Chapter 1
The Freedoms of (Guarded) Bisimulation

Erich Grädel and Martin Otto

Abstract We survey different notions of bisimulation equivalence that provide flexible and powerful concepts for understanding the expressive power as well as the model-theoretic and algorithmic properties of modal logics and of more and more powerful variants of guarded logics. An appropriate notion of bisimulation for a logic allows us to study the expressive power of that logic in terms of semantic invariance and logical indistinguishability. As bisimilar nodes or tuples in two structures cannot be distinguished by formulae of the logic, bisimulations may be used to control the complexity of the models under consideration. In this manner, bisimulation-respecting model constructions and transformations lead to results about model-theoretic properties of modal and guarded logics, such as the tree model property of modal logics and the fact that satisfiable guarded formulae have models of bounded tree width. A highlight of the bisimulation-based analysis are the characterisation theorems: inside a classical level of logical expressiveness such as first-order or monadic second-order definability, these provide a tight match between bisimulation invariance and logical definability. Typically such characterisation theorems state that a modal or guarded logic is not only invariant under bisimulation but, conversely, also expressively complete for the class of all bisimulation invariant properties at that level. Finally, the bisimulation-based analysis of modal and guarded logics also leads to important insights concerning their algorithmic properties. Since satisfiable formulae always admit simple models, for instance tree-like ones, and since modal and guarded logics can be embedded or interpreted in monadic second-order logic on trees, powerful automata theoretic methods become available for checking satisfiability and for evaluating formulae.

E. Grädel (✉)
RWTH Aachen University, Aachen, Germany
e-mail: graedel@logic.rwth-aachen.de

M. Otto
TU Darmstadt, Darmstadt, Germany
e-mail: otto@mathematik.tu-darmstadt.de

A. Baltag and S. Smets (eds.), *Johan van Benthem on Logic and Information Dynamics*, Outstanding Contributions to Logic 5,
DOI: 10.1007/978-3-319-06025-5_1,

1.1 Introduction

Bisimulation equivalence is one of the leading themes in modal logic. As the quintessential back-and-forth notion for two-player combinatorial games it may not only be regarded as a special case in the model-theoretic tradition of Ehrenfeucht–Fraïssé games but may also be seen as their common backbone. Bisimulation equivalence (of game graphs or transition systems) grasps the complex equivalence between dynamic behaviours as a natural structural equivalence. The generalisation of this graph-based bisimulation concept to higher dimensions in the form of guarded bisimulation opened up one further branch in the rich world of model-theoretic games; the study of guarded bisimulation in the wake of the inception of the guarded fragment of first-order logic in [1] has led to a new conceptual understanding of well-behaved logics that are 'modal' in a more general sense. Guarded logics far transcend basic modal logics while retaining some of the key features of modal model theory precisely through the parallelism between the underlying notions of bisimulation equivalence. Guarded bisimulation can be seen as derived from a hypergraph version of ordinary (modal, graph-based) bisimulation. And just as preservation under ordinary bisimulation accounts for much of the good model-theoretic behaviour of modal logics, so hypergraph bisimulation and guarded bisimulation are the keys to understanding the model theory of guarded logics. Model constructions and transformations that are compatible with guarded bisimulation account for the malleability of models and the tractability of the finite and algorithmic model theory of various guarded logics. We here survey and summarise a number of model-theoretic techniques and results, especially in the light of bisimulation respecting model constructions, including some more recent developments. Results to be surveyed include finite and small model properties, decidability results, complexity and expressive completeness issues. Among the more recent developments are notions of guardedness that focus on the role of negation rather than on just the quantification pattern. Unary and guarded negation bisimulation and the corresponding unary and guarded negation fragments of first-order logic from [10] and [3] have contributed yet another aspect to our understanding of the good behaviour of 'modal' logics with a yet wider scope.

