Modal and Guarded Characterisation Theorems over Finite Transition Systems

Martin Otto *
University of Wales Swansea
Swansea SA2 8PP, United Kingdom
m.otto@swan.ac.uk

Abstract

Characterisation theorems for modal and guarded fragments of first-order logic are explored over finite transition systems. We show that the classical characterisations in terms of semantic invariance under the appropriate forms of bisimulation equivalence can be recovered in finite model theory. The new, more constructive proofs naturally extend to alternative proofs of the classical variants. The finite model theory version of van Benthem's characterisation of basic modal logic is due to E. Rosen. That proof is simplified and the result slightly strengthened in terms of quantitative bounds. The main theme, however, is a uniform treatment that extends to incorporate universal and inverse modalities and guarded quantification over transition systems. Technically, the present treatment exploits first-order locality in the context of a new finitary construction of locally acyclic bisimilar covers. These serve as graded finite analogues of tree unravellings, giving local control over first-order logic in finite bisimilar companion structures.

1 Introduction

We are concerned with model theoretic characterisation theorems, i.e., correspondences between semantics and syntax of the following kind.

Theorem 1 (van Benthem). *Any first-order formula $\varphi(x)$ that is invariant under bisimulation is equivalent to a formula of basic modal logic, and vice versa.*

One may read this correspondence both as a semantic characterisation of modal logic as a fragment of first-order logic, or as a provision of effective syntax for the bisimulation invariant first-order properties. Regarding the latter, note that bisimulation invariance, like other non-trivial semantic conditions, is an undecidable property of first-order formulae.

For a syntactic characterisation of the bisimulation invariant first-order properties of *finite* transition systems, one would look to a finite model theory version of van Benthem's theorem. In fact, the same characterisation is valid in the sense of finite model theory.

Theorem 2 (Rosen). *Any first-order formula $\varphi(x)$ that is invariant under bisimulation over finite structures is equivalent over finite structures to a formula of basic modal logic, and vice versa.*

The finite model theory version is by no means a direct consequence of the classical version (and the finite model property for modal logic does not help here). For a simple example of a first-order formula that is bisimulation invariant over finite, but not over infinite structures, take any first-order $\psi(x)$ with infinite but without finite models (an infinity axiom). Any such ψ is trivially bisimulation invariant over finite structures, but cannot be so over infinite structures. Indeed if it were, it would, by Theorem 1, be equivalent to a modal formula, whence it would also have to have finite models. Moreover, the classical proof (involving compactness and typically the use of infinite ω-saturated models) does not translate into a finite model theory proof. Indeed, other closely related classical characterisation theorems, like that of the two-variable fragment of first-order logic in terms of two-pebble game invariance [2], are known to be false in the sense of finite model theory.

The van Benthem-Rosen link seems to be the only established positive example along these lines to date. And while the classical proof of van Benthem's theorem tells us nothing about the finite model theory version, the rather more constructive argument given by Rosen equally applies to the classical version, thus providing a new proof there as well.

Apart from re-proving Rosen's theorem, we shall extend corresponding techniques to obtain similar new characterisations of closely related but more expressive modal logics and also of the guarded fragment of first-order logic, over

* research partially supported by EPSRC grant GR/R11896/01

finite transition systems. These correspond to two natural strengthenings of ordinary (modal) bisimulation: *global two-way bisimulations* and *guarded bisimulations*. [For the precise definitions of the relevant technical concepts in the following discussion we refer to the main part of the paper.] *Two-way* bisimulations require the typical back-and-forth correspondence not only forward along edges but also backwards; in temporal terms they account for past as well as for future behaviour. They are therefore the natural choice in all considerations involving past tense operators, but also implicit whenever symmetry assumptions about the underlying binary relations are made (cf. bi-directional frames). *Global* bisimulations are used to overcome the restriction to just the local neighbourhood of the distinguished node, and are a more natural choice for applications that require a uniform global view – which otherwise can only be modelled in frames with a designated global accessibility relation.

These natural strengthenings of ordinary modal bisimulation are well understood and their association with the extensions of basic modal logic by inverse and universal modalities, respectively, is exactly as expected. On the other hand, global two-way bisimulation equivalence is much more discerning than ordinary bisimulation equivalence. It does therefore make much stronger demands on compatible finitary model constructions, as are at the root of the following new characterisation theorem.

Theorem 3. *Any first-order formula $\varphi(x)$ that is invariant under global two-way bisimulation over finite structures is equivalent over finite structures to a formula of modal logic with inverse and universal modalities, and vice versa.*

Guarded bisimulations between transition systems may be seen as a further strengthening of global two-way bisimulations. In general, however, they have much looser ties with ordinary bisimulation and go much further in lifting the concept of bisimulation equivalence from the modal world of transition systems to arbitrary relational structures. Over transition systems guarded bisimulation is sufficient to encompass not only inverse modalities but also positive boolean combinations of different modalities (and equality). The *guarded fragment* GF of first-order logic, introduced by Andréka, van Benthem and Németi in [1] as a powerful yet tractable generalisation of modal logic, captures many good model theoretic and algorithmic features of modal logics in a much more general setting; see also [7]. Its characterisation as a fragment of first-order logic, in the classical context, is analogous to that of modal logic in Theorem 1.

Theorem 4 (Andréka, van Benthem, Németi). *Any first-order formula $\varphi(\bar{x})$ that is invariant under guarded bisimulation is equivalent to a formula of* GF, *and vice versa.*

In restriction to just transition systems (or relational structures of width 2) our techniques yield a finite model theory version of this characterisation as well.

Theorem 5. *Any first-order formula $\varphi(x)$ in a purely relational vocabulary of width* 2 *that is invariant under guarded bisimulation over finite structures is equivalent over finite structures to a formula of the guarded fragment* GF *of first-order logic, and vice versa.*

Note that Theorem 5 is only a partial analogue of the more general characterisation result in the classical context, which does not put any restriction on the nature of the relational vocabulary. It remains an interesting open problem whether Theorem 4 translates into finite model theory in full generality.

The proofs of both our new characterisation theorems, Theorems 3 and 5 – which are stated as theorems of finite model theory – go through to provide alternative, more constructive proofs of the same characterisations in the classical context. Another interesting methodological point – in our proofs as well as in Rosen's – is the shift from classical first-order model construction techniques to combinatorial transformations of structures. This is also where our bisimilar covers come in. In a sense the corresponding proof strategies are really orthogonal with respect to the two criteria involved: first-order definability vs. bisimulation invariance.

The classical proofs rest on first-order model constructions that upgrade finite approximations of bisimulation to full bisimulation while preserving first-order theories; the proofs presented here use bisimilar covers to preserve full bisimulation equivalence while upgrading control over the (local) first-order theories.

Sections 1.1 and 1.2 provide basic definitions, which are standard apart from the important notion of bisimilar and guarded covers in Definition 9. Section 2 explores Gaifman locality in the context of bisimulation invariant formulae, with a preliminary application to a new proof of Rosen's result. Sections 3 and 4 contain the technical core of the paper, with a construction of locally acyclic covers and its application to the new characterisation theorems, respectively.

1.1 Basic definitions

Structures We look at purely relational structures $\mathfrak{A} = (A, \bar{P}, \bar{R})$ with unary predicates P and binary predicates R. Often one element $a \in A$ is a distinguished parameter, indicated as in $\mathfrak{A}, a$. It is a matter of context and preference whether we want to think of τ-structures as transition systems, as Kripke structures, or as edge and vertex coloured directed graphs. We think of a finite, relational vocabulary of width 2 as fixed throughout, $\tau = \{P, \dots, R, \dots\}$. Where necessary we let $\tau = \tau^{(1)} \dot{\cup} \tau^{(2)}$ to isolate the unary predicates P in $\tau^{(1)}$ and the binary predicates R in $\tau^{(2)}$.

Although we are mostly interested in finite model theory, all our considerations equally apply to infinite structures.

