', i.e., we seek some $b' \notin N^{2\ell+1}(\mathbf{a}')$ satisfying ψ^{ℓ}.

Firstly, if $\mathfrak{A}$ and hence also $\mathfrak{A}'$ have $(2\ell+1)$-scattered m-tuples of elements satisfying ψ^{ℓ}, then one of the components of any such tuple in $\mathfrak{A}'$ will serve as b'.

If, on the other hand, there are no such m-tuples, then the maximal size $n < m$ of $(2\ell + 1)$-scattered tuples for ψ^{ℓ} is the same in $\mathfrak{A}$ and $\mathfrak{A}'$. Now a comparison with n_0,

[9] Due to the absence of parameters this involves only the global condition (ii) of Definition 2.12.

the maximal size of $(2\ell+1)$-scattered tuples for ψ^ℓ within $N^{2\ell+1}(\mathbf{a})$ can help to locate b', provided $L \geqslant 3\ell+1$ and provided Q is large enough to force the same n_0 to work in $\mathfrak{A}'{\restriction}N^{2\ell+1}(\mathbf{a}')$ (via the local pre-condition).

If $n_0 < n$, then there must be realisations of ψ^ℓ outside $N^{2\ell+1}(\mathbf{a}')$ and any such is a good choice for b'.

The remaining subcase that $n_0 = n$ (no surplus of realisations of ψ^ℓ beyond $N^{2\ell+1}(\mathbf{a}')$), implies in particular that $d(\mathbf{a}, b) \leqslant 6\ell+3$ and the existence of such an element satisfying ψ^ℓ at distance greater than $2\ell+1$ but at most $6\ell+3$ is covered by the local pre-condition, provided $L \geqslant 7\ell+3$ and Q is large enough to cover this (under the local pre-condition), too.

3 Special classes of transition systems

Up to bisimulation, every transition system is equivalent to a tree via a bisimilar tree unfolding, just as every game graph can be replaced by the associated game tree, typically making the representation structurally simpler though less succinct. Correspondingly, any *bisimulation invariant* logic (logic whose formulae are preserved under bisimulation equivalence) has the tree model property. Because cycles are unfolded into infinite paths, bisimulation equivalent tree models may necessarily be infinite even though the original model was finite. So bisimilar unfoldings into tree models are typically not available within classes of finite models. In the investigation of the model theoretic relationship between bisimulation invariant fragments of FO with FO itself, however, Gaifman locality can be used to replace acyclicity by local acyclicity in key arguments. We briefly review the classical construction of bisimilar unfoldings into tree models and then review a construction of locally acyclic bisimilar companion structures from [41]. These are used to establish variants of the classical model theoretic characterisations of modal fragments of FO in terms of bisimulation preservation (van Benthem's theorem, cf. Corollary 3.5 below) over natural, restricted classes of transition systems in section 3.2.

3.1 Tree unfoldings and locally tree-like systems

3.1.1 Bisimulation invariance and the tree model property

Let $\mathfrak{A}$ be a transition system in a finite vocabulary τ consisting of binary relations E_α and unary predicates P_j. With $a \in \mathfrak{A}$ we associate the following *bisimilar unfolding of* $\mathfrak{A}$ *at* a, $\mathfrak{A}^*_a$. The universe of $\mathfrak{A}^*_a$ is the set of all finite, edge-labelled paths from a in $\mathfrak{A}$, $\sigma = (a_0, \alpha_1, a_1, \dots, \alpha_n, a_n)$, where $a_0 = a$ and $(a_{i-1}, a_i) \in E^{\mathfrak{A}}_{\alpha_i}$. The transition relation E_α of $\mathfrak{A}^*_a$ corresponds to path extensions by single $E^{\mathfrak{A}}_\alpha$ edges; the unary predicate P_j in $\mathfrak{A}^*_a$ consists of those paths that end in $P^{\mathfrak{A}}_j$. Then the map that associates to every path its last element, viewed as a map $\pi\colon \mathfrak{A}^*_a \to \mathfrak{A}$, induces a bisimulation:

$$\pi\colon \mathfrak{A}^*_a \xrightarrow{\sim} \mathfrak{A} \qquad \mathfrak{A}^*_a, \sigma \sim \mathfrak{A}, \pi(\sigma).$$

It follows that every bisimulation invariant logic has the *tree model property*: satisfiability implies satisfiability in a tree model. The tree model property has important algorithmic consequences. Since it reduces satisfiability issues to problems over trees, strong classical results like Rabin's decidability result for the MSO theory of trees [42] and in particular automata theoretic methods can be brought to bear, see also [46]. The example below illustrates the usefulness of this simple insight for the (classical) model theory of modal logic, in giving an alternative proof for van Benthem's classical characterisation theorem for modal logic (a *preservation theorem* in classical model theoretic terminology). We first discuss the classical argument, though, with emphasis on the more interesting aspect of expressive completeness.

Theorem 3.1 (van Benthem). *Any bisimulation invariant first-order formula $\varphi(x) \in \mathrm{FO}[\tau]$ is equivalent to a formula of $\mathsf{ML}[\tau]$ (and, conversely, this condition is sufficient to guarantee bisimulation invariance).*

A simple compactness argument shows that, if φ is not expressible in ML, then there are $\mathfrak{A}, a \equiv^{\mathsf{M}} \mathfrak{A}', a'$ such that $\mathfrak{A} \models \varphi[a]$ while $\mathfrak{A}' \not\models \varphi[a']$. In ω-saturated elementary extensions $\hat{\mathfrak{A}} \succcurlyeq \mathfrak{A}$ and $\hat{\mathfrak{A}}' \succcurlyeq \mathfrak{A}'$, which are modally saturated, one automatically upgrades $\mathfrak{A}, a \sim_\omega \mathfrak{A}', a'$ and $\mathfrak{A}, a \equiv^{\mathsf{M}} \mathfrak{A}', a'$ to $\hat{\mathfrak{A}}, a \sim \hat{\mathfrak{A}}', a'$ (cf. the Hennessy–Milner property in Lemma 2.8), whence $\hat{\mathfrak{A}} \models \varphi[a]$ and $\hat{\mathfrak{A}}' \not\models \varphi[a']$ refutes preservation under $\sim$.

We turn to alternative approaches that work with explicit model constructions and transformations. We shall later see how this alternative approach relativises to many restricted classes (in particular also of finite models) where compactness is not available. But even in the classical context, and working over the class of all frames, such an explicit and game-based approach yields extra benefits.

Example: van Benthem's theorem via explicit constructions The following auxiliary observation is straightforward from the bisimulation game: any common upper bound on the lengths of directed paths from the elements in a bisimulation game position is also a bound on the number of rounds that can be played by **I**.

Observation 3.2. *For directed, rooted trees $\mathfrak{A}, a$ and $\mathfrak{A}', a'$ of depths $\leqslant \ell$:*

$$\mathfrak{A}, a \sim_\ell \mathfrak{A}', a' \quad \Rightarrow \quad \mathfrak{A}, a \sim \mathfrak{A}', a'.$$

Combining this with the tree model property, we find the following.

Claim 3.3. *Any ℓ-local $\varphi(x) \in \mathrm{FO}[\tau]$ that is invariant under $\sim$ is invariant under $\sim_\ell$.*

Proof. We need to show for $\mathfrak{A}, a \sim_\ell \mathfrak{A}', a'$ that $\mathfrak{A} \models \varphi[a] \Leftrightarrow \mathfrak{A}' \models \varphi[a']$. Replacing both structures by their bisimilar unfoldings from the distinguished nodes (and appealing to $\sim$ invariance of φ), then truncating both tree structures at depth ℓ (and appealing to ℓ-locality of φ), we have transformed the given situation into

$$\mathfrak{A}, a \;\simeq^{(\ell)}\; \hat{\mathfrak{A}}, \hat{a} \;\sim\; \hat{\mathfrak{A}}', \hat{a}' \;\simeq^{(\ell)}\; \mathfrak{A}', a',$$

where $\simeq^{(\ell)}$stands for isomorphism up to depth ℓ from the distinguished node. The central bisimulation equivalence is based on Observation 3.2. But now $\mathfrak{A} \models \varphi[a] \Leftrightarrow \mathfrak{A}' \models \varphi[a']$ follows by $\sim$ invariance and ℓ-locality. ∎

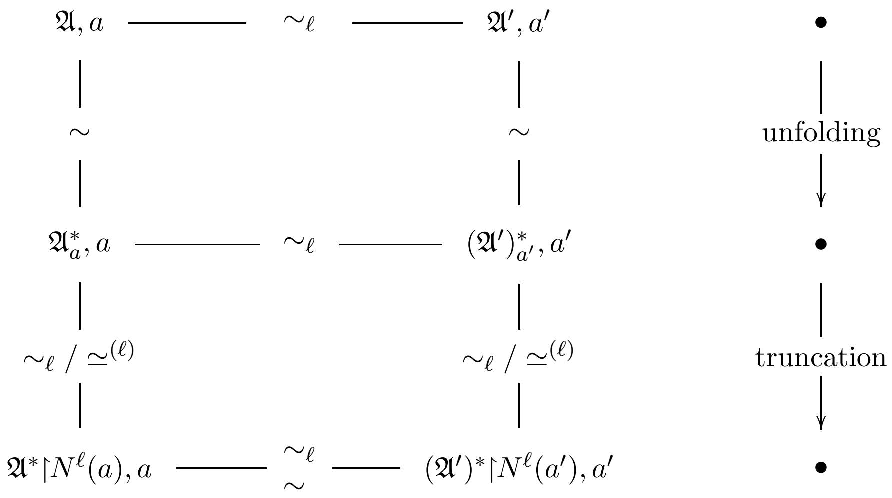
Claim 3.4. *If $\varphi(x) \in \mathrm{FO}[\tau]$ is preserved under $\sim$, then it is preserved under $\sim_\ell$ for $\ell = 2^{\mathrm{qr}(\varphi)} - 1$.*

Proof. As $\sim$ invariance implies invariance under disjoint unions, Lemma 2.4 shows that φ is ℓ-local, thus $\sim_\ell$ invariant by Claim 3.3. $\square$

As $\sim_\ell$ is of finite index, and each $\sim_\ell$ class definable in ML at nesting depth ℓ, we directly have the following version of van Benthem's theorem, which even gives a tight bound on the modal nesting depth which is not implicit in the classical proof.

Corollary 3.5. *Any quantifier-rank q formula $\varphi(x) \in \mathrm{FO}[\tau]$ that is preserved under bisimulation is equivalent to a formula of $\mathsf{ML}[\tau]$ of nesting depth $\leqslant 2^q - 1$.*

It may be worth representing the overall strategy of *upgrading a concrete level of* $\sim^\ell$ to preservation of φ in this approach. The transformations, from top to bottom in the diagram, involve firstly a tree unfolding and secondly truncation at depth ℓ. The first step preserves $\sim$, the second simultaneously preserves $\sim_\ell$ and φ (by Lemma 2.4). Consequently φ is preserved all along the vertical, but also along the bottom horizontal (as here $\sim_\ell$ guarantees full $\sim$ equivalence, by Observation 3.2). Thus φ is shown to be preserved along the top horizontal, too.



The construction of unfoldings shows that every τ transition system is bisimilar to a τ-tree, and (by taking disjoint unions of unfoldings at different elements as appropriate) globally bisimilar to a τ-forest. Obvious variations of these constructions provide acyclic companion structures that are (globally) two-way bisimilar.

As pointed out above, not every finite transition system is bisimilar to a finite acyclic system. Note that, for instance, the above proof of van Benthem's theorem fails to yield the finite model theory version (due to Rosen [43]): the argument crucially uses bisimulation invariance of φ in the transition from $\mathfrak{A}, a$ to $\mathfrak{A}^*_a$, where the target structure may be infinite.

In the case of Corollary 3.5 there is in fact an easy way out: the full (and potentially infinite) tree unfoldings of the given finite structures in the proof of Claim 3.3 can in that context be replaced by truncations to depth ℓ with isomorphic copies of the finite original structures attached at the cut-off points to yield fully bisimilar companions that are both finite and tree-like up to depth ℓ. This simple modification yields a proof of

Rosen's finite model theory analogue of van Benthem's theorem [43], including the tight bound on nesting depth in our version [41].

In connection with stronger and, in particular, global notions of bisimulation equivalence, however, better approximations to acyclicity in finite models are required. The upgrading will lead from suitable levels of finitary game equivalence to appropriate levels of local FO equivalence (Gaifman equivalence).