1.2 Bisimulation: Behavioural and Structural Equivalence

1.2.1 Ehrenfeucht–Fraïssé, Back-and-forth, Zig-zag, Pebble Games: Games Model-Theorists Play

Notions like 'behaviour' and 'strategies' seem to be quintessentially dynamic, while the analysis of structure and structural comparisons are mostly construed as static concerns. Yet modal logics, transition systems and game graphs bridge the apparent gap in a natural manner and typically allow us to understand behavioural comparisons as structural comparisons, and behavioural equivalences as structural equivalences. This is not even really surprising if we remind ourselves how, e.g., game graphs can be regarded as extensional (and static) descriptions of the possible plays (hence behaviours) of the game, so that, e.g., the existence of a winning strategy for one of the players can be determined by structural analysis. The dynamics and intuitive appeal of games can also be harnessed for the analysis of the semantics and expressive power of logics: model checking games account for the evaluation of logical formulae over structures, and model comparison games are used to account for distinctions and degrees of indistinguishability between structures w.r.t. properties expressible in a given logic. In the classical context of first-order logic the model comparison games are at the centre of the Ehrenfeucht–Fraïssé technique.

In the world of modal logics, the essential model comparison game is the *bisimulation* game. It is a typical model-theoretic back and forth game, played by two players over the two structures at hand (Kripke structures or transition systems). A position in the game is a pair of (similar) nodes, one from each of the two structures, marked by pebbles; players take turns to move the pebbles along available transitions in the respective structure; in each new round the first player is free to choose one of the structures and one of the available transitions to move the pebble across that transition, and the second player must respond likewise in the opposite structure. Overall, the game protocol ensures that the second player has a winning strategy in a position precisely if—recursively—every transition in the one structure can be matched by a transition in the opposite structure, ad infinitum. Bisimulation relations and bisimulation equivalence capture this notion of game equivalence by means of back&forth closure conditions on a (or the maximal) set of pairs that are winning positions for the second player.

Definition 1.1 For structures $\mathfrak{A} = (A, (R_i^{\mathfrak{A}}), (P_j^{\mathfrak{A}}))$ and $\mathfrak{B} = (B, (R_i^{\mathfrak{B}}), (P_j^{\mathfrak{B}}))$ with binary accessibility relations R_i and unary predicates P_j:

A binary relation $Z \subseteq A \times B$ between the nodes of $\mathfrak{A}$ and nodes of $\mathfrak{B}$ is a *bisimulation relation* if for all $(a, b) \in Z$:

(i) (*atom eq.*): for each P_j, $a \in P_j^{\mathfrak{A}}$ iff $b \in P_j^{\mathfrak{B}}$;
(ii) (R_i*-back*): for every b' with $(b, b') \in R_i^{\mathfrak{B}}$ there is some a' such that $(a, a') \in R_i^{\mathfrak{A}}$ and $(a', b') \in Z$;
(iii) (R_i*-forth*): for every a' with $(a, a') \in R_i^{\mathfrak{A}}$ there is some b' such that $(b, b') \in R_i^{\mathfrak{B}}$ and $(a', b') \in Z$.

As the union of bisimulation relations is again a bisimulation relation, there is a well-defined $\subseteq$-maximal *largest bisimulation* between $\mathfrak{A}$ and $\mathfrak{B}$. Pointed structures $\mathfrak{A}, a$ and $\mathfrak{B}, b$ are *bisimilar*, $\mathfrak{A}, a \sim \mathfrak{B}, b$, if (a, b) is in some (hence in the largest) bisimulation between $\mathfrak{A}$ and $\mathfrak{B}$.

Clearly $\sim$ captures a strong form of behavioural equivalence, if we think of 'behaviours' not just as traces of actions, but rather as the complex interactive and responsive patterns that can evolve in any step-wise alternating exploration of potential transitions. The conditions (R_i*-back*) and (R_i*-forth*) capture the challenge-response requirements posed for the second player by one additional round.

Correspondingly, the largest bisimulation on $A \times B$ forms a greatest fixed point w.r.t. the refinement operator induced by (*atom eq.*) and the (R_i*-back*) and (R_i*-forth*) conditions:

$$Z \longmapsto \mathscr{F}(Z),$$

where $\mathscr{F}(Z)$ consist of those pairs $(a, b) \in Z$ that satisfy (*atom eq.*) and the (R_i*-back*) and (R_i*-forth*) conditions w.r.t. Z. Locally, over every pair of structures, the bisimulation relation $\sim$ is the greatest fixed point of this operation $\mathscr{F}$ (which is guaranteed to exist since $\mathscr{F}$ is monotone w.r.t. $\subseteq$).

This direct—more static—description of the target equivalence as a greatest fixed point is typical for comparison games of this kind; in the case of bisimulation equivalence the typical back and forth conditions were introduced in the modal world under the name of *zig-zag* conditions by Johan van Benthem. The term *bisimulation* equivalence, which points to an intuition based on the behaviour of transition systems, was introduced by Milner and Park.