We therefore adopt the convention to mention finiteness explicitly where it matters. For common logical notions, like semantic implication, $\varphi \models \psi$, or logical equivalence (semantic bi-implication), $\varphi \equiv \psi$, we use modified notation like $\varphi \models_{\text{fin}} \psi$ or $\varphi \equiv_{\text{fin}} \psi$ to indicate the restriction to just finite models.

Basic modal logic For general background on propositional modal logic, here denoted ML, we refer to the comprehensive textbook [5]. The formulae of $\mathrm{ML}[\tau]$ are generated as follows:

- for each $P \in \tau^{(1)}$, P is an atomic formula of $\mathrm{ML}[\tau]$.
- if $\varphi \in \mathrm{ML}[\tau]$ then so is $\neg\varphi$.
- if $\varphi_1, \varphi_2 \in \mathrm{ML}[\tau]$, then so are $\varphi_1 \wedge \varphi_2$ and $\varphi_1 \vee \varphi_2$.
- if $R \in \tau$ is binary, $\varphi \in \mathrm{ML}[\tau]$, then the following are formulae of $\mathrm{ML}[\tau]$: $[R]\varphi$ and $\langle R\rangle\varphi$.

Instead of presenting the semantics in standard terms, we right away pass to a translation into first-order logic, which induces the usual semantics. Using variable symbols x and y (distinct from x) as meta-variables, we associate formulae $\varphi(x)$ and $\varphi(y)$ in $\mathrm{FO}[\tau]$ with every $\varphi \in \mathrm{ML}[\tau]$ according to the following. For atomic $\varphi = P$ we let $\varphi(x) = Px$ and $\varphi(y) = Py$; the boolean connectives carry over; and for $[R]\varphi$ or $\langle R\rangle\varphi$, respectively, we put

$$\begin{aligned} \big([R]\varphi\big)(x) &:= \forall y\big(Rxy \to \varphi(y)\big), \\ \big(\langle R\rangle\varphi\big)(x) &:= \exists y\big(Rxy \wedge \varphi(y)\big), \end{aligned}$$

and similarly with x and y exchanged. For the sake of a uniform treatment we identify $\mathrm{ML}[\tau]$ with its image under this standard translation, and regard the classical syntax as shorthand for the corresponding expansion into $\mathrm{FO}[\tau]$.

Modal logic with inverse and universal modalities The following modalities provide natural extensions of basic modal logic.

Universal modality: universal (and, by boolean duality, existential) quantification, corresponding to the modal operator [U] where U is the universal relation $\mathbf{U}^{\mathfrak{A}} = A \times A$.

Inverse modalities: modal quantification also w.r.t. to the inverses R^- of the binary relations R, corresponding to operators $[R^-]$ and $\langle R^-\rangle$ for $R^- = \{(a, a') \colon (a', a) \in R\}$.

We denote by $\mathrm{ML}^{-\forall}$ the simultaneous closure of ML with these operators, also allowing the new sentential formulae as constituents in the scope of boolean operators and modal quantification.

The guarded fragment The guarded fragment GF extends the modal quantification pattern to a more general form of relativised first-order quantification. We present the syntax in the general format of arbitrary relational vocabularies, but keep in mind that we shall only deal with GF in the setting of width 2 vocabularies where the similarity with ML is closer. See [1, 4, 7, 8, 10] on GF and related context. The formulae of $\mathrm{GF}[\tau]$ are generated as follows:

- all atomic τ-formulae are formulae of $\mathrm{GF}[\tau]$.
- $\mathrm{GF}[\tau]$ is closed under $\neg$, $\wedge$ and $\vee$.
- if $\varphi(\bar{x}, \bar{y})$ is a formula of $\mathrm{GF}[\tau]$ and if $\alpha(\bar{x}, \bar{y})$ is a τ-atom (also allowing equality) such that $\mathrm{free}(\varphi) \subseteq \mathrm{var}(\alpha)$, then the following are formulae of $\mathrm{GF}[\tau]$:

$$\begin{aligned} \big(\forall\bar{y}.\alpha\big)\varphi(\bar{x}, \bar{y}) &:= \forall\bar{y}\big(\alpha(\bar{x}, \bar{y}) \to \varphi(\bar{x}, \bar{y})\big), \\ \big(\exists\bar{y}.\alpha\big)\varphi(\bar{x}, \bar{y}) &:= \exists\bar{y}\big(\alpha(\bar{x}, \bar{y}) \wedge \varphi(\bar{x}, \bar{y})\big). \end{aligned}$$

The atom α in the last clause is called the *guard* of the (universal or existential) quantification.

The *nesting depth* $\mathrm{d}(\varphi)$ is defined inductively for $\varphi \in \mathrm{GF}$ as follows. $\mathrm{d}(\varphi) = 0$ for atomic φ; $\mathrm{d}(\neg\varphi) = \mathrm{d}(\varphi)$; $\mathrm{d}(\varphi_1 \wedge \varphi_2) = \max(\mathrm{d}(\varphi_1), \mathrm{d}(\varphi_2))$; and $\mathrm{d}((\forall\bar{y}.\alpha)\varphi) = \mathrm{d}((\exists\bar{y}.\alpha)\varphi) = \mathrm{d}(\varphi) + 1$.

For ML or $\mathrm{ML}^{-\forall}$ (viewed as fragments of GF), this guarded nesting depth is just the modal nesting depth or quantifier rank; for GF itself, however, it can be more indicative than ordinary quantifier rank.

Clearly $\mathrm{ML} \subseteq \mathrm{GF}$, and even $\mathrm{ML}^{-\forall} \subseteq \mathrm{GF}$. These inclusions are strict; e.g., $\exists x Rxx$ or $\forall y(Rxy \to (Ryx \vee Py))$ cannot be expressed in $\mathrm{ML}^{-\forall}$.

1.2 Bisimulations

Modal bisimulations Bisimulation equivalences capture notions of behavioural equivalence between transition systems. They can equivalently be presented either in terms of games or in terms of back-and-forth systems. Many variations of the basic notion of plain bisimulation equivalence have been considered. We here deal with plain bisimulation equivalence and the variation involving unrestricted moves to fresh start states (cf. universal modality) and backward transitions (cf. inverse modalities).

Definition 6. Let $\mathfrak{A}$ and $\mathfrak{B}$ be τ-structures, $Z \subseteq A \times B$, $Z \neq \emptyset$. Z is a *bisimulation* between $\mathfrak{A}$ and $\mathfrak{B}$ if for all $(a, b) \in Z$ the following are satisfied:

(i) for all $P \in \tau^{(1)}$: $\mathfrak{A}, a \models P \Leftrightarrow \mathfrak{B}, b \models P$.

(ii) back-and-forth conditions w.r.t. every $R \in \tau^{(2)}$:
forth: for all a' such that $(a, a') \in R$ there is some b' such that $(b, b') \in R$ and $(a', b') \in Z$.
back: for all b' such that $(b, b') \in R$ there is some a' such that $(a, a') \in R$ and $(a', b') \in Z$.

A bisimulation Z between $\mathfrak{A}$ and $\mathfrak{B}$ is a *global bisimulation* if in addition $\pi_1(Z) = A$ and $\pi_2(Z) = B$. A (global) bisimulation between $\mathfrak{A}$ and $\mathfrak{B}$ is a (global) *two-way bisimulation* if it also satisfies back-and-forth conditions w.r.t. the inverses R^- of the R.

We write $Z\colon \mathfrak{A} \sim \mathfrak{B}$ to indicate that Z is a bisimulation between $\mathfrak{A}$ and $\mathfrak{B}$; and $Z\colon \mathfrak{A} \approx \mathfrak{B}$ to indicate that Z is a global two-way bisimulation. Two structures are bisimilar, $\mathfrak{A} \sim \mathfrak{B}$ or $\mathfrak{A} \approx \mathfrak{B}$, if there is a corresponding bisimulation. Two structures with distinguished nodes are bisimilar, $\mathfrak{A}, a \sim \mathfrak{B}, b$ or $\mathfrak{A}, a \approx \mathfrak{B}, b$, if there is a corresponding bisimulation with $(a, b) \in Z$.