3.1.2 Locally acyclic bisimilar covers

Recall that a transition system is *simple* if it does not have loops or multiple edges (not even in opposite directions); it is called *ℓ-acyclic* if every ℓ-neighbourhood in its Gaifman graph is acyclic (this forbids undirected cycles of lengths up to $2\ell + 1$).

Definition 3.6. A *bisimilar cover* $\pi \colon \hat{\mathfrak{A}} \xrightarrow{\approx} \mathfrak{A}$ is a homomorphism π whose graph is a global two-way bisimulation: $\hat{\mathfrak{A}}, \hat{a} \approx \mathfrak{A}, \pi(\hat{a})$ for all $\hat{a} \in \hat{\mathfrak{A}}$. We call π *faithful* if it preserves in- and out-degrees w.r.t. each individual relation $E_\alpha \in \tau$.

A (faithful) *simple ℓ-acyclic cover* of $\mathfrak{A}$ is a (faithful) bisimilar cover $\pi \colon \hat{\mathfrak{A}} \xrightarrow{\approx} \mathfrak{A}$ by a simple ℓ-acyclic τ-structure $\hat{\mathfrak{A}}$.

Lemma 3.7. *Every finite τ transition system admits, for every ℓ, a finite faithful simple ℓ-acyclic cover.*

The construction in [41] uses for $\hat{\mathfrak{A}}$ a product of the given $\mathfrak{A}$ with a finite group G which has a generator g_e for every edge $e \in \dot{\bigcup}_\alpha E_\alpha^{\mathfrak{A}}$ and such that the Cayley graph of G w.r.t. this set of generators has girth greater than $2\ell + 1$ (compare [1] for such groups) – much as the tree unfolding could be described in terms of a product with the infinite free group of this set of generators. Over the cartesian product $A \times G$ one puts an E_α-edge precisely from (a, h) to (b, k) if $e = (a, b) \in E_\alpha^{\mathfrak{A}}$ and $k = h \circ g_e$. In this fashion, any cycle in the product projects to a cycle in the Cayley graph of G, and hence its length is bounded from below by the girth of that graph.

The following is a simple auxiliary observation towards an ℓ-local upgrading of ℓ-bisimulation equivalence to $\equiv_q$. A natural strategy for **II** can be based on maintaining full isomorphism of the substructures generated by the paths connecting the elements of the current configurations to the roots [13].

Observation 3.8. *Let $\mathfrak{A}, a \sim_\ell \mathfrak{A}', a'$ be two directed τ-trees of depths $\leqslant \ell$ with roots a and a', such that every node apart from the root is one of at least q bisimilar siblings. Then $\mathfrak{A}, a \equiv_q \mathfrak{A}', a'$. The same holds w.r.t. two-way ℓ-bisimulation equivalence in acyclic ℓ-neighbourhoods $\mathfrak{A} \restriction N^\ell(a)$ and $\mathfrak{A}' \restriction N^\ell(a')$ with at least q equivalent siblings to choose from in every node.*

Structures that have at least q equivalent successors/predecessors in every node are easily obtained by taking products with $\{1, \ldots, q\}$ in the natural manner. We write $\mathfrak{A} \mapsto \mathfrak{A} \otimes q$ for this transformation, and identify a distinguished element a with $(a, 1)$ in the new structure were appropriate.

Faithful bisimilar covers preserve this property, and can be used to achieve local acyclicity and therefore local $\equiv_q$-equivalence, viz. $\equiv_{q,0}^\ell$, by the above observation.

Example: van Benthem–Rosen once more Combining the passage to $\mathfrak{A} \otimes q$ (boosting multiplicities) with a bisimilar unfolding, one obtains a variant proof of Claim 3.4 (and through it Corollary 3.5 and its finite model theory analogue, too). Let $\mathrm{qr}(\varphi) = q$ and $\ell := 2^q - 1$. Let $\hat{\mathfrak{A}}$ be the tree unfolding from $(a, 1)$ in $\mathfrak{A} \otimes q$ (or the truncation of this unfolding glued with copies of $\mathfrak{A}$ if we want to deal with finite structures exclusively), similarly for $\mathfrak{A}', a'$.

$$
\begin{array}{ccc}
\mathfrak{A}, a & \text{——} \sim_\ell \text{——} & \mathfrak{A}', a' \\
| & & | \\
\sim & & \sim \\
| & & | \\
\hat{\mathfrak{A}}, \hat{a} & \text{——} \begin{smallmatrix} \sim_\ell \\ \equiv^\ell_{q,0} \end{smallmatrix} \text{——} & \hat{\mathfrak{A}}', \hat{a}'
\end{array}
$$

Now $\equiv^\ell_{q,0}$ equivalence in the bottom horizontal follows from Observation 3.8; preservation of φ along the bottom horizontal additionally uses Lemma 2.4 again.

Acyclic bisimilar covers really come into their own in upgradings to some target level $\equiv^\ell_{q,m}$ of Gaifman equivalence with $m > 0$, i.e., if the first-order property at hand really does express non-trivial global conditions on the existence or non-existence of certain local types – global in the sense of not only involving the ℓ-neighbourhood of the distinguished element.[10]

We look, as a typical example, at the characterisation of $\mathsf{ML}^{\overleftarrow{\forall}} \subseteq \mathrm{FO}$ in terms of invariance under $\approx$ (global two-way bisimulation) [41]. Again, we stress the expressive completeness phenomenon, as preservation of $\mathsf{ML}^{\overleftarrow{\forall}}$ under $\approx$ is obvious.

Theorem 3.9. *Both classically and in the sense of finite model theory: any first-order formula $\varphi(x) \in \mathrm{FO}[\tau]$ that is preserved under $\approx$ is equivalent to a formula of $\mathsf{ML}^{\overleftarrow{\forall}}[\tau]$.*

This follows from the following claim, based on an upgrading of $\approx_\ell$ to $\equiv^\ell_{q,m}$ in an explicit $\approx$ preserving model transformation, under which in particular the class of finite structures is closed.

Claim 3.10. *If $\varphi(x) \in \mathrm{FO}[\tau]$ is preserved under $\approx$ (over finite structures), then it is preserved under $\approx_\ell$ and hence expressible in $\mathsf{ML}^{\overleftarrow{\forall}}[\tau]$ at nesting depth ℓ, for some ℓ. Any ℓ such that φ is preserved under $\equiv^\ell_{q,m}$ for some q, m will do, i.e., the Gaifman locality radius of φ gives a bound on the necessary modal nesting depth.*[11]

Proof. We just mention the upgrading steps towards the proof of the claim, also indicated in the diagram below.

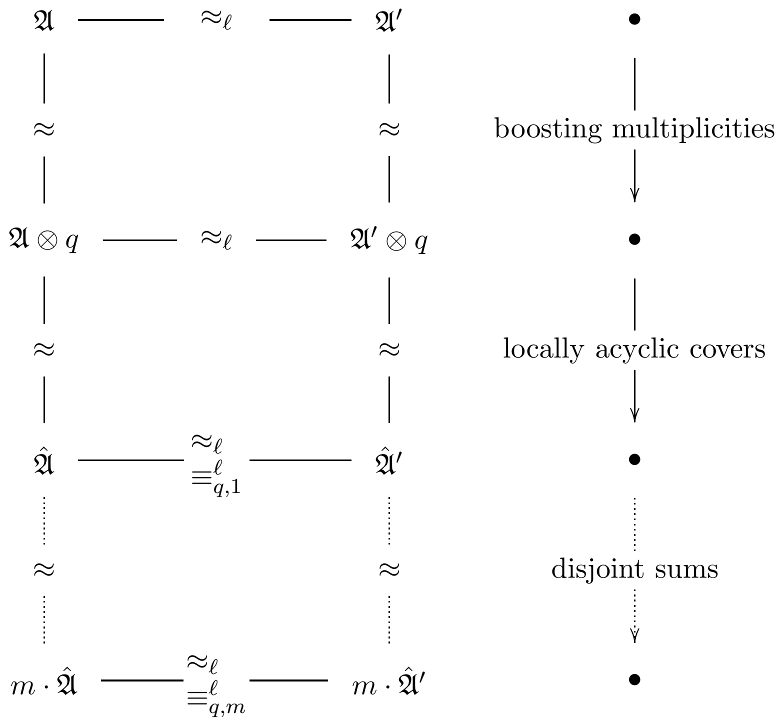
[10] See [41] for a discussion that for any $\varphi(x) \in \mathrm{FO}$ that is invariant under disjoint sums (over finite structures, or indeed over some other class which itself is closed under disjoint sums), only $\equiv^\ell_{q,m}$ for $m = 0, 1$ can matter.

[11] For simplicity, the modal nesting depth in $\mathsf{ML}^{\overleftarrow{\forall}}$ discounts $\forall/\exists$ quantifiers, which w.l.o.g. can be eliminated from within the scope of modal quantifications so that they only occur 'on the outside.'

The first step, passage to a product with $\{1, \ldots, q\}$, serves to boost all multiplicities to at least q: every E_α successor or predecessor of any node belongs to a group of at least q siblings related by automorphisms of the entire structure.

The second step yields an ℓ-acyclic bisimilar cover of the resulting structures so that the ℓ-neighbourhood of any node will be acyclic, and maintains the at-least-q-similar-siblings property due to the preservation of in- and out-degrees in faithful covers. In these circumstances, the $\approx_\ell$ relationship between the two structures guarantees $\equiv^{\ell}_{q,1}$ equivalence, by Observation 3.8.

Finally we can, if we wish, upgrade $\equiv^{\ell}_{q,1}$ further to $\equiv^{\ell}_{q,m}$, for any desired level m, by just passing to m disjoint copies of the structures obtained so far. This step guarantees that any local isomorphism type that is realised at all is a member of a scattered set of at least m many nodes of the same local isomorphism type, so that $\equiv^{\ell}_{q,1}$ implies $\equiv^{\ell}_{q,m}$. As pointed out above, however, this last upgrading can be made redundant by showing right away that φ must be preserved under some $\equiv^{\ell}_{q,1}$ (i.e., $m = 1$ suffices). □



It is clear that arguments of the kind explored here may have entirely different relativisations from the classical arguments. While classical model theoretic arguments based on compactness go through in restriction to any elementary class of structures, the above argument goes through, for instance, in restriction to any class of (finite) transition systems that is closed under $\approx$. But while this upgrading argument, and hence the expressive completeness result, does relativise to the class of all finite transition systems, it does for instance not immediately relativise to the class of connected or rooted (finite) transitions systems: clearly the last step does not preserve connectivity (and there is no immediate reason why a first-order formula $\varphi(x)$ that is invariant under

$\approx$ over connected structures should be preserved by some $\equiv^{\ell}_{q,1}$), and even the first step does not preserve rootedness.

3.2 Non-classical modal characterisation theorems

3.2.1 The general format

Analogues of the van Benthem theorem in classical and finite model theory for stronger and in particular global forms of bisimulation in the style of Theorem 3.9 are pursued in [41]. Many further natural variations of the underlying class of (finite) structures are explored in [13], with an emphasis also on methodological distinctions. In all these cases, concrete and explicit model transformations adapted to the classes at hand are used, which in many cases also provide alternative routes to characterisations over some interesting elementary classes of not necessarily finite structures.

We highlight the general format of a characterisation theorem for a fragment L of FO of this kind. Let $\mathsf{L} \subseteq \mathrm{FO}$ be a fragment of FO with

(1) equivalences $\leftrightarroweq_q$ for the relation of L-equivalence up to rank q, which we assume to have finite index; it follows that $\leftrightarroweq_q$ classes are L-definable at rank q. ($\leftrightarroweq_q$ is induced by the q-round game $\mathbb{G}^{\mathsf{L}}_q$.)

(2) the common refinement of the $(\leftrightarroweq_q)_{q\in\omega}$, $\leftrightarroweq_\omega$, capturing $\equiv^{\mathsf{L}}$. ($\leftrightarroweq_\omega$ is induced by $\mathbb{G}^{\mathsf{L}}_\omega$.)

(3) the full infinitary equivalence $\leftrightarroweq$ associated with $\mathbb{G}^{\mathsf{L}}_\infty$.

The assumptions that each $\leftrightarroweq_q$ has finite index and that $\equiv^{\mathsf{L}}$ is the limit of these finitary game equivalences reflect the 'finitary nature' of L. In this context we want to show, over a given class $\mathcal{C}$ of τ-structures, that the following are equivalent for $\varphi(x) \in \mathrm{FO}[\tau]$:

(i) φ is preserved under $\leftrightarroweq$ over $\mathcal{C}$, i.e., for all $\mathfrak{A}, a$ and $\mathfrak{A}', a'$ in $\mathcal{C}$: $\mathfrak{A}, a \leftrightarroweq \mathfrak{A}', a' \;\Rightarrow\; \big(\mathfrak{A} \models \varphi[a] \;\Leftrightarrow\; \mathfrak{A}' \models \varphi[a']\big)$.