Guarded bisimulations Guarded bisimulations are the adequate counterparts to bisimulations in the context of GF.

Definition 7. Let $\mathfrak{A}$ be a τ-structure. A subset $s \subseteq A$ is *guarded* if s is a singleton set or if $s = \{a_1, \dots, a_k\}$ for some tuple $(a_1, \dots, a_k) \in R$ for some relation $R \in \tau$. A tuple $\bar{a}$ over $\mathfrak{A}$ is *guarded* if its components are elements of some common guarded subset.

Definition 8. Let $\mathfrak{A}$ and $\mathfrak{B}$ be τ-structures, $Z \subseteq \mathrm{Part}(\mathfrak{A}, \mathfrak{B})$ a non-empty set of partial isomorphisms between $\mathfrak{A}$ and $\mathfrak{B}$. Z is a *guarded bisimulation* between $\mathfrak{A}$ and $\mathfrak{B}$ if

(i) for every $p \in Z$, the domain and image of p are guarded subsets of $\mathfrak{A}$ and $\mathfrak{B}$, respectively.

(ii) Z satisfies the following back-and-forth conditions w.r.t. guarded subsets:
forth: for all $p \in Z$ with $\mathrm{dom}(p) = s$ and all guarded subsets s' of $\mathfrak{A}$, there is some $p' \in Z$ with $\mathrm{dom}(p') = s'$ such that p and p' agree on their common domain.
back: for all $p \in Z$ with $\mathrm{im}(p) = t$ and all guarded subsets t' of $\mathfrak{B}$, there is some $p' \in Z$ with $\mathrm{im}(p') = t'$ such that p^{-1} and p'^{-1} agree on their common domain.

We write $Z\colon \mathfrak{A} \sim_{\mathrm{g}} \mathfrak{B}$, and $Z\colon \mathfrak{A}, \bar{a} \sim_{\mathrm{g}} \mathfrak{B}, \bar{b}$ to indicate that $p\colon \bar{a} \mapsto \bar{b}$ for some $p \in Z$. Note that $\sim_{\mathrm{g}}$ refines $\approx$.

Bisimilar covers A special and very natural kind of bisimulations – familiar from companion structures obtained as 'unravellings' – are those induced by homomorphisms. We present both the modal and the guarded versions in the following definition.

Definition 9. (a) A homomorphism $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$ is a *bisimilar cover* of $\mathfrak{A}$ by $\hat{\mathfrak{A}}$ if $Z_\pi = \big\{(\hat{a}, a)\colon a = \pi(\hat{a})\big\}$ is a global two-way bisimulation between $\hat{\mathfrak{A}}$ and $\mathfrak{A}$.

(b) A homomorphism $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$ is a *guarded cover* of $\mathfrak{A}$ by $\hat{\mathfrak{A}}$ if $Z_\pi = \big\{\pi\restriction\hat{s}\colon \hat{s} \subseteq \hat{A} \text{ guarded in } \hat{\mathfrak{A}}\big\}$ is a guarded bisimulation between $\hat{\mathfrak{A}}$ and $\mathfrak{A}$.

Bisimulation games It is well understood how the above notions of modal and guarded bisimulation capture corresponding Ehrenfeucht-Fraïssé games adapted to the specific quantification patterns of ML, $\mathrm{ML}^{-\forall}$ and GF. In the ordinary *modal bisimulation game* players move single markers (one in each structure) forward along relational edges; the two-way bisimulation game also has backward moves along edges; for the global bisimulation game, the first move can be made to any position in the respective structure; and for the guarded bisimulation game, groups of markers are relocated in such a way that the tuples of marked elements form guarded sets. In any of these games it is the second player's task to copy the first player's moves in the opposite structure in such a way that the correspondence between the marked configurations is compatible with the basic propositions or the local isomorphism type, respectively. In each case, a bisimulation between the two structures is nothing but a formalisation of a winning strategy for the second player, which allows her continue her play through an infinite sequence of rounds.

Finite approximations Besides strategies in the infinite games one naturally considers strategies in corresponding games with a fixed finite number of ℓ rounds. These induce finite approximations to full bisimulation equivalence, at successively refined finite levels $\ell \in \mathbb{N}$. We denote these approximations by superscripts as in $\sim^\ell$, $\approx^\ell$ and $\sim_{\mathrm{g}}^\ell$. In complete analogy with classical Ehrenfeucht-Fraïssé theory, the respective equivalence relations have finite index for each ℓ (over the class of all τ-structures, for fixed finite relational τ). Furthermore, each of these equivalence classes is definable by a formula in the corresponding fragment of first-order logic (ML, $\mathrm{ML}^{-\forall}$, or GF) at nesting depth ℓ. As a consequence one has straightforward level-by-level correspondences between invariance under ℓ-bisimulations (of the respective kind) and formulae of nesting depth ℓ – both classically and in the sense of finite model theory. These ℓ-approximations, however, do not prove the full characterisation theorems. Over infinite structures, full bisimulation equivalence is stronger than simultaneous ℓ-bisimulation for all ℓ; and even over finite (or finitely branching) structures, where $\sim\, = \bigcap_{\ell \in \mathbb{N}} \sim^\ell$, there is no direct link. Rather, the actual connection is captured in the following observation, whose straightforward proof is here omitted.

Observation 10. *Let $\leftrightarrows$ stand for any bisimulation-like equivalence with finite approximations $\leftrightarrows^\ell$ of finite index. Let the logic $\mathcal{L} = \bigcup_\ell \mathcal{L}^\ell$ be closed under disjunction and such that each formula of $\mathcal{L}^\ell$ is invariant under $\leftrightarrows^\ell$ and each $\leftrightarrows^\ell$-class definable in $\mathcal{L}^\ell$. Then the following are equivalent, both in the sense of classical model theory and of finite model theory:*

(i) *Every first-order formula that is $\leftrightarrows$-invariant is invariant under $\leftrightarrows^\ell$ for some ℓ.*

(ii) *Every first-order formula that is $\leftrightarrows$-invariant is equivalent to some formula in $\mathcal{L}$.*

One major step towards establishing (i) in the cases considered here relies on Gaifman locality for first-order logic.

2 Gaifman locality and invariance

Recall that the Gaifman graph $G(\mathfrak{A})$ of a relational structure $\mathfrak{A} = (A, \dots)$ is the undirected graph over A with edges (a, a') for all non-degenerate guarded pairs (a, a'). *Gaifman distance* d is induced by ordinary graph distance in $G(\mathfrak{A})$. The *ℓ-neighbourhood* of $a \in A$ is the subset $N^\ell(a) = \{a' \in A \colon d(a, a') \leqslant \ell\}$. An *$\ell$-scattered* subset of A is a set of elements with disjoint ℓ-neighbourhoods.

Definition 11. (i) A formula $\psi(x)$ is *ℓ-local* if it is equivalent to its relativisation to $N^\ell(x)$.

(ii) A *basic ℓ-local sentence* is one that asserts the existence of an ℓ-scattered set of m elements x all of which satisfy the same ℓ-local formula $\psi(x)$.

(iii) A *simple ℓ-local sentence* is one that asserts the existence of an element x satisfying some ℓ-local formula $\psi(x)$.

Gaifman distance d is first-order, in the sense that for every ℓ there is a first-order formula $\eta_\ell(x, y)$ which says that $d(x, y) \leqslant \ell$. The above semantic criteria therefore have syntactic counterparts, where, e.g., ℓ-local formulae $\varphi(x)$ are equivalent to formulae whose quantifiers are explicitly relativised to the ℓ-neighbourhood of x. The following highlights the precise sense in which FO can only express local properties; compare also [6].

Theorem 12 (Gaifman). *Any first-order formula $\varphi(x)$ is equivalent to one that is a boolean combination of local formulae and basic local sentences.*

We write $\mathfrak{A} + \mathfrak{B}$ for the disjoint sum of two τ-structures $\mathfrak{A}$ and $\mathfrak{B}$. We also write $q \cdot \mathfrak{A}$ for the q-fold disjoint sum of $\mathfrak{A}$ with (isomorphic copies of) itself.

Definition 13. A τ-formula φ is

(i) *invariant under disjoint sums* (meaning: with arbitrary other τ-structures) if for all $\mathfrak{A}, \bar{a}$ and $\mathfrak{B}$:
$\mathfrak{A}, \bar{a} \models \varphi \;\Leftrightarrow\; \mathfrak{A} + \mathfrak{B}, \bar{a} \models \varphi$.

(ii) *invariant under disjoint copies* (meaning: of the same τ-structure) if for all $\mathfrak{A}, \bar{a}$ and $q \geqslant 1$:
$\mathfrak{A}, \bar{a} \models \varphi \;\Leftrightarrow\; q \cdot \mathfrak{A}, \bar{a} \models \varphi$.

As always $\mathfrak{A} + \mathfrak{B}, a \sim \mathfrak{A}, a$ and $q \cdot \mathfrak{A}, \bar{a} \sim_{\mathsf{g}} \mathfrak{A}, \bar{a}$, any formula that is invariant under ordinary bisimulation is invariant under disjoint sums; any formula that is invariant under global bisimulation or under guarded bisimulation is invariant under disjoint copies. The following sums up the specialisations of Gaifman's theorem to be brought into play.

Proposition 14. *Both in the sense of classical and finite model theory:*

(a) *Any $\varphi(x) \in$ FO that is invariant under disjoint sums is equivalent to a local formula.*

(b) *Any $\varphi(x) \in$ FO that is invariant under disjoint copies is equivalent to a boolean combination of local formulae about x and simple local sentences.*

In both cases the converses are trivially valid, so that we actually have characterisation theorems: e.g. in connection with (a), a first-order formula $\varphi(x)$ is invariant under disjoint sums if and only if it is (equivalent to) a local formula. While the statements of the proposition are rather intuitive, their proofs from Gaifman's theorem itself require some book-keeping at the level of propositional logic.

Proof of Proposition 14. We explicitly prove these statements in their reading for finite model theory. For the classical case one can follow exactly the same lines.

(a) Assume that over finite structures, $\varphi(x)$ is invariant under disjoint sums. According to Gaifman's theorem, φ is equivalent (over all structures) to a formula of the form

$$\varphi(x) \equiv \bigvee_i \left(\varphi_0^i(x) \wedge \chi^i\right)$$

where the $\varphi_0^i(x)$ are ℓ_i-local about x and the χ^i are boolean combinations of basic local sentences such that w.l.o.g.

(i) every $\varphi_0^i(x) \wedge \chi^i$ is satisfiable.

(ii) any two distinct $\varphi_0^i(x)$ are mutually exclusive.

We may delete any disjuncts $\varphi_0^i(x) \wedge \chi^i$ that have no finite models, and still retain a formula that is equivalent to φ over all finite structures, where even

(i') every $\varphi_0^i(x) \wedge \chi^i$ is satisfiable in a finite model.

We claim that then necessarily

$$\varphi(x) \equiv_{\text{fin}} \bigvee_i \varphi_0^i(x).$$

Clearly $\varphi(x) \models_{\text{fin}} \bigvee_i \varphi_0^i(x)$. Conversely we show that also $\bigvee_i \varphi_0^i(x) \models_{\text{fin}} \varphi(x)$.

Let to this end $\mathfrak{A}, a \models \bigvee_i \varphi_0^i(x)$ be a finite model of $\bigvee_i \varphi_0^i(x)$. From (i') we obtain finite models $\mathfrak{B}_i, b_i \models \varphi_0^i(x) \wedge \chi^i$. Note that this implies $\mathfrak{B}_i, b_i \models \varphi(x)$. Let $\mathfrak{B}$ be the disjoint union of the $\mathfrak{B}_i$ and $\mathfrak{A}$. From invariance under disjoint sums we get $\mathfrak{B}, b_i \models \varphi(x)$ for each i. As $\mathfrak{B}, b_i \models \neg\varphi_0^j(x)$ for every $j \neq i$ by (ii), inspection of $\varphi(x)$ shows that necessarily $\mathfrak{B}, b_i \models \varphi_0^i(x) \wedge \chi^i$ for each i. Therefore $\mathfrak{B} \models \bigwedge_i \chi^i$. So $\mathfrak{B}, a \models \bigvee_i \varphi_0^i(x) \wedge \bigwedge_i \chi^i$. The latter formula clearly implies $\varphi(x)$. So $\mathfrak{B}, a \models \varphi(x)$, and using invariance under disjoint sums again, also $\mathfrak{A}, a \models \varphi(x)$.

(b) Let $\varphi(x)$ be invariant under disjoint copies over finite structures. Using Gaifman's theorem we obtain a presentation of $\varphi(x)$ of the following form:

$$\varphi \equiv \bigvee \Big(\psi^i \wedge \bigvee_j \left(\varphi_0^{ij}(x) \wedge \chi^{ij}\right)\Big)$$

where

(i) the formulae $\varphi_0^{ij}(x)$ are local about x.

(ii) the sentences χ^{ij} are boolean combinations of basic local sentences talking about scattered sets of size greater than 1.

(iii) the sentences ψ^i are complete boolean combinations of all formulae of the form $\exists y \rho(y)$ where ρ ranges over all the local formulae that occur in any of the ψ^i or χ^{ij}.

(iv) Any two ψ^i and $\psi^{i'}$ are mutually exclusive for $i \neq i'$.

(v) For every i: $\varphi_0^{ij}(x)$ and $\varphi_0^{ij'}(x)$ are mutually exclusive whenever $j \neq j'$.

(vi) For every i, j: $\psi^i \wedge (\varphi_0^{ij}(x) \wedge \chi^{ij})$ is satisfiable.

Again, where just $\equiv_{\text{fin}}$ is concerned, we further assume

(vi') For every i, j: $\psi^i \wedge (\varphi_0^{ij}(x) \wedge \chi^{ij})$ is satisfiable in a finite model.

We consider $\varphi^i(x) := \bigvee_j (\varphi_0^{ij}(x) \wedge \chi^{ij})$ on $\mathrm{MOD}(\psi^i)$ and claim that

$$\varphi^i(x) \wedge \psi^i \equiv_{\text{fin}} \psi^i \wedge \bigvee_j \varphi_0^{ij}(x).$$

Clearly this proves the claim of part (b).

The implication from left to right is obvious. For the opposite direction let $\mathfrak{A}, a \models \psi^i \wedge \bigvee_j \varphi_0^{ij}(x)$ be a finite model of the right-hand formula. We need to show that $\mathfrak{A}, a \models \varphi^i$. Choose $q \in \mathbb{N}$ to be greater than the cardinality of any scattered set mentioned in the χ^{ij}. By (vi') we find finite models $\mathfrak{B}_j, b_j \models \psi^i \wedge \varphi_0^{ij}(x) \wedge \chi^{ij}$. As $\mathfrak{B}_j, b_j \models \varphi(x)$ and by invariance of φ under disjoint copies, we have $q \cdot \mathfrak{B}_j, b_j \models \varphi(x)$. Clearly still $q \cdot \mathfrak{B}_j, b_j \models \psi^i$. Therefore, $q \cdot \mathfrak{B}_j, b_j \models \varphi^i(x)$ and as clearly also still $q \cdot \mathfrak{B}_j, b_j \models \varphi_0^{ij}(x)$, condition (v) implies that $q \cdot \mathfrak{B}_j, b_j \models \chi^{ij}$.

It follows that χ^{ij} can only make positive existential claims about large scattered sets whose members satisfy some local formula that is realised according to ψ^i; conversely, any negative statement in the χ^{ij} can only forbid (large) scattered sets satisfying some local formula that according to ψ^i cannot be realised at all.