(ii) φ is equivalent over $\mathcal{C}$ to a formula $\tilde{\varphi} \in \mathsf{L}[\tau]$, i.e., there is some $\tilde{\varphi} \in \mathsf{L}[\tau]$ s.t. for all $\mathfrak{A}, a$ in $\mathcal{C}$: $\mathfrak{A} \models \varphi[a] \;\Leftrightarrow\; \mathfrak{A} \models \tilde{\varphi}[a]$.

It is worth looking at the two implications separately:

Preservation, (ii) $\Rightarrow$ (i), is a trivial consequence of the game analysis of L-equivalence (our assumptions above). Moreover, the validity of this implication over the class of all structures trivially implies its validity in restriction to any subclass $\mathcal{C}$. In particular a *preservation* statement trivially implies its finite model theory analogue.

Expressive completeness, (i) $\Rightarrow$ (ii), is the crucial and non-trivial part of the equivalence, which is sensitive to the class $\mathcal{C}$. In particular, expressive completeness does not generally relativise to subclasses, and a classical result cannot generally be expected to persist in the sense of finite model theory.[12]

[12] An easy example of a known failure of the finite model theory version of a classically valid expressive completeness result close to our concerns is provided by $\equiv^2_\infty$ and FO^2: the class of all finite linear orderings is closed under $\equiv^2_\infty$ within the class of finite structures, but not defnable in FO^2.

If $\leftrightharpoons_\omega$ coincides with $\leftrightharpoons_\infty$ in ω-saturated structures, as is typically the case,[13] then expressive completeness of L for first-order properties that are preserved under $\leftrightharpoons$ over the class of all τ-structures follows from the assumptions (along the lines of the classical proof outlined for van Benthem's theorem above, for instance).

Under the assumptions made, expressibility of φ in L (over $\mathcal{C}$) is equivalent to preservation of φ under some level of $\leftrightharpoons_\ell$ (over $\mathcal{C}$). Therefore, the expressive completeness of L for $\leftrightharpoons$ invariance over $\mathcal{C}$ is equivalent (for any $\mathcal{C}$) to the implication

$$\begin{array}{rl} & \varphi(x) \text{ preserved under } \leftrightharpoons \text{ (over } \mathcal{C}) \\ \Rightarrow & \varphi(x) \text{ preserved under } \leftrightharpoons_\ell \text{ (over } \mathcal{C}) \text{ for some } \ell \in \omega, \end{array}$$

which is a particular 'compactness property' that may or may not be valid, depending on the nature of $\leftrightharpoons$ and $\mathcal{C}$. The classical manner of establishing this compactness property, as well as the alternative explicit and game-oriented constructions indicated above may both be cast as upgradings of equivalences, albeit in orthogonal directions. The juxtaposition of the generic diagrams below may serve to make this distinction apparent. While the classical upgrading involves a transformation of structures up to full FO equivalence (passage to ω-saturated elementary extensions say) to boost $\leftrightharpoons_\omega$ to $\leftrightharpoons$, the alternative upgrading consists of a transformation of structures up to full (infinitary) $\leftrightharpoons$ to boost a concrete finitary level of $\leftrightharpoons_\ell$ to an approximate level $\doteqdot$ of first-order equivalence that is good enough to preserve φ. In the examples encountered here, $\doteqdot$ is either some level $\equiv_q$ or $\equiv^\ell_{q,m}$. The following two sections will review and summarise some of the results obtained along these lines in [13].

$$\begin{array}{ccc} \mathfrak{A}, a & \text{——} \leftrightharpoons_\omega \text{——} & \mathfrak{A}', a' \\ | & & | \\ \equiv & & \equiv \\ | & & | \\ \mathfrak{A}^*, a & \text{——} \leftrightharpoons \text{——} & (\mathfrak{A}')^*, a' \end{array} \qquad \begin{array}{ccc} \mathfrak{A}, a & \text{——} \leftrightharpoons_\ell \text{——} & \mathfrak{A}, a' \\ | & & | \\ \leftrightharpoons & & \leftrightharpoons \\ | & & | \\ \hat{\mathfrak{A}}, a & \text{——} \underset{\varphi}{\doteqdot} \text{——} & \hat{\mathfrak{A}}', a' \end{array}$$

3.2.2 Explicit upgrading through local control

By approximating FO equivalence by a concrete level of Gaifman equivalence we shift the emphasis to local control over FO equivalence. This allows us to make use of explicit model constructions that lead to locally acyclic structures, as in Lemma 3.7, which means that *locally* $\approx^\ell$ can be upgraded to $\equiv_q$ (if multiplicities have been boosted in preparation) via Observation 3.8. For characterisations of $\sim_\forall$ invariance rather than $\approx$ (global but only forward bisimulation, related to $\mathsf{ML}^\forall$), a correspondingly higher level of global ℓ_0-bisimulation equivalence can first be upgraded (in a transformation up to full global forward bisimulation $\sim_\forall$) to $\approx^{\ell_1}$, which can then be further upgraded to some

[13] Our discussion of saturation and the Hennessy–Milner property in section 2 and especially section 2.4 indicates that this is true whenever the corresponding $\mathcal{O}^{\mathsf{L}}(\mathfrak{A})$ is uniformly FO-interpretable over $\mathfrak{A}$.

$\equiv^{\ell}_{q,m}$ as above. In this manner, for example the expressive completeness results below are proved in [13].

A *rooted* structure is a τ-structure $\mathfrak{A}, a$ with distinguished element a as a *root* from which all elements of $\mathfrak{A}$ are reachable on directed paths. For tree structures compare section 1.2. Note that even the class of not necessarily finite rooted structures is not elementary. Also note that for rooted structures, the full infinitary equivalences $\sim_{\forall}$ and $\sim$ coincide at the roots, while the finite levels clearly do not.

Theorem 3.11. $\mathsf{ML}^{\forall}$ *is expressively complete for first-order properties that are preserved under* $\sim$ *over the following classes* $\mathcal{C}$ *of structures:*

(i) the class of rooted structures.

(ii) the class of finite rooted structures.

(iii) the class of tree structures.

(iv) the class of finite tree structures.

Another natural and classically important class of transition systems (as Kripke structures in the context of knowledge representation) is the class of *equivalence structures*: τ-structures in which all transition relations E_α are interpreted as equivalence relations. And even though transitivity requirements tend to trivialise locality analysis (also compare the next section), equivalence structures are amenable to an analysis and to upgrading transformations based on locally acyclic covers. Here FO interpretations can be used to adapt both the construction of suitable covers and the analysis of bisimulation invariant FO properties. As far as local acyclicity in bisimilar covers is concerned, the following can be obtained from Lemma 3.7 via simple FO translations.

Lemma 3.12. *Every finite equivalence structure admits, for every* ℓ*, a faithful bisimilar cover by some finite equivalence structure in which*

(i) any two equivalence classes (w.r.t. to distinct E_α*) intersect in at most one element,*

(ii) all cycles of lengths up to $2\ell + 1$ *stay within a single* E_α *class for some* α*.*

Over such essentially ℓ-acyclic structures, an analogue of Observation 3.8 is available to show that global ℓ-bisimulation can be upgraded to $\equiv^{\ell}_{q,m}$ for any required level of q and m. Therefore, $\sim$ invariance implies $\sim_\ell$ invariance also over the class of finite equivalence structures.

Corollary 3.13. $\mathsf{ML}^{\forall}$ *is expressively complete for first-order properties that are preserved under global bisimulation* $\sim_{\forall}$ *over the class of finite equivalence structures.*

3.2.3 Explicit upgrading through decomposition

Locality arguments cannot be used to great effect over structures that trivialise Gaifman locality. For instance, the Gaifman graph of directed transitive trees (trees with a partial order) has diameter 2, and $\equiv^{\ell}_{q,m}$ is essentially just $\equiv_q$, for $\ell \geqslant 1$. On some related and particularly interesting classes of transition systems with one transitive transition relation, however, one may instead base expressive completeness proofs for modal fragments on another classical constructive approach to the analysis of games: composition arguments w.r.t. order. We saw a glimpse of this in the Ehrenfeucht–Fraïssé analysis of finite linear orderings in section 2.2 (Lemma 2.3).

We consider the example of rooted, irreflexive transitive tree structures with a single transition relation E: $\mathfrak{A} = (A, E^{\mathfrak{A}}, (P_i^{\mathfrak{A}}))$ with distinguished root a, with a transitive and irreflexive partial order relation $E^{\mathfrak{A}}$ such that the set of E-predecessors of any element $b \in A$ is well-ordered by $E^{\mathfrak{A}}$ with minimal element a. For succinctness we refer to such structures as $\prec$-trees. The class of all $\prec$-trees (finite and infinite ones) is non-elementary (due to the well-foundedness condition); and so is the class of all finite $\prec$-trees (due to the finiteness condition).

We review the key decomposition idea from [13] that allows us to upgrade ℓ-bisimulation equivalence between (finite) $\prec$-trees $\mathfrak{A}, a \sim_\ell \mathfrak{A}', a'$ to quantifier-rank q first-order equivalence $\equiv_q$ through a transformation that preserves full bisimulation equivalence.

In a preparatory step, we boost multiplicities and unravel in order to achieve some homogeneity w.r.t. paths in $\prec$-trees.

For a given q let the $\prec$-trees $\mathfrak{A}_0^q$ and $\mathfrak{A}_{q-1}^q$ (an expansion of $\mathfrak{A}_0^q$ by colours for certain $\equiv_{q-1}$ types) be obtained from $\mathfrak{A}, a$ as follows.

The universe and the interpretation of the unary predicates of $\mathfrak{A}_0^q$ are those of the bisimilar unfolding of $\mathfrak{A} \otimes \{1, \ldots, q\}$ from one of the representatives of the root a (say we identify a with $(a, 1)$); for its transition relation we pass to the transitive closure of the transition relation in the unfolding. It is easily checked that this transformation leads to a bisimilar $\prec$-tree $\mathfrak{A}_0^q$, which is finite if $\mathfrak{A}$ is. Even for infinite $\mathfrak{A}$ the $\prec$-tree $\mathfrak{A}_0^q$ has predecessor sets that are finite linear orderings rather than arbitrary well-orderings. In addition, due to the unfolding step in its construction, $\mathfrak{A}_0^q$ has the following useful representation property for its paths. Any path $a_0 = a, a_1, \ldots, a_n$ from the root in $\mathfrak{A}_0^q$, can be matched with some *full path* $\hat{a}_0 = a, \hat{a}_1, \ldots, \hat{a}_n$ consisting of the full predecessor set of the target node $\hat{a}_n$ in $\mathfrak{A}_0^q$, such that a_i and $\hat{a}_i$ are not only bisimilar but even are the roots of isomorphic subtrees.

Towards an inductive analysis of $\equiv_q$, we use $\mathfrak{A}_{q-1}^q$, which is the expansion of $\mathfrak{A}_0^q$ with new unary predicates that colour every node with the $\equiv_{q-1}$-class of the subtree rooted at this node in $\mathfrak{A}_0^q$.

In order to show how suitable levels of ℓ-bisimulation between $\prec$-trees $\mathfrak{A}, a$ and $\mathfrak{A}', a'$ can be upgraded to $\equiv_q$ equivalence in bisimilar $\prec$-trees, we firstly replace $\mathfrak{A}$ and $\mathfrak{A}'$ by the $\prec$-trees $\mathfrak{A}_0^q, a \sim \mathfrak{A}, a$ and $(\mathfrak{A}')_0^q, a' \sim \mathfrak{A}', a'$. It then suffices to show, in the context of an induction on q, that for some sufficiently large ℓ (depending on q):

$$(*) \qquad \mathfrak{A}_{q-1}^q, a \sim_\ell (\mathfrak{A}')_{q-1}^q, a' \quad \Rightarrow \quad \mathfrak{A}_0^q, a \equiv_q (\mathfrak{A}')_0^q, a'.$$

For this, a composition argument can be used towards a reduction to the analysis of Ehrenfeucht–Fraïssé games over finite coloured linear orderings. We associate with an element b in $\mathfrak{A}^q, a$ the coloured finite linear ordering $\mathfrak{I}_b$ induced on the interval $[a, b]$ in $\mathfrak{A}^q$; similarly $\mathfrak{I}'_{b'}$ with any b' in $(\mathfrak{A}')^q, a'$. Then

$$\mathfrak{I}_b, a, b \equiv_{q-1} \mathfrak{I}'_{b'}, a', b' \quad \Rightarrow \quad \mathfrak{A}_{q-1}^q, a, b \equiv_{q-1} (\mathfrak{A}')_{q-1}^q, a', b',$$

due to compositionality of strategies in the games. A winning strategy for **II** in the remaining $(q-1)$-round game on the $\prec$-trees can be based on

(a) a strategy in the $(q-1)$-round game on the induced linear orderings: this provides a match between subtrees rooted along the coloured paths $[a, b]$ and $[a', b']$.