It follows that also $q \cdot \mathfrak{A} \models \chi^{ij}$, and as this is the case for all choices of j, in fact $q \cdot \mathfrak{A} \models \bigwedge_j \chi^{ij}$. But then $q \cdot \mathfrak{A}, a \models \psi^i \wedge \bigvee_j \varphi_0^{ij} \wedge \bigwedge_j \chi^{ij}$. Therefore $q \cdot \mathfrak{A}, a \models \varphi$ and, by invariance of φ, also $\mathfrak{A}, a \models \varphi$. As the ψ^i are mutually exclusive, (iv), it must be that $\mathfrak{A}, a \models \varphi^i(x)$, as desired. ☐

A simple proof of Rosen's theorem This proof preserves core aspects of Rosen's proof [11], but uses Proposition 14 (a) to avoid surgery on finite transition systems. Let $\varphi(x) \in \mathrm{FO}$ be invariant under bisimulation over finite structures, hence invariant under disjoint sums over finite structures and therefore equivalent to an ℓ-local formula about x over finite structures. We may conclude the proof of the characterisation theorem with the following observation, and an application of Observation 10. (The argument is good for the classical case as well.)

Observation 15. *If $\varphi(x)$ is ℓ-local about x and $\sim$-invariant (over finite structures) then φ is invariant under ℓ-bisimulation (over finite structures).*

Proof. Let $\mathfrak{A}, a \sim^\ell \mathfrak{B}, b$ and assume that $\mathfrak{A}, a \models \varphi$ while $\mathfrak{B}, b \models \neg\varphi$. Let $\mathfrak{A}^*$ be the result of unravelling $\mathfrak{A}$ from a, restricting the resulting tree to depth ℓ, and identifying leaf nodes in this truncated tree with corresponding nodes in disjoint isomorphic copies of $\mathfrak{A}$. Let $\mathfrak{A}^*_\ell = \mathfrak{A}^* \restriction N^\ell(a)$ be the truncated unravelling of depth ℓ of $\mathfrak{A}, a$. Clearly $\mathfrak{A}^*$ and $\mathfrak{A}^*_\ell$ are finite if $\mathfrak{A}$ is finite. Let $\mathfrak{B}^*$ and $\mathfrak{B}^*_\ell$ be similarly obtained from $\mathfrak{B}, b$. Then $\mathfrak{A}^*_\ell \sim \mathfrak{B}^*_\ell$, as $\mathfrak{A}, a \sim^\ell \mathfrak{B}, b$ and as both $\mathfrak{A}^*_\ell$ and $\mathfrak{B}^*_\ell$ are trees of depth ℓ. By $\sim$-invariance of φ, $\mathfrak{A}^*, a \models \varphi$ and $\mathfrak{B}^*, b \models \neg\varphi$. By ℓ-locality of φ, also $\mathfrak{A}^*_\ell, a \models \varphi$ while $\mathfrak{B}^*_\ell, b \models \neg\varphi$. Contradiction. ☐

A weaker locality result suffices We find a better proof of Rosen's theorem, which replaces the use of Gaifman's theorem by a direct Ehrenfeucht-Fräïssé argument and also yields tight bounds on the modal nesting depth in terms of the first-order quantifier rank. As usual we let $\equiv_q$ stand for for first-order equivalence up to quantifier rank q, as characterised by the existence of a winning strategy for the second player, player II, in the corresponding q-round Ehrenfeucht-Fraïssé game (see, e.g., [6]).

Lemma 16. *Let $\mathfrak{A}, a$ and $\mathfrak{B}, b$ agree on all ℓ-local* FO *formulae of quantifier rank up to q, where $\ell = 2^q$. Then $q \cdot \mathfrak{A} + q \cdot \mathfrak{B}, a \equiv_q q \cdot \mathfrak{A} + q \cdot \mathfrak{B}, b$.*

Proof sketch. The proof is by a simple game argument to show that player II wins the q round Ehrenfeucht-Fraïssé game on the composite structures provided she has a winning strategy for q rounds on $\mathfrak{A} \restriction N^\ell(a), a$ against $\mathfrak{B} \restriction N^\ell(b), b$ for $\ell = 2^q$.

In round k of the game we let player II answer isomorphically any move of player I that goes to a position whose distance from the distinguished elements and all previously played positions is greater than $d_k = 2^{q-k}$ in a spare, i.e., as yet unused, isomorphic copy of the respective structure in the opposite domain. Other moves are either played isomorphically into copies determined by the closest position previously played and/or according to the local strategy in the copies with the distinguished elements. In other words, player II can essentially make do with preserving locality up to d_k and regarding distances greater than d_k as infinite. ☐

As $q \cdot \mathfrak{A} + q \cdot \mathfrak{B}, a \sim \mathfrak{A}, a$ and $q \cdot \mathfrak{A} + q \cdot \mathfrak{B}, b \sim \mathfrak{B}, b$, we immediately get the following, which together with Observations 15 and 10 proves the quantitative strengthening of the crucial direction in the van Benthem and Rosen theorems indicated in the theorem below.

Corollary 17. *A first-order formula $\varphi(x)$ of quantifier rank q that is bisimulation invariant (over finite structures) is ℓ-local (over finite structures) for $\ell = 2^q$.*

Theorem 18. *Any first-order formula $\varphi(x)$ of quantifier rank q that is invariant under bisimulation (over finite structures) is equivalent (over finite structures) to a formula of basic modal logic of nesting depth bounded by 2^q.*

The quantifier rank $q + 1$ FO property that "there is a red node within distance 2^q of x" provides a matching exponential lower bound for the modal nesting depth.

3 Avoiding short cycles in finite transition systems

Definition 19. Let $\mathfrak{A}$ be a τ-structure, $G(\mathfrak{A})$ its Gaifman graph.

(i) A *cycle* (of length ℓ) in $\mathfrak{A}$ is an ℓ-cycle in $G(\mathfrak{A})$ in the graph theoretic sense: a sequence of vertices $a_0, \ldots, a_{\ell-1}$, where for each consecutive pair of indices $(i, i+1)$ (cyclically understood in the sense of $\mathbb{Z}_\ell$) we have $(a_i, a_{i+1}) \in R$ or $(a_{i+1}, a_i) \in R$ for some R in τ.

(ii) A cycle is *non-degenerate* if always $a_{i-1} \neq a_{i+1}$; it is called *simple* if $a_i \neq a_j$ for $i \neq j$.

(iii) $\mathfrak{A}$ is *acyclic* if it does not have non-degenerate cycles.

(iv) $\mathfrak{A}$ is *k-acyclic* if it does not have non-degenerate cycles of lengths $< k$.

A k-acyclic transition system is locally acyclic in the sense that its ℓ-neighbourhoods are acyclic if $k > 2\ell$. Degenerate cycles cannot be avoided at all, as every edge gives rise to a degenerate cycle of length 2. For technical purposes we introduce the following notion.

Definition 20. Let $\mathfrak{A}$ be a τ-structure, E the union of the binary relations $R \in \tau$ over $\mathfrak{A}$. $\mathfrak{A}$ is *simple* if the R are mutually disjoint and E is anti-symmetric and irreflexive.

In graph theoretic terms one might consider a simple structure as an edge-partitioned and vertex-coloured tournament. The usual two-way unravellings provide bisimilar covers by simple acyclic transition systems, albeit generally infinite ones. Clearly no cyclic finite $\mathfrak{A}$ can have a finite acyclic bisimilar companion. Our aim is the following.

Proposition 21. *Every finite transition system $\mathfrak{A}$ admits, for every $k > 3$, a bisimilar cover $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$ by a finite k-acyclic transition system $\hat{\mathfrak{A}}$. Moreover, $\hat{\mathfrak{A}}$ and π can be chosen such that $\hat{\mathfrak{A}}$ is simple and π preserves the in- and out-degrees w.r.t. each individual binary relation R in τ.*

A first attempt would use tree-like unravellings to a certain depth beyond which further requirements are realised through links going back to nodes close to the root. It is not at all clear, however, where the target nodes for these necessary extra links are to be chosen so as to avoid the generation of new short cycles – note that the cycles under consideration are cycles in $G(\mathfrak{A})$ where edges are undirected. The following construction overcomes this difficulty in a uniform, symmetry and multiplicity preserving manner, albeit with a hyper-exponential blow-up in size. It remains open whether there is also a more feasible construction.