(b) strategies to play within colour-matched subtrees based on their $\equiv_{q-1}$ equivalence.

Therefore, it suffices to guarantee that for every b there is some b' (and vice versa, for every b' a b) such that $\mathfrak{I}_b, a, b \equiv_{q-1} \mathfrak{I}'_{b'}, a', b'$, provided only that $\mathfrak{A}^q_{q-1}, a \sim_\ell (\mathfrak{A}')^q_{q-1}, a'$. A bound on such an ℓ can now be extracted from the Ehrenfeucht–Fraïssé game on finite coloured linear orderings. The following is a consequence of the compatibility of the game with ordered sums or concatenation (we leave it as a nice exercise; see [14] and also [13] for details).

Observation 3.14. *There is a bound N (depending on q and the number of colours) such that any finite coloured linear ordering (with constants for the first and last elements) of length greater than N has some proper $\equiv_{q-1}$ equivalent substructure.*

In the case of the finite coloured orderings $\mathfrak{I}_b$ this means that, up to $\equiv_{q-1}$, only those of lengths up to N need to be taken into account (any substructure of an $\mathfrak{I}_b$ is realised as $\mathfrak{I}_{\hat{b}}$ for suitable $\hat{b}$ by the homogeneity property of $\mathfrak{A}^q_{q-1}$). But the isomorphism types of (substructures of) $\mathfrak{I}_b$ of size up to N are clearly governed by the $\sim_{N-1}$ type of $\mathfrak{A}^q_{q-1}, a$, whence we get $(*)$ for $\ell = N - 1$.

Based on this decomposition approach, the following are obtained in [13].

Theorem 3.15. ML *is expressively complete for first-order properties that are preserved under bisimulation over the following classes $\mathcal{C}$ of partially ordered trees:*

(i) the class of irreflexive transitive trees.

(ii) the class of finite irreflexive transitive trees.

While the classes of rooted reflexive transitive structures or reflexive transitive trees display a similar behaviour [13], the picture changes if reflexivity is not uniformly prescribed. For transitive tree-like structures in which some nodes *may* be reflexive, a marked difference between finite and not necessarily finite structures becomes important. The first-order formula

$$\varphi(x) = \exists y(Exy \land Eyy),$$

expressing accessibility of a reflexive node, is

(a) invariant under bisimulation over the class of *finite* transitive structures, but

(b) not invariant under bisimulation over the class of all transitive structures.

Point (b) is illustrated by the simple example of the infinite irreflexive unfolding of a structure consisting of a single reflexive node. For (a) consider finite transitive structures $\mathfrak{A}, a \sim \mathfrak{A}', a'$ and assume that $\mathfrak{A} \models \varphi[a]$. Consider a play of **I** from a to some reflexive b in $\mathfrak{A}$ followed by a sequence of stationary moves at b (b is reflexive) that is long enough to force the sequence of responses by **II** to visit some node b' twice: as b' is on a cycle, it is reflexive.

[13] shows that an extension of basic modal logic with a modality as suggested by φ above, asserting that there is some *reflexive* successor satisfying ψ, is expressively complete for bisimulation invariant first-order properties over *finite* transitive tree-like structures. For expressive completeness over the wider classes of all finite transitive structures a stronger variant of this new modality is required, which also captures reachability of an E-clique (rather than a single reflexive node) realising several distinct formulae. As indicated above, such extra modalities are necessary in the finite, but not compatible

with bisimulation in transitive structures in general. (In fact it is not finiteness, but the absence of infinite strictly forward-directed E-paths, that matters, see [13].)

Over finite transitive structures and some related restricted classes of transitive transition systems, the decomposition based analysis in [13] also extends from first-order to monadic second-order logic.

Among the long open questions in this area remain the finite model theory status of

- the Janin–Walukiewicz result [33] that the modal μ-calculus is expressively complete for monadic second-order properties preserved under bisimulation, and
- expressive completeness of the guarded fragment for the first-order properties preserved under guarded bisimulation, established in the classical setting in [2].

The second issue, concerning guarded bisimulation as a generalisation of modal bisimulation, also leads over to the following section.

4 From graphs to hypergraphs

The guarded fragment of FO and, more fundamentally, the concept of guarded bisimulations (compare section 2.3.2) point to a hypergraph structure induced by a relational structure, over and above the graph structure embodied in the Gaifman graph. With the relational τ-structure $\mathfrak{A}$ we can associate the hypergraph of guarded subsets of $\mathfrak{A}$, whose universe is the universe A of $\mathfrak{A}$ and whose hyperedges are the guarded subsets $s \subseteq A$ of $\mathfrak{A}$:

$$H(\mathfrak{A}) = \big(A, \{s \subseteq A \colon s \text{ a guarded subset }\}\big).$$

Generally, with any hypergraph $H = (A, S)$, one also associates the graph over the same universe A whose edge relation is precisely the union of the cliques induced by the hyperedges of H:

$$G(H) = (A, E) \quad \text{where} \quad E = \textstyle\bigcup_{s \in S} \{(a, b) \colon a, b \in s, a \neq b\}.$$

In the case of the hypergraph $H(\mathfrak{A})$ this just returns the Gaifman graph $G(\mathfrak{A})$.

The graph $G(H)$, however, contains less information, since not every clique in $G(H)$ need be induced by a hyperedge. The complete graph on three elements, K_3, for instance, occurs as $G(H)$ for $H = K_3$ as well as for any hypergraph that has the full set of three elements as one of its hyperedges. In the classical literature on hypergraphs [6], a hypergraph H such that all cliques in $G(H)$ are induced by hyperedges is called *conformal*; conformality plays a role in acyclicity criteria for hypergraphs. In the next section we briefly look at the natural notion of hypergraph bisimulation and discuss corresponding notions of acyclicity and unfoldings.

4.1 Hypergraph bisimulation

If we disregard the local relational content in guarded bisimulations, i.e., if we relax the soundness condition on positions in the game from local isomorphism of relational substructures to just local bijections, we obtain a natural notion of hypergraph bisimulation. Guarded bisimulations become a special case of hypergraph bisimulations between the associated hypergraphs of guarded subsets. For questions of acyclicity and of tree decomposability, the actual local relational content does not matter and it makes sense to work with the more fundamental notion of hypergraph bisimulation.

The hypergraph bisimulation game The positions in the bisimulation game on hypergraphs $H = (A, S)$ and $H' = (A', S')$ are local bijections $\rho\colon s \to s'$ between hyperedges $s \in S$ and $s' \in S'$. The challenge/response exchange between players **I** and **II** in a single round, from position $\rho\colon s \to s'$, is played as follows:

- **I** selects either some hyperedge $t \in S$ or some hyperedge $t' \in S'$;
- **II** has to respond with a position $\sigma\colon t \to t'$ (involving the hyperedge proposed by **I** and a match with a hyperedge in the opposite structure) such that ρ agrees with σ on the overlap (between s and t if **I** chose t, or between s' and t' if **I** chose t').

II loses if she has no such response. Otherwise, winning conditions in the q-round game, the finite-round game and the infinite game are as usual. We correspondingly define equivalences in terms of winning positions for **II**.

Definition 4.1. For hypergraphs $H = (A, S)$ and $H' = (A', S')$: $H, \mathbf{a} \sim_q H, \mathbf{a}'$ if the bijection $\rho\colon \mathbf{a} \mapsto \mathbf{a}'$ is a winning position in the q-round bisimulation game on the hypergraphs H and H'. Equivalences $H, \mathbf{a} \sim_\omega H', \mathbf{a}'$ and $H, \mathbf{a} \sim H', \mathbf{a}'$ are similarly defined w.r.t. the finite-round and infinite games.

Definition 4.2. A *bisimilar cover* of the hypergraph $H = (A, S)$ by the hypergraph $\hat{H} = (\hat{A}, \hat{S})$ is a map $\pi\colon \hat{A} \to A$ such that

(i) π is injective in restriction to every $\hat{s} \in \hat{S}$.

(ii) $S = \{\pi(\hat{s})\colon \hat{s} \in \hat{S}\}$.

(iii) π comprises a winning strategy for **II** in the infinite bisimulation game in the sense that **II** can maintain positions in which hyperedges are matched through π.

Consider the special case of $H = H(\mathfrak{A})$, the hypergraph of guarded subsets of the τ-structure $\mathfrak{A}$. It is not hard to see that any bisimilar cover $\pi\colon \hat{H} \to H$ by a hypergraph $\hat{H} = (\hat{A}, \hat{S})$ induces a *guarded cover*

$$\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A},$$

where $\hat{\mathfrak{A}}$ is simply obtained by pulling the relational interpretation on A back to $\hat{A}$ in such a way that every restriction of π to a hyperedge of $\hat{H}$ becomes a local isomorphism. One checks that this leads to a well-defined interpretation of a τ-structure over the universe $\hat{A}$, for which indeed also $H(\hat{\mathfrak{A}}) = \hat{H}$. In particular π now comprises a winning strategy for **II** in the infinite guarded bisimulation game on $\hat{\mathfrak{A}}$ and $\mathfrak{A}$ (compare (iii) above). These simple considerations suggest to view hypergraph bisimulation just as 'guarded bisimulations without relations' – or to view guarded bisimulation as a relational incarnation of a possibly more fundamental notion of hypergraph bisimulation.

4.2 Tree-likeness: acyclicity criteria

Full acyclicity (in the hypergraph sense) can be achieved, up to bisimulation, through a process of bisimilar unfolding in close analogy with the tree unfolding of transition systems. We present this basic construction before relating it to the relevant notions of acyclicity and tree-likeness that it exemplifies.

Bisimilar hypergraph unfolding Consider a hypergraph $H = (A, S)$. We want to find a tree-like hypergraph $\hat{H}$ that provides a bisimilar cover for H; while overlaps between hyperedges have to be reproduced in $\hat{H}$, it should otherwise and in particular globally be as free (free of incidental overlaps) as possible. The construction follows the idea of a tree unfolding of a transition system, but instead of nodes, subsets need to be joined – joined through identifications in overlaps as prescribed in H, compare [22].

With H firstly associate the tree S^* of all finite sequences of hyperedges, with a successor relation linking a sequence $\sigma \in S^*$ to its immediate extensions $\sigma\hat{\ }s$ for $s \in S$. We obtain the universe $\hat{A}$ of the desired hypergraph $\hat{H}$ as a quotient of the following auxiliary set D, which may be seen as a disjoint union of path-labelled copies of hyperedges $s \in S$:

$$D := \big\{(\sigma\hat{\ }s, a) \in S^+ \times A \colon a \in s\big\} \subseteq S^* \times A.$$

In this set, we want to identify same elements in nodes that are labelled with next-neighbour paths. Let $\doteq$ be the reflexive, symmetric, transitive closure of the relation that links (σ, a) to $(\sigma\hat{\ }s, a)$ in D. In the following we write $[\sigma, a]$ for the $\doteq$ equivalence class of $(\sigma, a) \in D$. We put

$$\begin{array}{rcl} \hat{A} & := & D \,/\, \doteq, \\ \hat{S} & := & \big\{\hat{s}_\sigma \colon s \in S, \sigma \in S^*\big\}, \\ & & \text{where } \hat{s}_\sigma = \{[\sigma\hat{\ }s, a] \colon a \in s\} \text{ for } \sigma \in S^*, s \in S. \end{array}$$

One checks that $\pi \colon \hat{H} \to H,\ \ [\sigma, a] \mapsto a$ is well-defined and a bisimilar hypergraph cover. In line with the above remarks, if the same construction is applied to the hypergraph $H = H(\mathfrak{A})$ associated with the guarded subsets of a τ-structure $\mathfrak{A}$, then the obvious expansion of $\hat{A}$ to a τ-structure yields a guarded bisimilar cover of $\mathfrak{A}$. In both cases, the tree structure of S^* also provides a tree decomposition of the new hypergraph $\hat{H}$, or of the τ-structure $\hat{\mathfrak{A}}$.