The case of simple transition systems The assumption of simplicity simplifies the proof of the proposition; the general case can then be reduced to it. Let $\mathfrak{A}$ be simple and ℓ-acyclic ($\ell \geqslant 3$). We shall produce a finite bisimilar cover by a simple transition system that is n-acyclic for $n = 2(\ell+1)$, also preserving degrees. Repeated application yields the required cover for any given value of k. Put

$$\hat{A} = A \times \prod_{e \in E} \mathbb{Z}_n, \qquad \pi\colon \hat{A} \to A \text{ the natural projection.}$$

Consider $\prod_{e\in E} \mathbb{Z}_n$ as an additive group with generators $\mathfrak{n}_e$ indexed by the edges $e \in E$, where $\mathfrak{n}_e$ has a 1 in the e-component, and 0 else. We write $\hat{a} = (a, \mathfrak{a})$ for a typical element $\hat{a}$ above a in $\hat{A}$. Let $e = (a, a') \in E$. We put an edge (of type R or R^{-1}, as the case may be, depending on the nature of e) between $(a, \mathfrak{a})$ and $(a', \mathfrak{a}')$ if, and only if $\mathfrak{a}' = \mathfrak{a} + \mathfrak{n}_e$. Together with the obvious choice for the unary predicates this fixes $\hat{\mathfrak{A}}$. $\hat{\mathfrak{A}}$ is simple and $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$ a homomorphism by construction. Z_π as induced by π is a global two-way (in fact even guarded) bisimulation. The forth-property is obvious; for the back-property consider $\hat{a} = (a, \mathfrak{a})$ above a and, e.g., $(a', a) \in R$ in $\mathfrak{A}$; then $e = (a', a) \in E$ and we find that $((a', \mathfrak{a}'), (a, \mathfrak{a})) \in R$ in $\hat{\mathfrak{A}}$ for $\mathfrak{a}' := \mathfrak{a} - \mathfrak{n}_e$. π is 1–1 in restriction to any 1-neighbourhood in $\hat{\mathfrak{A}}$, and therefore also degree preserving. It also follows that every $(E \cup E^{-1})$-path from some start vertex a in $\mathfrak{A}$ has a unique lift to a path $\hat{\mathfrak{A}}$ from any given $\hat{a}$ above a.

It remains to check that $\hat{\mathfrak{A}}$ is n-acyclic, $n = 2(\ell + 1)$. Let $((a_i, \mathfrak{a}_i))_{0 \leqslant i < m}$ be a non-degenerate cycle of length m in $G(\hat{\mathfrak{A}})$. Let $e_i \in E$ be the underlying edges in $\mathfrak{A}$: $e_i = (a_i, a_{i+1})$ or $e_i = (a_{i+1}, a_i)$ as he case may be. Let further $\mathfrak{n}_i \in \prod_{e\in E} \mathbb{Z}_n$ be chosen as $\mathfrak{n}_i = \mathfrak{n}_{e_i}$ if $e_i = (a_i, a_{i+1})$, and as $\mathfrak{n}_i = -\mathfrak{n}_{e_i}$ if $e_i = (a_{i+1}, a_i)$. It follows that $\sum_i \mathfrak{n}_i = \mathfrak{o}$ in $\prod_{e\in E} \mathbb{Z}_n$. Therefore, the contributions in the e-component of $\prod_{e\in E} \mathbb{Z}_n$ must cancel for every single e among the e_i. For each such e there are two (non-exclusive) cases:

(i) $e_i = e$ for at least n values of i that all contribute summands of the same sign (the cycle in $\hat{\mathfrak{A}}$ crosses above e in the same direction at least n times); so $m \geqslant n$.

(ii) $e_i = e$ for at least two values of i that contribute summands of opposite signs (the cycle in $\hat{\mathfrak{A}}$ crosses above e in opposite directions).

If case (ii) applies to at least one edge, then the projection of the cycle must traverse that edge e at least twice, in opposite directions. By non-degeneracy in $\hat{\mathfrak{A}}$, and as we have locally unique lifts, no conjugate pairs of edges can occur in contiguous positions along the cycle. It follows that the trace of the projection to $\mathfrak{A}$ must contain at least two distinct simple cycles apart from the two traversals of e. By assumption, the length of each of these simple parts is at least ℓ. This also implies that $m \geqslant 2\ell + 2 = n$. □

The general case The case of structures $\mathfrak{A} = (A, \bar{R}, \bar{P})$ that are not simple is reduced to the simple case by means of a bisimilar cover $\pi : \mathfrak{A}' \to \mathfrak{A}$ by some simple $\mathfrak{A}'$ such that π preserves in- and out-degrees. This can be achieved using an intermediate encoding in which edges of $\mathfrak{A}$ are replaced by paths of length two that pass through new vertices whose colour characterises the kind of edge involved.

Let $\mathfrak{A} = (A, \bar{R}, \bar{P})$ be a not necessarily simple τ-structure. Associate with τ a new vocabulary τ_s consisting of $\tau^{(1)}$ (the unary predicates in τ) and the following:

Q_R	a new unary predicate for each $R \in \tau^{(2)}$,
S	a new binary predicate.

With an arbitrary τ-transition system $\mathfrak{A}$ we associate the following simple τ_s-transition system $\mathfrak{A}_s$.

$$\begin{aligned}
\mathfrak{A}_s &= \big(A_s, S, (P)_{P \in \tau^{(1)}}, (Q_R)_{Q \in \tau^{(2)}}\big) \\
A_s &= A \cup \dot{\bigcup}_{R \in \tau^{(2)}} R^{\mathfrak{A}} \quad \text{(disjoint union)} \\
P &= P^{\mathfrak{A}} \\
Q_R &= \{e : e \in R^{\mathfrak{A}}\} \\
S &= \{(a, e), (e, a') : (a, a') \in R^{\mathfrak{A}}, R \in \tau\}
\end{aligned}$$

Clearly $\mathfrak{A}_s$ is simple. We may now apply the above construction to obtain a bisimilar cover $\pi_s : \hat{\mathfrak{A}}_s \to \mathfrak{A}_s$ by a simple, 5-acyclic τ_s-transition system $\hat{\mathfrak{A}}_s$, where π_s preserves in- and out-degrees of S.

It follows that directed S-paths of length two in $\mathfrak{A}_s$ of the form a, e, a' have unique lifts to any $\hat{a} \in \pi_s^{-1}(a)$ or any $\hat{a}' \in \pi_s^{-1}(a')$. Conversely any length two directed S-path of the form $\hat{a}, \hat{e}, \hat{a}'$ in $\hat{\mathfrak{A}}_s$ with $\hat{e} \in Q_R$ projects to a path a, e, a' in $\mathfrak{A}_s$ with $e \in Q_R$, whence it corresponds to an R-edge in $\mathfrak{A}$.

Any such τ_s transition system induces a simple τ transition system, according to the following straightforward reverse transformation.

$$\begin{aligned}
\hat{\mathfrak{A}} &= \big(\hat{A}, \bar{P}, \bar{R}\big) \\
\hat{A} &= \pi_s^{-1}(A) \\
P &= P^{\hat{\mathfrak{A}}_s} \\
R &= \big\{(\hat{a}, \hat{a}') : (\hat{a}, \hat{e}), (\hat{e}, \hat{a}') \in S^{\hat{\mathfrak{A}}_s}, \text{ for some } e \in Q_R^{\hat{\mathfrak{A}}_s}\big\}
\end{aligned}$$

Since $\hat{\mathfrak{A}}_s$ does not have any non-degenerate 4-cycles, $\hat{\mathfrak{A}}$ turns out simple. The above considerations about projections and unique lifts of paths imply that π_s induces a homomorphism $\pi : \hat{\mathfrak{A}} \to \mathfrak{A}$ which moreover still is a bisimilar cover. Further, π preserves in- and out-degrees for every individual R. To sum up, we have found a bisimilar cover $\pi : \hat{\mathfrak{A}} \to \mathfrak{A}$ as in the following corollary.