Definition 4.3. A *tree decomposition* of $H = (A, S)$ consists of a tree T together with a surjective map $\rho \colon T \to S$ such that for every $a \in A$ the subset $\{t \in T \colon a \in \rho(t)\} \subseteq T$ is connected in T.

It may be intuitive that the existence of a tree decomposition is an acyclicity condition. Consider a tree decomposition ρ of a finite hypergraph H with finite tree T. One can use ρ to reduce H to the empty hypergraph by repeated application of the following two reduction steps

- removal of an element $a \in A$ that is covered by at most one hyperedge (more precisely, a is removed from A and from the hyperedge covering it).
- removal of a hyperedge s that is contained in some other hyperedge that is retained.

For the claimed reduction, essentially just proceed from the leaves of T: a leaf of T is mapped to a hyperedge that is either contained in the hyperedge at its predecessor node, or it contains some elements not covered by any other hyperedge. Removal of hyperedges or elements based on this procedure is compatible with maintaining a tree decomposition.[14]

[14] Note that to deal with infinite hypergraphs, it is necessary to phrase the reduction condition for finite sub-hypergraphs rather than the full graph; e.g., a two-way infinite edge chain is not decomposable as such.

If we transfer this notion of a hypergraph tree decomposition to relational structures (cf. Definition 5.1 for tree decompositions in that sense), there is an important difference: the usual notion of tree decomposition is more liberal in allowing arbitrary subsets of A to be associated with the nodes of the representation tree, while here we would only admit guarded sets. A cycle (viewed as a hypergraph with size 2 hyperedges) does not admit a hypergraph tree decomposition, but it does admit tree decompositions based on subsets of size 3. We return to ordinary tree decompositions of relational structures in section 5.1 below.

It follows that every logic invariant under guarded bisimulation (i.e., whose formulae are preserved under guarded bisimulations) has a *bounded treewidth model property* or *generalised tree model property* [20]. This property is of great value in the algorithmic model theory of GF and of its extensions that still are sublogics of GF_∞ like guarded fixpoint logic [24], because it allows a reduction of satisfiability issues to the model theory of trees, via a coding of models in tree representations.

Proposition 4.4 (Grädel)**.** GF *has the following generalised tree model property: any satisfiable* $\varphi \in \mathsf{GF}[\tau]$ *is satisfiable in a model that admits a tree decomposition w.r.t. guarded subsets, and in particular one of treewidth less than the width of* τ.[15]

Returning to hypergraphs, the classical criterion for hypergraph acyclicity is the following. As shown in [5] it coincides (for finite hypergraphs) with hypergraph tree decomposability in the sense of Definition 4.3 above, as well as with several other criteria. For classical hypergraph theory compare [6].

Definition 4.5. A hypergraph $H = (A, S)$ with associated graph $G(H)$ is called *acyclic* if it satisfies the following two conditions:

(i) *conformality:* every clique in $G(H)$ is contained in some hyperedge of H.

(ii) *chordality:* every cycle of length at least 4 in $G(H)$ has a chord: there are two nodes that are not next neighbours along the cycle that are linked (by an edge of $G(H)$/hyperedge of H).

Clearly hypergraph unfoldings are acyclic in this sense, so that every hypergraph admits a bisimilar cover by an acyclic hypergraph. The following, however, is open.

Question 4.6. *Does every* finite *hypergraph admit bisimilar covers by* finite *ℓ-acyclic hypergraphs, for all ℓ?*

Here one hopefully sensible notion of ℓ-acyclicity would be to postulate that all non-trivial, chordless cycles must have lengths greater than $2\ell + 1$. Note that this is different form the requirement that the induced hypergraphs on ℓ-neighbourhoods in $G(H)$ be acyclic. For that latter, stronger notion, the answer to the question is negative.

Example Consider a cartwheel hypergraph H_n consisting of an exterior cycle of nodes $a_1, \ldots, a_n, a_1$ plus a central node a, and with hyperedges $\{a, a_i, a_{i+1}\}$ for $i \in \mathbb{Z}/n\mathbb{Z}$. The exterior cycle of length n is without chord, and any bisimilar cover of H_n will still have cycles in the 1-neighbourhood of any node related to a, albeit possibly longer cycles.

[15] Treewidth is defined to be one less than the maximal size of sets needed in a tree decomposition, here bounded by the width of τ minus 1.

A positive resolution to Question 4.6 would possibly be a starting point for proving the finite model theory analogue of the classical characterisation theorem for GF, due to [2]. So far only the graph case, or the case of $\mathsf{GF}[\tau]$ for relational vocabularies of width 2, is settled positively in [41].

4.3 Excursion: extension properties

In contrast with the open status of Question 4.6, we know from [30] that conformality can always be achieved in finite bisimilar hypergraph covers. The basic idea towards the construction of such covers in [30] is quite simple – and surprisingly contrary to the intuition of an unfolding. It essentially focuses on the footprints of forbidden cliques in the associated graph $G(\hat{H})$. We illustrate the key idea with a (generic) local example of the task.

Let, for instance, $H = (A, S)$ be a finite hypergraph with a tuple of pairwise distinct nodes $\mathbf{a} = (a_1, \dots, a_n)$ such that $[\mathbf{a}] = \{a_1, \dots, a_n\}$ is a clique in $G(H)$ not contained in any hyperedge of H. We want to construct a bisimilar cover $\pi \colon \hat{H} \to H$ by a finite hypergraph $\hat{H} = (\hat{A}, \hat{S})$ such that no lift $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_n)$ with $\hat{a}_i \in \pi^{-1}(a_i)$ forms a clique in $G(\hat{H})$. Let $A_0 := A \setminus [\mathbf{a}]$ and put

$$\hat{A} := A_0 \cup ([\mathbf{a}] \times \{1, \dots, n-1\}); \qquad \pi \restriction A_0 = \mathrm{id}, \quad \pi(a_j, i) = a_j.$$

We now set $\hat{S}$ to be the set of all subsets $\hat{s} \subseteq \hat{A}$ such that

(i) $\pi \restriction \hat{s}$ is a bijection onto some $s \in S$.

(ii) for $(a_j, i), (a_{j'}, i') \in \hat{s}$, if $(a_j, i) \neq (a_{j'}, i')$, then $i \neq i'$:
any two distinct nodes in $\hat{s}$ above $\mathbf{a}$ must have distinct tags in $\{1, \dots, n-1\}$.

On one hand, one checks that $\pi \colon \hat{H} \to H$ is a bisimilar cover. Crucially, the back-and-forth conditions do not give rise to requirements (of the *back* kind) to produce a hyperedge $\hat{s}$ whose projection to A would cover all of $[\mathbf{a}]$: this is clear, since $[\mathbf{a}]$ is not contained in any hyperedge of H. On the other hand, condition (ii) rules out the possibility of a clique in $G(\hat{H})$ above $\mathbf{a}$: if each pair of components in $\hat{\mathbf{a}}$ were to be linked by a hyperedge, then they would have to have pairwise distinct tags, which is impossible simply by the pigeon-hole principle.

A uniform application of this idea, for all forbidden cliques simultaneously, yields a finite conformal cover which moreover has useful automorphism properties [30].

An automorphism of a hypergraph is a permutation of its universe that preserves the set of hyperedges. We say that the cover $\pi \colon \hat{H} \to H$ *lifts automorphisms* of H if for every automorphism ρ of H there is an automorphism of $\hat{H}$ such that $\rho \circ \pi = \pi \circ \hat{\rho}$. The cover is *homogeneous*, if for every pair of hyperedges $\hat{s}_1, \hat{s}_2 \in \hat{S}$ above the same $s \in S$, there is some automorphism σ of $\hat{H}$ mapping $\hat{s}_1$ to $\hat{s}_2$.

Lemma 4.7. *Every finite hypergraph $H = (A, S)$ admits a bisimilar cover $\pi \colon \hat{H} \to H$ by some finite conformal hypergraph $\hat{H} = (\hat{A}, \hat{S})$.*

Every finite relational τ-structure $\mathfrak{A}$ admits a guarded cover by some finite τ-structure $\hat{\mathfrak{A}}$ whose hypergraph of guarded subsets $H(\hat{\mathfrak{A}})$ is conformal.

Moreover, the cover can be chosen homogeneous and such that it lifts all automorphisms of the base structure.

Herwig–Lascar extension theorems, EPPA A *local automorphism* of a τ-structure $\mathfrak{A}$ is a partial bijection p of A that is an isomorphism between the substructures induced on $\mathrm{dom}(p)$ and $\mathrm{image}(p)$. The following extension theorem for local automorphisms is from [26], also compare [28].

Theorem 4.8 (Herwig)**.** *Let $\mathfrak{A}_0$ be a finite τ-structure. Then there is a finite extension $\mathfrak{A}_1 \supseteq \mathfrak{A}_0$ such that every local automorphism of $\mathfrak{A}_0$ extends to a full automorphism of $\mathfrak{A}_1$. $\mathfrak{A}_1$ can be chosen such that every guarded subset of $\mathfrak{A}_1$ is the image under some automorphism of $\mathfrak{A}_1$ of a guarded subset of $\mathfrak{A}_0$.*

The last condition is in the given situation in fact equivalent to saying that, for every relation $R \in \tau$:

$$R^{\mathfrak{A}_1} = \bigcup_{\rho \in \mathrm{Aut}(\mathfrak{A}_1)} \rho(R^{\mathfrak{A}_0}).$$

If $\mathfrak{A}_1'$ at first only satisfies the automorphism extension property, and $G' = \mathrm{Aut}(\mathfrak{A}_1')$, then replacing $R^{\mathfrak{A}_1'}$ by $\bigcup_{\rho \in G'} \rho(R^{\mathfrak{A}_0})$ preserves the automorphism extension property and yields a structure that also satisfies the additional requirement on guarded subsets. A combination with Lemma 4.7 then gives the following strengthening of the theorem [30].

Corollary 4.9. *For every finite $\mathfrak{A}_0$ there is a finite extension $\mathfrak{A}_2 \supseteq \mathfrak{A}_0$ such that every local automorphism of $\mathfrak{A}_0$ extends to a full automorphism of $\mathfrak{A}_2$ and such that every clique in $G(\mathfrak{A}_2)$ is the image under some automorphism of $\mathfrak{A}_2$ of some clique in $G(\mathfrak{A}_0)$.*

Proof. Let $\mathfrak{A}_1 \supseteq \mathfrak{A}$ as in Theorem 4.8. Let $H_1 = (A_1, S)$ be the hypergraph with hyperedges

$$S = \big\{\rho(A_0) \colon \rho \in \mathrm{Aut}(\mathfrak{A}_1)\big\}.$$

We may now apply Lemma 4.7 to obtain a conformal bisimilar cover $\pi \colon \hat{H} \to H_1$ with hypergraph $\hat{H} = (\hat{A}, \hat{S})$. The desired τ-structure $\mathfrak{A}_2 = \hat{\mathfrak{A}}$ is obtained by interpreting the relations over the universe $\hat{A}$ such that, for every $\hat{s} \in \hat{S}$, the local bijection $\pi \restriction \hat{s} \colon \hat{s} \to s$ becomes a local isomorphism between $\hat{\mathfrak{A}}\restriction\hat{s}$ and $\mathfrak{A}_1\restriction s$. $\mathfrak{A}_0$ may be isomorphically embedded into this new structure $\hat{\mathfrak{A}}$ by singling out any particular $\hat{s} \in \hat{S}$ above $s = A_0 \in S$. The automorphism properties of the cover as stated in Lemma 4.7 guarantee that the local automorphisms of the embedded $\mathfrak{A}_0$ still extend to automorphisms of $\hat{\mathfrak{A}}$. And $G(\hat{\mathfrak{A}})$ does not have any cliques other than those that are unavoidable automorphic copies of cliques already present in the embedded $\mathfrak{A}_0$: this is a consequence of the conformality of $\hat{H}$ and the fact that $G(\hat{\mathfrak{A}})$ consists of the union of the $G(\hat{\mathfrak{A}})\restriction\hat{s}$ for $\hat{s} \in \hat{S}$, each of which is an isomorphic copy of $G(\mathfrak{A}_0)$ by construction. □

Further corollaries of this are (simpler proofs of) the extension theorem for local automorphisms within the class of finite triangle-free graphs [26], the class of finite clique-free graphs [27], or the class of finite relational structures with conformal hypergraphs of guarded sets.

The corollary as stated has also been employed in [30] to yield a very transparent proof of the finite model property of the clique guarded fragment, just as Theorem 4.8 itself yields a very natural proof of the finite model property for basic GF first given by Grädel [20].