Corollary 22. *Any finite transition system $\mathfrak{A}$ admits a bisimilar cover $\pi : \hat{\mathfrak{A}} \to \mathfrak{A}$ by a finite simple transition system $\hat{\mathfrak{A}}$ where π moreover preserves in- and out-degrees for all binary predicates.*

Variation for guarded covers The above construction naturally extends to guarded covers of τ-structures. One can use a simple encoding based on new binary relations for every non-degenerate guarded quantifier-free 2-type. Similar considerations for guarded logics on graphs are presented in [8]. Some care has to be taken here with respect to symmetric types and types of non-guarded pairs. The extra layer of encoding uses new binary relations for every non-degenerate guarded quantifier-free 2-type.

A (non-degenerate, quantifier-free) 2-type over τ is a full description of the isomorphism type of a two element τ-structure in variables x, y. It may be formalised as a conjunction over a maximally consistent set of atomic and negated atomic τ-formulae in variables x and y including the conjunct $x \neq y$. We write $p(x, y)$ for a 2-type, and $\mathrm{tp}_{\mathfrak{A}}(a, a')$ for the unique 2-type satisfied by (a, a') in $\mathfrak{A}$, for $a \neq a'$.

A 1-type $q(x)$ over τ similarly is a full description of a one element τ-structure (which apart from monadic information contains the information about loops w.r.t. the binary predicates). Obvious notation like $\mathrm{tp}_{\mathfrak{A}}(a)$ applies.

For a 2-type $p = p(x, y)$ we let p_x and p_y be the unique 1-types obtained as the restrictions of p to its x-part or y-part, respectively.

Let p^{-1} stand for the result of swapping x and y in p. A 2-type p is *symmetric* if $p = p^{-1}$, asymmetric otherwise.

Definition 23. A 2-type $p(x, y)$ over τ is *guarded* if it includes a conjunct Rxy or Ryx for some binary relation R of τ. In other words, guarded 2-types are those 2-types that are realised by non-degenerate guarded tuples.

Let Λ_2 be a fixed maximal set of guarded 2-types over τ containing all symmetric guarded 2-types, and precisely one of p or p^{-1} for every asymmetric guarded 2-type. Let Λ_1 be the set of all 1-types over τ. Associate with τ a new vocabulary τ_g consisting of a unary predicate P_q for each $q \in \Lambda_1$ and a binary predicate R_p for each $p \in \Lambda_2$.

With a τ-structure $\mathfrak{A}$, and an auxiliary linear ordering $<$ on its universe, we associate the following simple τ_g-

transition system $\mathfrak{A}_g$.

$$\begin{array}{rcl} \mathfrak{A}_g & = & \big(A, (P_q)_{q\in\Lambda_1}, (R_p)_{p\in\Lambda_2}\big) \\ P_q & = & \big\{a \colon q = \mathrm{tp}_{\mathfrak{A}}(a)\big\} \qquad \text{for } q \in \Lambda_1, \\ R_p & = & \left\{ \begin{array}{l} \big\{(a,a') \colon a < a' \text{ and } \mathrm{tp}_{\mathfrak{A}}(a,a') = p\big\} \\ \qquad\qquad \text{for } p \in \Lambda_2 \text{ symmetric,} \\ \big\{(a,a') \colon \mathrm{tp}_{\mathfrak{A}}(a,a') = p\big\} \\ \qquad\qquad \text{for } p \in \Lambda_2 \text{ asymmetric.} \end{array} \right. \end{array}$$

Clearly $\mathfrak{A}_g$ is simple and satisfies the following compatibility conditions:

(i) the P_q partition the universe.

(ii) if $(a, a') \in R_p$ then $a \in P_{p_x}$ and $a' \in P_{p_y}$.

(iii) for any non-degenerate pair (a, a'), at most one binary relation R_p can link a with a' (simplicity).

Conversely, for any simple τ_g-transition system $\mathfrak{B}_g$ satisfying (i) and (ii) there is a unique associated τ-structure $\mathfrak{B}$. The universe of $\mathfrak{B}$ is that of $\mathfrak{B}_g$. Monadic and binary predicates from τ are interpreted so as to be consistent with the 1- and 2-types prescribed by the P_q and the R_p, and such that a non-degenerate pair (b, b') will be guarded in $\mathfrak{B}$ if and only if b and b' are linked by some R_p in $\mathfrak{B}_g$. With this it is not hard to check the following.

Lemma 24. (i) *Let $\mathfrak{B}_g$ and $\mathfrak{B}'_g$ be simple τ_g transition systems. Let $\mathfrak{B}_g \approx \mathfrak{B}'_g$ and let $\mathfrak{B}_g$ satisfy conditions (i) and (ii) above. Then $\mathfrak{B}'_g$ also satisfies (i) and (ii), and $\mathfrak{B} \sim_{\mathrm{g}} \mathfrak{B}'$ for the associated τ-structures.*

(ii) *Let $\mathfrak{A}$ be a τ-structure with an associated τ_g-transition system $\mathfrak{A}_g$. Let $\pi\colon \hat{\mathfrak{A}}_g \to \mathfrak{A}_g$ be a bisimilar cover of $\mathfrak{A}_g$ by a simple τ_g-transition system $\hat{\mathfrak{A}}_g$. Then $\hat{\mathfrak{A}}_g$ satisfies (i) and (ii) and for the induced τ-structure $\hat{\mathfrak{A}}$ we have that $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$ is a guarded cover. If $\hat{\mathfrak{A}}_g$ is k-acyclic then so is $\hat{\mathfrak{A}}$, and if $\pi\colon \hat{\mathfrak{A}}_g \to \mathfrak{A}_g$ preserves in- and out-degrees then $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$ preserves incidence degrees w.r.t. guarded 2-types.*

Together with the covering results for (simple) transition systems this gives the following.

Corollary 25. *Let τ be any finite relational vocabulary of width two, $\mathfrak{A}$ a finite τ-structure and $k > 3$. Then there is a guarded cover $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$ by a finite k-acyclic structure $\hat{\mathfrak{A}}$. π can be chosen such that π preserves incidence degrees with guarded 2-types.*

An open issue related to this result concerns potential extensions to the setting of arbitrary relational vocabularies. We do not know whether one can similarly achieve finite guarded covers of finite relational structures that avoid short chordless cycles. See [9] for a discussion. In that paper another aspect of acyclicity (in hypergraphs) – to do with the avoidance of bad cliques rather than cycles in the Gaifman graph – is shown to be realisable in finite guarded covers, with applications to the clique guarded fragment and extension theorems for partial isomorphisms.

4 Locally nice covers and characterisation theorems

We use the results from the previous section to obtain bisimilar companions which can serve to lift ℓ-bisimulations between finite transition systems to stronger forms of local first-order equivalence. These techniques can then be employed in the proof of our characterisation theorems on the basis of Observation 10. The degree of similarity between structures to which we boost ℓ-bisimulation equivalence in the following lemma captures the local first-order properties addressed in Proposition 14.

Lemma 26. *Let Ψ be a finite collection of ℓ-local first-order formulae. For any two (finite) τ transition systems $\mathfrak{A}$ and $\mathfrak{B}$: if $\mathfrak{A}, a \approx^{\ell} \mathfrak{B}, b$ then there are (finite) bisimilar companions $\mathfrak{A}^*, a \approx \mathfrak{A}, a$ and $\mathfrak{B}^*, b \approx \mathfrak{B}, b$ such that for all $\psi \in \Psi$:*

(i) $\mathfrak{A}^*, a \models \psi \;\Leftrightarrow\; \mathfrak{B}^*, b \models \psi$.

(ii) $\mathfrak{A}^* \models \exists x \psi(x) \;\Leftrightarrow\; \mathfrak{B}^* \models \exists x \psi(x)$.