5 Locality and special classes of relational structures

5.1 Tree-decompositions and treewidth

Bounded treewidth has emerged as one central notion of 'tameness' or 'well-behavedness' in finite relational structures, which is useful both algorithmically and model theoretically. For instance, model checking for first-order or monadic second-order formulae becomes more tractable if the input is restricted to finite structures of bounded treewidth. But also decidability issues, in particular satisfiability, can often be linked to a priori bounds on the treewidth of target models – a phenomenon best known, and in its purest form, for logics with the tree model property, e.g., due to bisimulation invariance. As pointed out above, the bounded treewidth model property of logics invariant under guarded bisimulation extends this benefit to richer settings. Moreover, bounded treewidth has featured in recent analogues to classical expressive completeness issues over finite structures. While bounded treewidth certainly is not the only structural restriction that helps to overcome well known obstacles in finite model theory, it seems to occupy a central place in such concerns. We here mainly want to discuss several such results, especially results concerning expressive completeness for fragments of FO, in the light of connections with techniques stemming form the fundamental notion of Gaifman locality.

Bounded treewidth is also at the center of Stephan Kreutzer's chapter [38] in this volume, where the algorithmic impact of bounded treewidth, among other structural criteria, is treated in depth. There the reader will also find a much more detailed account of the connections between bounded treewidth and model checking complexities for first- and monadic second-order logic than what is sketchily hinted at below.

Relational structures of bounded treewidth We have already come across a special form of tree decompositions in section 4.2, cf. Definition 4.3, and now briefly review the general notion of a tree decomposition underlying the definition of treewidth.

Definition 5.1. A *tree decomposition* of the finite relational structure $\mathfrak{A}$ consists of a tree T together with a map $\rho\colon T \to \mathcal{P}(A)$ associating subsets of A with the nodes of T in such a manner that

(i) every relational ground atom of $\mathfrak{A}$ is contained in some $\rho(t)$.

(ii) for all $a \in A$, $\{t \in T\colon a \in \rho(t)\} \subseteq T$ is connected in T.

The width of the tree decomposition (T, ρ) is $\max_{t \in T} |\rho(t)| - 1$.
The *treewidth of* $\mathfrak{A}$, $\mathrm{tw}(\mathfrak{A})$ is the minimal width among all tree decompositions of $\mathfrak{A}$.
$\mathcal{C}_k[\tau] := \{\mathfrak{A}\colon \mathrm{tw}(\mathfrak{A}) \leqslant k\}$ denotes the class of finite τ-structures of treewidth up to k.

Note that (i) is the same as to say that the subsets used in a tree decomposition of $\mathfrak{A}$ must cover the guarded subsets.[16] The correction by -1 in the definition of treewidth is so that trees get treewidth 1 (rather than 2, which is the required patch size).

[16]That they must not themselves be guarded subsets accounts for the difference in comparison with Definition 4.3; a tree decomposition of $\mathfrak{A}$ is a hypergraph decomposition of some hypergraph that may be coarser than the hypergraph $H(\mathfrak{A})$ of guarded subsets.

Model checking complexity For the complexity of the model checking problem for some logic L over the class $\mathcal{C}$, one distinguishes

- *combined complexity*, where both $\varphi \in \mathsf{L}$ and $\mathfrak{A} \in \mathcal{C}$ vary, and the input size is the sum of the input sizes, $|\varphi| + ||\mathfrak{A}||$;[17]
- *data complexity*, where the formula $\varphi \in \mathsf{L}$ is fixed, and the variation is in the structure, with input size measure $||\mathfrak{A}||$;
- *expression complexity*, with fixed $\mathfrak{A}$ and varying $\varphi \in \mathsf{L}$.

The following are some well known cornerstones for the model checking complexity of monadic second-order logic MSO, FO and some fragments of FO considered above:

- MSO model checking over $\mathcal{C}_k$ (treewidth k structures) has linear combined complexity due to a fundamental theorem of Courcelle [11], where "linear" refers to a complexity in $\mathcal{O}(||\mathfrak{A}|| \cdot |\varphi|)$. On the class of all finite graphs, on the other hand, MSO clearly captures graph properties at any level of the polynomial hierarchy (this is w.r.t. data complexity).
- FO-formulae have logarithmic data complexity (i.e., poly-logarithmic in $||\mathfrak{A}||$ or $|A|$, but with syntactic parameters of the formula in the exponent) [14, 32].
- The combined complexity of FO model checking is complete for Pspace (this is even true for formula complexity over the fixed naked two-element structure, by a simple reduction of the Pspace complete satisfiability problem for quantified boolean formulae).
- The combined complexity for model checking FO^k, on the other hand, is complete for Ptime for every $k \geqslant 2$, and even linear for FO^2 as well as for GF, and still Ptime complete even for ML, [45, 23, 7, 17].

Interestingly, measures of *tree-likeness* improve model checking complexities – both on the side of the structure (e.g., model checking over bounded treewidth structures) and on the side of the formula input (e.g., model checking conjunctive queries with templates of bounded treewidth). We just mention some key results with pointers to the literature, and again refer to [38] for a more thorough treatment of some of these.

FO data and combined complexity and local constraints For FO data complexity, Frick and Grohe [15] establish a linear bound over any class $\mathcal{C}$ of structures whose treewidth is *locally bounded*. A class $\mathcal{C}$ of structures has locally bounded treewidth if the treewidth of ℓ-neighbourhoods in structures from $\mathcal{C}$ is uniformly bounded by some function in the radius ℓ. The underlying model checking algorithm is based on a presentation of the formula in Gaifman form. With this, the checking of 'global' structural properties reduces to local evaluation of FO formulae in ℓ-neighbourhoods and a graph theoretic core algorithm that checks for existence of scattered tuples in the Gaifman graph, vertex-coloured according to the local pre-processing. For generalisations and more recent successes of this approach to first-order model checking complexity in

[17] $||\mathfrak{A}||$ stands for the size of a succinct encoding of the relational structure $\mathfrak{A}$. E.g., for graphs $\mathfrak{A}$ in an adjacency list encoding, $||\mathfrak{A}|| \in \mathcal{O}(n^2)$, but it can be sub-quadratic in the number $n = |A|$ of vertices for graphs with few edges. Finer complexity accounts need to be based on a random access model of computation, so that access to the input structure does not distort the real algorithmic content of formula evaluation.

classes tamed by local conditions on graph invariants see Grohe's survey [25] as well as Kreutzer's chapter [38] in this valume, with a view also to the parameterised complexity of the combined model checking problem.

Combined complexity for fragments of FO The combined complexity of conjunctive query evaluation has been studied intensively, with a natural motivation central to database theory and with interesting connections to constraint satisfaction problems. Also in these investigations tree-likeness (in this case of syntactic features of very special FO formulae) plays a major role. *Conjunctive query evaluation* is the model checking of existential positive prenex FO formulae whose quantifier-free core is just a conjunction of relational atoms, $\varphi = \exists \mathbf{x} \bigwedge_i \alpha_i(\mathbf{x}_i)$ with atomic α_i (in subtuples of variables $\mathbf{x}_i$ of $\mathbf{x}$). The link with homomorphism problems and hence with constraint satisfaction (see for instance [35, 37]) is natural and straightforward. The desired assignment to variables $\mathbf{x}$ over the τ-structure $\mathfrak{A}$ *is* a homomorphism from a τ-structure $\mathfrak{X}_\varphi$ induced by the conjuncts α_i on the set of variables $[\mathbf{x}]$ into $\mathfrak{A}$,

$$\beta \colon \mathfrak{X}_\varphi \xrightarrow{\text{hom}} \mathfrak{A}.$$

Note that while the data complexity is poly-logarithmic for each individual (first-order) φ or $\mathfrak{X}$, in general one expects an exponential dependency on the number of variables in φ or on the size of $\mathfrak{X}$.

It turns out that the hypergraph $H(\mathfrak{X}_\varphi)$ holds one key to better bounds on the complexity of the associated homomorphism/query evaluation problems. In fact φ is (equivalent to a formula) in GF if $H(\mathfrak{X}_\varphi)$ is acyclic, in which case model checking becomes linear in $|\varphi|$. Indeed, a tree decomposition of $H(\mathfrak{X}_\varphi)$ yields a translation into GF and hence a reduction to the linear model checking of GF. This generalises to φ with a fixed bound on the treewidth of $\mathfrak{X}_\varphi$, where the model checking can be based on the auxiliary acyclic hypergraph of bounded width extracted from the tree decomposition (instead of $H(\mathfrak{X}_\varphi)$ itself). In these cases, which admit considerable further extensions in terms of weaker notions of bounded widths (e.g., bounded hypertreewidth rather than treewidth [19]), combined model checking remains in Ptime [18, 19].

But also reductions to FO^k can be seen as essential for tractability. For any finite τ-structures $\mathfrak{X}$ and $\mathfrak{A}$, the following are equivalent [12, 37]:

- (i) existence of a homomorphism from $\mathfrak{X}$ to $\mathfrak{A}$, $\mathfrak{X} \xrightarrow{\text{hom}} \mathfrak{A}$;
- (ii) $\mathfrak{A} \models \exists \mathbf{x} \eta_{\mathfrak{X}}$, where $\eta_{\mathfrak{X}}$ is the positive diagram of $\mathfrak{X}$;
- (iii) the *transfer property* $\mathfrak{X} \Rightarrow_{\mathrm{pos}\,\exists^*} \mathfrak{A}$, meaning that every positive existential sentence true in $\mathfrak{X}$ is also true in $\mathfrak{A}$.

For $\mathfrak{X} = \mathfrak{X}_\varphi$, where φ is a conjunctive query, φ is equivalent to $\exists \mathbf{x} \eta_{\mathfrak{X}}$ (cf. (ii)). For $\mathfrak{X} \in \mathcal{C}_k$, this sentence is expressible in positive existential FO^{k+1} [36, 37], so that (iii) above can be replaced by a transfer condition for all positive existential FO^{k+1} rather than FO. In this context, therefore, the Ptime analysis of winning positions in the (positively restricted, one-sided) $(k+1)$-pebble game [34] on $\mathfrak{X}$ versus $\mathfrak{A}$ decides the homomorphism problem.

5.2 Non-classical proofs for (variants of) classical characterisations

With this section we return to expressive completeness issues, related to the existential and the existential positive fragments of FO over classes of finite structures. A first-order sentence $\varphi \in \mathrm{FO}[\tau]$ is *preserved under extensions* if in every substructure relationship $\mathfrak{A} \subseteq \mathfrak{B}$ between τ-structures, $\mathfrak{A} \models \varphi$ implies $\mathfrak{B} \models \varphi$. Similarly, φ is *preserved under homomorphisms* if for every homomorphism $\mathfrak{A} \xrightarrow{\text{hom}} \mathfrak{B}$ between τ-structures, $\mathfrak{A} \models \varphi$ implies $\mathfrak{B} \models \varphi$. As an embedding of a substructure is a special case of homomorphism, preservation under homomorphisms implies preservation under extensions. Clearly, existential sentences are preserved under extensions, while existential positive sentences are even preserved under homomorphisms.

The classical results are the following. We explicitly state the more interesting expressive completeness statements.

Theorem 5.2 (Łos–Tarski). *The existential fragment of first-order logic is expressively complete for first-order properties that are preserved under extensions.*

Theorem 5.3 (Lyndon–Tarski). *The existential positive fragment of first-order logic is expressively complete for first-order properties that are preserved under homomorphisms.*

These are proved classically, e.g. in [10], by means of a compactness argument for the construction of suitable elementary extensions, respectively elementary chain constructions.

Classically, as well as towards possible restrictions of the expressive completeness claim to some class $\mathcal{C}$ other than the class of all τ-structures, both essentially amount to finiteness claims for classes of *minimal models* (within $\mathcal{C}$).

We refer to *substructure minimal* models as generators w.r.t. extensions, and, as generators w.r.t. homomorphisms, also to so-called *cores*. In a class closed under homomorphisms, the natural generators are simultaneously minimal w.r.t. the weak substructure relationship and w.r.t. inverse homomorphisms. We review some standard terminology in this connection.

A *weak substructure* relationship between τ-structures, denoted $\mathfrak{A} \subseteq_w \mathfrak{B}$, requires that $A \subseteq B$ and $R^{\mathfrak{A}} \subseteq R^{\mathfrak{B}}$ for every relation R in τ (rather than $R^{\mathfrak{A}} = R^{\mathfrak{B}} {\restriction} A$ as in the substructure relationship $\mathfrak{A} \subseteq \mathfrak{B}$). A *retraction* is a homomorphism h from some structure $\mathfrak{A}$ onto a weak substructure $\mathfrak{A}_0 \subseteq_w \mathfrak{A}$ such that $h{\restriction}A_0 = \mathrm{id}$. It is worth noting that a retraction $h\colon \mathfrak{A} \xrightarrow{\text{ret}} \mathfrak{A}_0$ is accompanied by a trivial inclusion homomorphism back from $\mathfrak{A}_0$ into $\mathfrak{A}$, since $\mathfrak{A}_0 \subseteq_w \mathfrak{A}$. A structure whose only retraction is the identity is called a *core*. Every finite relational structure $\mathfrak{A}$ possesses a retract onto some core and this core is unique up to isomorphism. It is then straightforward to see that a homomorphism closed class of finite structures is generated by its members that are cores; viz., generated as the class of all weak extensions of these. But the subclass of $\subseteq_w$-minimal members generates the same class.

Definition 5.4. (a) $\mathfrak{A}$ is a *substructure minimal* ($\subseteq$-minimal) model of φ if $\mathfrak{A} \models \varphi$ and $\mathfrak{A}' \not\models \varphi$ for all $\mathfrak{A}' \subsetneq \mathfrak{A}$.

(b) $\mathfrak{A}$ is a *weak-substructure minimal* ($\subseteq_w$-minimal) model of φ if $\mathfrak{A} \models \varphi$ and $\mathfrak{A}' \not\models \varphi$ for all $\mathfrak{A}' \subsetneq_w \mathfrak{A}$.

(c) $\mathfrak{A}$ is a *core* model of φ if $\mathfrak{A} \models \varphi$ and $\mathfrak{A}$ is a core.

Observation 5.5. *Let $\mathcal{C}_0$ be a class of finite τ-structures that is closed under extensions. Then the following are equivalent:*

(i) $\mathcal{C}_0$ is definable (within the class of finite τ-structures) by an existential first-order sentence.

(ii) $\mathcal{C}_0$ has, up to isomorphism, finitely many substructure minimal members.

For the crucial direction, (ii) $\Rightarrow$ (i): if $\mathfrak{A}_1, \ldots, \mathfrak{A}_N$ are the isomorphism types of substructure minimal members in $\mathcal{C}_0$, then $\mathcal{C}_0$ is definable by the disjunction over the existentially quantified algebraic diagrams of the $\mathfrak{A}_i$. For (i) $\Rightarrow$ (ii) it suffices to observe that the size of substructure minimal models of an existential prenex sentence φ is bounded by the number of variables.

The above equivalence persists in restriction to any class $\mathcal{C}$ of τ-structures that is itself closed under substructures (some such extra condition on the surrounding class $\mathcal{C}$ is necessary for (i) $\Rightarrow$ (ii), not for (ii) $\Rightarrow$ (i)).

Similarly one obtains the following, where a disjunction over the existentially quantified *positive* diagrams of $\subseteq_w$-minimal models, which are cores, provides a canonical definition in existential positive FO. We state the equivalence relative to the class of all (finite) τ-structures, but it similarly holds in restriction to any class $\mathcal{C}$ of τ-structures that is closed, e.g., under substructures.

Observation 5.6. *For any class $\mathcal{C}_0$ of (finite) τ-structures that is closed under homomorphisms, the following are equivalent:*

(i) $\mathcal{C}_0$ is definable (within the class of finite τ-structures) by a sentence in existential positive FO.

(ii) $\mathcal{C}_0$ has, up to isomorphism, finitely many $\subseteq_w$-minimal members.

(iii) $\mathcal{C}_0$ has, up to isomorphism, finitely many $\subseteq$-minimal members.

(iv) $\mathcal{C}_0$ has, up to isomorphism, finitely many homomorphism minimal core members.

As we are dealing with finite relational vocabularies τ, a finite bound on the number of isomorphism types of minimal models is equivalent to a bound on the size of minimal models.

It has been known for a long time that the Łos–Tarski theorem (Theorem 5.2) fails in the sense of finite model theory (with counterexamples due to Tait and Gurevich, see e.g. [14]).

The status of the Lyndon–Tarski theorem (Theorem 5.3) in finite model theory, on the other hand, had been an important open problem for quite some time when it was resolved, positively, by Rossman [44].

Beside the overall finite model theory version, however, one may of course investigate the status of these expressive completeness issues in restriction to various classes of (finite) structures of interest. In the following sections we outline a particular criterion of well-behavedness motivated by considerations of Gaifman locality, which has led to interesting results along these lines.

5.2.1 Wideness criteria

The wideness criteria proposed in [4, 3] couple the existence of large scattered subsets to the size of structures. In the context of the minimal model criteria as in Observations 5.5

and 5.6 above they can be used to derive upper bounds on the size of minimal models. Models exceeding a certain size cannot be minimal if their richness in scattered sets allows one to extract smaller models on the basis of a Gaifman representation of the first-order property at hand.

Definition 5.7. A structure is (ℓ, m)-*wide* if its Gaifman graph contains an ℓ-scattered subset of size m.

A class $\mathcal{C}$ of τ-structures is called *wide* if there is a function $N : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, such that, for all ℓ and m and $\mathfrak{A} \in \mathcal{C}$, if $|A| \geqslant N(\ell, m)$, then $\mathfrak{A}$ is (ℓ, m)-wide.

$\mathcal{C}$ is called *almost wide* if, for some fixed k, the analogous condition applies after the removal of a suitable subset of at most k elements from the structures $\mathfrak{A}$ at hand.

A typical example of a wide class is the class of graphs of fixed bounded degree. The class of trees, on the other hand, is not wide (there are arbitrarily large trees of diameter 2), but almost wide: a large tree either has long branches or a node of high degree; removal of a single node of high degree also produces a large scattered set. Similarly, in a tree decomposition of fixed bounded width of a sufficiently large graph or relational structure, a large scattered set becomes available at least after the removal of the elements associated with some high degree node in the decomposition tree. A much more profound analysis is necessary to show almost wideness for every class of graphs that excludes a minor [4].

Proposition 5.8 (Atserias–Dawar–Kolaitis)**.** *The class of treewidth k graphs is almost wide. By extension, $\mathcal{C}_k[\tau]$, the class of τ-structures of treewidth up to k, is almost wide.*

More generally, any class of graphs with excluded minor is almost wide, and by extension any class of τ-structures whose Gaifman graphs avoid some minor.

5.2.2 Expressive completeness for extension preservation

The following summarises key results from [3].

Theorem 5.9 (Atserias–Dawar–Grohe)**.** *The size of $\subseteq$-minimal models of a first-order sentence φ that is preserved under extensions can be bounded over the following classes of finite structures:*

(i) acyclic relational structures (i.e., directed coloured graphs with acyclic Gaifman graphs);

(ii) wide classes $\mathcal{C}$, like any class of graphs of bounded degree.

(iii) $\mathcal{C}_k$, the class of all finite structures of treewidth up to k.

As a consequence, existential FO *is expressively complete for first-order properties preserved under extensions over these classes.*

Interestingly, there are almost wide classes over which existential FO is not expressively complete for first-order properties preserved under extensions. A counterexample over the class of planar graphs is given in [3].

The underlying idea in the proof of the theorem is to choose parameters ℓ, q, m from a Gaifman representation of φ, such that φ is preserved under $\equiv^{\ell}_{q,m}$, and then to isolate a proper substructure $\mathfrak{A}_0 \subsetneq \mathfrak{A}$ that at the same time is $\equiv^{\ell}_{q,m}$ equivalent to some extension $\hat{\mathfrak{A}} \supseteq \mathfrak{A}$, in any large enough model $\mathfrak{A}$. The actual argument in [3] involves a sophisticated finite chain construction.

5.2.3 Expressive completeness for homomorphism preservation

The connection between wideness criteria and bounds on the number (or size) of $\subseteq_w$-minimal models, which is crucial according to Observation 5.6, is provided by the following theorem. It stems from the analysis of the boundedness problem for Datalog programs (least fixpoint recursion over positive existential FO) over finite structures.

Theorem 5.10 (Ajtai–Gurevich)**.** *Let $\mathcal{C}$ be a class of finite τ-structures that is closed under substructures and disjoint unions. If $\varphi \in$ FO is preserved under homomorphisms within $\mathcal{C}$, then there are $\ell, m \in \mathbb{N}$ such that no (ℓ, m)-wide model of φ can be $\subseteq$-minimal.*

The same applies w.r.t. wideness after removal of up to k elements, for fixed k.

Corollary 5.11 (Atserias–Dawar–Kolaitis)**.** *Over any class of finite structures that is almost wide and closed under substructures and disjoint unions, existential positive* FO *is expressively complete for first-order properties preserved under homomorphisms.*

That minimal models cannot be too wide in the sense of Theorem 5.10, comes from a Gaifman representation of φ. We sketch the argument that, for a first-order sentence φ that is preserved under $\equiv^{\ell}_{q,m}$ and under homomorphisms (within $\mathcal{C}$), there are $L, M \in \mathbb{N}$ such that no (L, M)-wide model of φ can be $\subseteq_w$-minimal. More precisely, there are

- M, large enough w.r.t. L, Q, such that within any L-scattered subset of size M in $\mathfrak{A} \models \varphi$ we find some pair of elements $a \neq b$ for which $\mathfrak{A}, a \equiv^{L}_{Q,0} \mathfrak{A}, b$;
- L and Q, large enough w.r.t. ℓ, q, such that $\mathfrak{A}, a \equiv^{L}_{Q,0} \mathfrak{A}, b$ implies the following transfer property for Gaifman rank $(\ell, q, 1)$-assertions:

  $$\mathfrak{A} \ \Rrightarrow^{\ell}_{q,1} \ \mathfrak{B} := \mathfrak{A}{\restriction}(A \setminus \{b\}),$$

  meaning that every sentence of the form $\exists x \chi^{\ell}(x)$ where $\mathrm{qr}(\chi) \leqslant q$ that is true in $\mathfrak{A}$ remains true in $\mathfrak{B}$ ($\mathfrak{A}$ with b removed).[18]

M simply needs to be chosen large w.r.t. the number of quantifier-rank Q types of single elements (in their L-neighbourhood) in order to guarantee the existence of distinct but $\equiv^{L}_{Q,0}$ equivalent nodes by the pigeon-hole principle.

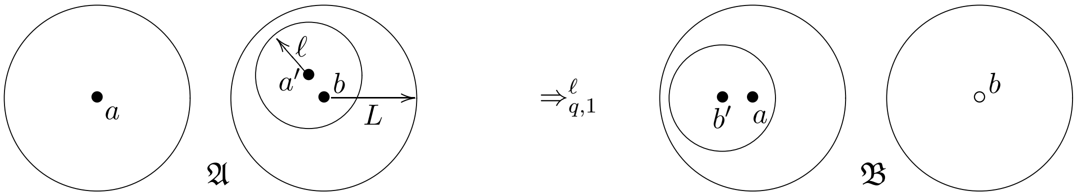
For such a and b, the desired transfer of $\exists x \chi^{\ell}(x)$-assertions follows from $\equiv^{L}_{Q,0}$ equivalence provided $L \geqslant 2\ell$ and Q large enough so that for all $\mathrm{qr}(\chi) \leqslant q$, the assertion

$$\exists x' \big(d(x, x') \leqslant \ell \wedge \chi^{\ell}(x') \big) \qquad (*)$$

is L-local and of quantifier rank $\leqslant Q$. Compare the diagram below for this proof sketch. In the non-trivial case $\mathfrak{A} \models \chi^{\ell}[a']$ for some $a' \in N^{\ell}(b)$, so that after the removal of b, there is no guarantee that still $\mathfrak{B} \models \chi^{\ell}[a']$. Using $\equiv^{L}_{Q,0}$ equivalence between a and b, though, $(*)$ is true of a if it is true at b. Hence there is a corresponding $b' \in N^{\ell}(a)$ such

[18]Note that this is a one-directional transfer rather than an equivalence. E.g., in a graph consisting just of a large cycle, the removal of any single element results in a structure that is inequivalent in the sense of $\equiv^{1}_{1,1}$.

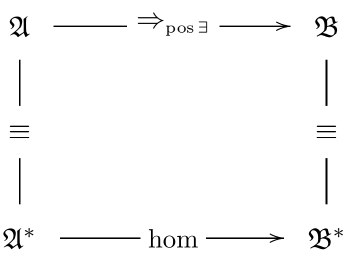
that $\mathfrak{A} \models \chi^\ell[b']$. So $\mathfrak{B} \models \chi^\ell[b']$ follows, since the L-neighbourhood of a is unaffected by the removal of b.