For the proof we need another lemma, which is obtained as a straightforward application of first-order Ehrenfeucht-Fraïssé techniques. The following definition captures a notion of local richness in a simple acyclic structure.

Definition 27. Let $\mathfrak{A}{\restriction}N^{\ell}(a)$ be simple and acyclic. We say that $\mathfrak{A}{\restriction}N^{\ell}(a)$ is *q-rich* if every path $a, a_1, \ldots, a_m$ from a of length $m \leqslant \ell$ has at least q disjoint isomorphic copies $a, a'_1, \ldots, a'_m$ (disjoint apart from their common start node a, that is).

Lemma 28. *Let $\mathfrak{A}, a \approx^{\ell} \mathfrak{B}, b$ and assume that $\mathfrak{A}{\restriction}N^{\ell}(a), a$ and $\mathfrak{B}{\restriction}N^{\ell}(b), b$ are both acyclic and q-rich. Then $\mathfrak{A}, a$ and $\mathfrak{B}, b$ agree on all ℓ-local* FO *formulae of quantifier rank up to q.*

Proof of Lemma 28. It suffices to exhibit a winning strategy for the second player in the q-round Ehrenfeucht-Fraïssé game on $\mathfrak{A}{\restriction}N^{\ell}(a), a$ and $\mathfrak{B}{\restriction}N^{\ell}(b), b$. Thinking of these acyclic structures as trees rooted in a and b respectively, we describe a strategy for the second player. For pebble placements $\bar{a}$ in $\mathfrak{A}{\restriction}N^{\ell}(a)$ we let $\mathrm{tree}(\bar{a}, a)$ stand for the set of those nodes that lie on shortest paths between the root a and any one of the pebbled nodes $\bar{a}$; and similarly in $\mathfrak{B}, b$. The strategy for the second player consists of maintaining an isomorphism $\rho\colon \mathfrak{A}{\restriction}\mathrm{tree}(\bar{a}, a) \simeq \mathfrak{B}{\restriction}\mathrm{tree}(\bar{b}, b)$ such that $\rho(\bar{a}, a) = (\bar{b}, b)$ and if $c \in \mathrm{tree}(\bar{a}, a)$ has distance d from a then $\mathfrak{A}, c \sim^{\ell-d} \mathfrak{B}, \rho(c)$. This condition can be maintained for q rounds, as there are always 'new' paths available where needed, thanks to q-richness. ☐

Proof of Lemma 26. We may w.l.o.g. assume that the given transition systems $\mathfrak{A}$ and $\mathfrak{B}$ are simple; otherwise we pass to simple bisimilar covers according to Corollary 22. Let q be the maximal quantifier rank of any of the local formulae in Ψ.

Let $\mathfrak{A}, a \approx^\ell \mathfrak{B}, b$, both simple. To both $\mathfrak{A}$ and $\mathfrak{B}$ we apply the following construction, which we explicitly describe for $\mathfrak{A}$. First inflate $\mathfrak{A}$ to obtain a structure $\mathfrak{A} * q$ as follows:

$$\mathfrak{A} * q = \big(A \times \{1, \dots, q\}, (\pi_q^{-1}(R)), (\pi_q^{-1}(P))\big)$$

where $\pi_q \colon \mathfrak{A} * q \to \mathfrak{A}$ is the natural projection. $\pi_q \colon \mathfrak{A} * q \to \mathfrak{A}$ is a bisimilar cover. Now choose $k > 2\ell$ and obtain $\pi_0 \colon \mathfrak{A}^* \to \mathfrak{A} * q$ as a degree-preserving k-acyclic bisimilar cover of $\mathfrak{A} * q$, according to Lemma 21. Altogether we have obtained a bisimilar cover $\pi \colon \mathfrak{A}^* \to \mathfrak{A}$ of $\mathfrak{A}$ by a structure $\mathfrak{A}^*$ whose ℓ-neighbourhoods are simple, acyclic and q-rich. Here q-richness is a consequence of the fact that $\mathfrak{A} * q$ has q-fold multiplicities and that π_0 preserves in- and out-degrees.

Let $\mathfrak{B}^*$ similarly be obtained as a cover of $\mathfrak{B}$. By Lemma 28, $\mathfrak{A}^*$ and $\mathfrak{B}^*$ are as desired. □

The following is obtained as an exact analogue of Lemma 26, based on simple encodings of guarded 2-types (cf. the construction used for Corollary 25).

Lemma 29. *Let Ψ be a finite collection of ℓ-local first-order formulae. For any two (finite) τ-structures $\mathfrak{A}$ and $\mathfrak{B}$: if $\mathfrak{A}, a \sim_{\mathrm{g}}^\ell \mathfrak{B}, b$ then there are (finite) guarded bisimilar companions $\mathfrak{A}^*, a \sim_{\mathrm{g}} \mathfrak{A}, a$ and $\mathfrak{B}^*, b \sim_{\mathrm{g}} \mathfrak{B}, b$ such that for all $\psi \in \Psi$:*

(i) $\mathfrak{A}^*, a \models \psi \;\Leftrightarrow\; \mathfrak{B}^*, b \models \psi$.

(ii) $\mathfrak{A}^* \models \exists x \psi(x) \;\Leftrightarrow\; \mathfrak{B}^* \models \exists x \psi(x)$.

To finish the arguments for Theorem 3 and 5 we follow the pattern of Observation 10 and finally prove the following.

Proposition 30. *Let $\varphi(x)$ be a first-order formula in a vocabulary of width 2. The following are valid classically and in the sense of finite model theory.*

(a) *If φ is invariant under global two-way bisimulation $\approx$ then φ is invariant under $\approx^\ell$ for some finite ℓ.*

(b) *If φ is invariant under guarded bisimulation $\sim_{\mathrm{g}}$ then φ is invariant under $\sim_{\mathrm{g}}^\ell$ for some finite ℓ.*

Proof. In both cases, according to Proposition 14, φ is equivalent (over finite structures) to a boolean combination of (possibly existentially quantified) local formulae. Let Ψ be a finite collection of ℓ-local formulae such that φ can be represented as a boolean combination over (existentially quantified) formulae in Ψ, for suitable ℓ.

In (a) φ is invariant (over finite structures) under $\approx^\ell$ by the following reasoning. Let $\mathfrak{A}, a \approx^\ell \mathfrak{B}, b$ and assume that $\mathfrak{A}, a \models \varphi$. Find bisimilar companions $\mathfrak{A}^*, a \approx \mathfrak{A}, a$ and $\mathfrak{B}^*, b \approx \mathfrak{B}, b$ according to Lemma 26. It follows that $\mathfrak{A}^*, a \models \varphi$, by invariance under $\approx$. But then $\mathfrak{B}^*, b \models \varphi$, as $\mathfrak{A}^*, a$ and $\mathfrak{B}^*, b$ agree on all (existentially quantified) formulae in Ψ. So we get $\mathfrak{B}, b \models \varphi$, by invariance under $\approx$ again.

(b) is completely analogous, with an appeal to Lemma 29. □

References

[1] 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.

[2] J. Barwise and J. van Benthem, *Interpolation, preservation, and pebble games*. Journal of Symbolic Logic 64 (1999), pp. 881–903.

[3] J. van Benthem, Modal Logic and Classical Logic, Bibliopolis, Napoli, 1983.

[4] J. van Benthem, Exploring Logical Dynamics, CSLI Publications, Stanford University, 1996.

[5] P. Blackburn, M. de Rijke, and Y. Venema, Modal Logic, Cambridge University Press, 2001.

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

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

[8] E. Grädel, C. Hirsch, and M. Otto, *Back and forth between guarded and modal logics*, to appear in ACM Transactions of Computational Logic, 2002. Also compare: Proc. 15th IEEE Symp. on Logic in Computer Science, LICS 2000, 217 – 228.

[9] I. Hodkinson and M. Otto, *Finite conformal hypergraph covers and Gaifman cliques in finite structures*, preprint 2001.

[10] E. Hoogland and M. Marx, *Interpolation in guarded fragments*, to appear in Studia Logica.

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