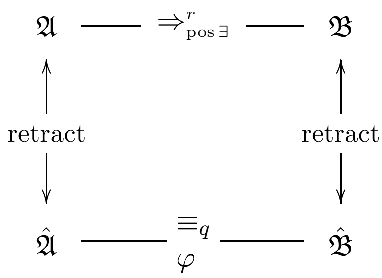


It follows that $\mathfrak{A} \oplus m \cdot \mathfrak{B} \equiv^\ell_{q,m} m \cdot \mathfrak{B}$ (with disjoint sums of m isomorphic copies of $\mathfrak{B}$ plus one copy of $\mathfrak{A}$ on the left-hand side). Therefore, $\mathfrak{B} \models \varphi$ is a smaller model of φ:

$$\mathfrak{A} \xrightarrow{\text{hom}} \mathfrak{A} \oplus m \cdot \mathfrak{B} \equiv^\ell_{q,m} m \cdot \mathfrak{B} \xrightarrow{\text{hom}} \mathfrak{B}.$$

Expressive completeness of the existential positive fragment of FO for homomorphism preservation over the class of all finite relational structures – the finite model theory version of the Lyndon–Tarski Theorem – has been shown by Rossman in [44]. His approach is based on a combinatorial analysis of existential positive types and saturation arguments for these, which can be brought to a sufficient level of closure in a finite chain construction. Leaving aside much of the actual sophistication of the combinatorial analysis, there is one aspect of Rossman's approach that may deserve to be highlighted in connection with the leading themes of this survey. That is the manner in which the new argument is based on explicit model construction (as opposed to an abstract model existence argument), and can be viewed as an upgrading (not of an equivalence, but of a unidirectional transfer relationship) to approximate first-order equivalence, which is orthogonal to the classical argument. This is an interesting parallel with the observations in section 3.2.1. While a traditional proof of the Lyndon-Tarski Theorem can be based on the upgrading indicated in the left-hand diagram, Rossman's proof amounts to the upgrading indicated in the right-hand diagram. In the traditional picture, transfer w.r.t. the full existential positive fragment of FO is upgraded, in a classical saturation argument based on compactness, to yield a homomorphism between structures that are elementarily equivalent to the original ones. In the 'explicit' construction of Rossman's, on the other hand, a specific finite level of transfer (existential positive formulae of quantifier rank up to r) is upgraded to a specific finite level of first-order equivalence that preserves the given sentence φ.





Moreover, Rossman's proof has a classical variant, in which the chain construction is extended to an infinite limit, that yields a completely new, alternative proof of the classical Lyndon–Tarski result with added value. In fact, Rossman shows that existential positive FO is expressively complete for first-order sentences preserved under homomorphisms, level-by-level w.r.t. quantifier-rank. In the classical model theory version of his proof, Rossman realises the above upgrading for $r = q$, while in the finite model theory version there is no elementary bound on r in terms of q.

6 Concluding remarks

The focus on a model theory of well-behaved classes of (finite) structures – adapted to specific application areas, or to the study of specific logics, or to other specific model theoretic themes – offers promising perspectives for the development and ramification of finite model theory. Finiteness as the only constraint, which often entails 'negative' results, may not be the best choice for many reasons.

It can be that the class of all finite structure is still not a good match for the natural domain of reasoning for certain application areas; some model theoretic answers – 'positive' or 'negative' – may still be 'too easy' over the class of all finite structures. In modal reasoning, for instance, rootedness or connectivity constraints are arguably essential in the intuitive modelling. More generally, the 'generic finite structure that we mean' may well have more specific structural properties than an 'arbitrary finite structure.'

It can also be that the class of all finite structures is too liberal a setting for structural insights into certain issues. Definability and expressive completeness results, for instance, that fail over the class of all finite structures may not just be recovered but also clarified overall through a better understanding of the structural conditions that support them. In this sense there is not just finite model theory, but there may be many adequate domains of structures for individual issues.

I think both aspects are important from the modelling point of view (i.e., in relation to applications), also clearly from an algorithmic point of view, but also from the point of view of classical issues in model theory. Sophisticated adaptations of classical techniques, like the analysis of types and the use of chain constructions in Rossman's result, enrich finite model theory but also cast fresh light on long-standing classical results. In this context the constructive aspect of explicit model constructions or model transformations – in contrast with smooth abstract existence proofs in classical model theory – is an important methodological contribution.

It seems that the modularity in game-oriented arguments and model constructions, as illustrated by the power of an analysis in terms of Gaifman locality, has had comparatively little impact on traditional classical model theory. The great potential of another aspect of modularity, viz. decomposition techniques, has apparently been realised more fully. The combination of such aspects may lead to a better model theoretic view of more complex hierarchical decompositions in particular for finite structures; and there may be more flavours of structural regularity, smoothness or tameness in finite structures to be discovered.

References

[1] N. Alon, *Tools from higher algebra*, in Handbook of Combinatorics, R. Graham et al., ed., vol. II, North-Holland, 1995, pp. 1749–1783.

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

[3] A. Atserias, A. Dawar, and M. Grohe, *Preservation under extensions on well-behaved finite structures*, in Proceedings of 32nd International Colloquium on Automata, Languages and Programming ICALP'05, vol. 3580 of LNCS, 2005, pp. 1437–1449.

[4] A. Atserias, A. Dawar, and P. Kolaitis, *On preservation under homomorphisms and unions of conjunctive queries*, Journal of the ACM, 53 (2006), pp. 208–237.

[5] C. Beeri, R. Fagin, D. Maier, and M. Yannakakis, *On the desirability of acyclic database schemes*, Journal of the ACM, 30 (1983), pp. 497–513.

[6] C. Berge, *Graphs and Hypergraphs*, North-Holland, 1973.

[7] D. Berwanger and E. Grädel, *Games and model checking for guarded logics*, in Proceedings of the 8th International Conference on Logic for Programming and Automated Reasoning LPAR'01, vol. 2250 of LNCS, 2001, pp. 70–84.

[8] P. Blackburn, M. de Rijke, and Y. Venema, *Modal Logic*, Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, 2001.

[9] P. Blackburn, J. van Benthem, and F. Wolter, eds., *Handbook of Modal Logic*, Elsevier, 2007.

[10] C. C. Chang and H. J. Keisler, *Model Theory*, North-Holland, 1990.

[11] B. Courcelle, *Graph rewriting: An algebraic and logic approach*, in Handbok of Theoretical Computer Science, volume B, J. van Leeuwen, ed., Elsevier, 1990, pp. 194–242.

[12] V. Dalmau, P. Kolaitis, and M. Vardi, *Constraint satisfaction, bounded treewidth, and finite-varaible logics*, in Proceedings of 8th International Conference on Constraint Programming, vol. 2470 of LNCS, 2002, pp. 310–326.

[13] A. Dawar and M. Otto, *Modal characterisation theorems over special classes of frames.* Extended and corrected journal version of LICS'05 paper, revised, 2008.

[14] H.-D. Ebbinghaus and J. Flum, *Finite Model Theory*, Springer, 2nd ed., 1999.

[15] M. Frick and M. Grohe, *Deciding first-order properties of locally tree-decomposable structures*, Journal of the ACM, 48 (2001), pp. 1184–1206.

[16] V. Goranko and M. Otto, *Model theory of modal logic*, in Handbook of Modal Logic, P. Blackburn, J. van Benthem, and F. Wolter, eds., Elsevier, 2007, pp. 249–329.

[17] G. Gottlob, E. Grädel, and H. Veith, *Datalog LITE: A deductive query language with linear time model checking*, ACM Transactions on Computational Logic, 3 (2002), pp. 1–35.

[18] G. Gottlob, N. Leone, and F. Scarcello, *The complexity of acyclic conjunctive queries*, Journal of the ACM, 43 (2001), pp. 431–498.

[19] ——, *Hypertree decompositions and tractable queries*, Journal of Computer and System Sciences, 64 (2002), pp. 579–627.

[20] E. Grädel, *On the restraining power of guards*, Journal of Symbolic Logic, 64 (1999), pp. 1719–1742.

[21] ——, *Finite model theory and descriptive complexity*, in Finite Model Theory and Its Applications, Springer, 2007, pp. 125–230.

[22] E. Grädel, C. Hirsch, and M. Otto, *Back and forth between guarded and modal logics*, ACM Transactions on Computational Logics, 3 (2002), pp. 418 – 463.

[23] E. Grädel and M. Otto, *On logics with two variables*, Theoretical Computer Science, 224 (1999), pp. 73–113.

[24] E. Grädel and I. Walukiewicz, *Guarded fixed point logic*, in Proceedings of 14th Annual IEEE Symposium on Logic in Computer Science LICS'99, 1999, pp. 45–54.

[25] M. Grohe, *Logic, graphs, and algorithms*, in Logic and Automata, History and Perspectives, J. Flum, E. Grädel, and T. Wilke, eds., Amsterdam University Press, 2008, pp. 357–422.

[26] B. Herwig, *Extending partial isomorphisms on finite structures*, Combinatorica, 15 (1995), pp. 365–371.

[27] ——, *Extending partial isomorphisms for the small index property of many omega-categorical structures*, Israel Journal of Mathematics, 107 (1998), pp. 93–124.

[28] B. Herwig and D. Lascar, *Extending partial isomorphisms and the profinite topology on free groups*, Transactions of the AMS, 352 (2000), pp. 1985–2021.

[29] W. Hodges, *Model Theory*, Cambridge University Press, 1993.

[30] I. Hodkinson and M. Otto, *Finite conformal hypergraph covers and Gaifman cliques in finite structures*, Bulletin of Symbolic Logic, 9 (2003), pp. 387–405.

[31] E. Hoogland, M. Marx, and M. Otto, *Beth definability for the guarded fragment*, in JFAK – Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday, J. Gebrandy, M. Marx, M. de Rijke, and Y. Venema, eds., Amsterdam University Press, 1999. CD-ROM.

[32] N. Immerman, *Decsriptive Complexity*, Graduate Texts in Computer Science, Springer, 1998.

[33] D. Janin and I. Walukiewicz, *On the expressive completeness of the propositional mu-calculus with respect to monadic second order logic*, in Proceedings of 7th International Conference on Concurrency Theory CONCUR'96, vol. 1119 of LNCS, 1996, pp. 263–277.

[34] P. Kolaitis, *On the expressive power of logics on finite models*, in Finite Model Theory and Its Applications, Springer, 2007, pp. 27–123.

[35] P. Kolaitis and M. Vardi, *Conjunctive-query containment and constraint satisfaction*, Journal of Computer and System Sciences, 61 (2000), pp. 302–332.

[36] ——, *A game-theoretic approach to constraint satisfaction*, in Proceedings of 0f 17th Conference on Artificial Intelligence AAAI'00, 2000, pp. 175–181.

[37] ——, *A logical approach to constraint satisfaction*, in Finite Model Theory and Its Applications, Springer, 2007, pp. 339–370.

[38] S. Kreutzer, *Algorithmic meta-theorems*, in Finite and Algorithmic Model Theory, J. Esparza et al., ed., CUP, 2008. (this volume).

[39] K. Lorenz, *Dialogspiele als semantische Grundlage von Logikkalkülen*, Archiv für Mathematische Logik und Grundlagenforschung, 11 (1968), pp. 32–55 and 73–100.

[40] P. Lorenzen, *Ein dialogisches Konstruktivitätskriterium*, in Infinitistic Methods, Proceedings of the Symposium on Foundations of Mathematics, Warsaw 1959, Oxford University Press, 1961, pp. 193–200.

[41] M. Otto, *Modal and guarded characterisation theorems over finite transition systems*, Annals of Pure and Applied Logic, 130 (2004), pp. 173–205.

[42] M. Rabin, *Decidability of second order theories and automata on infinite trees*, Transactions of the AMS, 141 (1969), pp. 1–35.

[43] E. Rosen, *Modal logic over finite structures*, Journal of Logic, Language and Information, 6 (1997), pp. 427–439.

[44] B. Rossman, *Homomorphism preservation theorems*, Journal of the ACM, 55 (2008).

[45] M. Vardi, *On the complexity of bounded-variable queries*, in Proceedings of 14th Annual ACM Symposium on Principles of Database Systems PODS'95, 1995, pp. 266–276.

[46] ——, *Why is modal logic so robustly decidable?*, in Descriptive Complexity and Finite Models, N. Immerman and P. Kolaitis, eds., vol. 31 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science, AMS, 1997, pp. 149–184.

[47] ——, *Games as an algorithmic construct.* Tutorial, Isaac Newton Institute workshop on Games and Verification (Logic and Algorithms programme), 2006.

[48] S. Weinstein, *Unifying themes in finite model theory*, in Finite Model Theory and Its Applications, Springer, 2007, pp. 1–25.