A more dynamic view is also extracted from the greatest fixed point characterisation, if we look at the refinement process that recursively generates the fixed point $\sim$ as a limit of relations $\sim^\alpha$:

$$\sim \; = \bigcap_\alpha \sim^\alpha, \quad \text{where}$$

$$\begin{aligned} \sim^0 &= \text{atom equivalence,} \\ \sim^{\alpha+1} &= \mathscr{F}(\sim^\alpha), \\ \sim^\lambda &= \bigcap_{\alpha<\lambda} \sim^\alpha \quad \text{for limit ordinals } \lambda. \end{aligned}$$

Formally, the intersection in the above definition of $\sim$ is over all ordinal levels α, but in restriction to any two concrete structures can be bounded by any infinite ordinal that is of cardinality greater than the structures at hand. Over all finite, and indeed over finitely branching structures and also over the class of all ω-saturated or the class of all modally saturated structures, the limit is reached by stage ω, i.e., coincides with the limit of the finite approximations $\sim^\ell$ for $\ell \in \mathbb{N}$,

$$\sim^\omega = \bigcap\nolimits_{\ell\in\omega} \sim^\ell .$$

Over finite $\mathfrak{A}$ and $\mathfrak{B}$ of sizes $|A|$ and $|B|$, the natural game analysis even shows that full bisimulation is reached no later than by level $\sim^\ell$, where $\ell = \max(|A|, |B|)$.

The game counterpart of $\sim^\ell$ for $\ell \in \mathbb{N}$ is the ℓ-round bisimulation game, which is won by the second player if she does not lose during the first ℓ rounds. Bisimulation equivalence and its infinite game, and especially its finite approximations $\sim^\ell$ for $\ell \in \mathbb{N}$ in relation to the ℓ-round game, can be viewed as a special adaptation to the modal scenario of the classical back&forth games in the Ehrenfeucht–Fraïssé tradition.

We write $\mathfrak{A}, a \equiv^{\ell}_{\mathrm{ML}} \mathfrak{B}, b$ for the modal levels of elementary equivalence up to quantifier rank (modal nesting depth) ℓ: $\mathfrak{A}, a \equiv^{\ell}_{\mathrm{ML}} \mathfrak{B}, b$ if $\mathfrak{A}, a \models \varphi \Leftrightarrow \mathfrak{B}, b \models \varphi$ for all $\varphi \in \mathrm{ML}$ of nesting depth up to ℓ. Similarly, $\mathfrak{A}, a \equiv_{\mathrm{ML}} \mathfrak{B}, b$ stands for full modal equivalence, and $\mathfrak{A}, a \equiv^{\infty}_{\mathrm{ML}} \mathfrak{B}, b$ for equivalence w.r.t. the infinitary variant of modal logic which allows for infinite conjunctions and disjunctions.

Theorem 1.2 (Ehrenfeucht–Fraïssé and Karp theorems for ML) *In restriction to finite modal vocabularies, and for every $\ell \in \mathbb{N}$:*

$$\mathfrak{A}, a \sim^{\ell} \mathfrak{B}, b \quad \textit{if, and only if,} \quad \mathfrak{A}, a \equiv^{\ell}_{\mathrm{ML}} \mathfrak{B}, b.$$

Consequently, in restriction to finite modal vocabularies $\mathfrak{A}, a \sim^{\omega} \mathfrak{B}, b$ if, and only if, $\mathfrak{A}, a \equiv_{\mathrm{ML}} \mathfrak{B}, b$. Without any restriction on the size of the modal vocabulary,

$$\mathfrak{A}, a \sim \mathfrak{B}, b \quad \textit{if, and only if,} \quad \mathfrak{A}, a \equiv^{\infty}_{\mathrm{ML}} \mathfrak{B}, b.$$

Many other logics, and in particular other fragments of first-order logic besides the modal fragment, can be analysed via specifically associated Ehrenfeucht–Fraïssé games. The analysis of the guarded fragment GF of first-order logic in the light of its invariance under guarded bisimulation equivalence is a prime example to be discussed in Sect. 1.3. The very proposal of GF in [1] was inspired by considerations concerning the taming of first-order logic through variations that involve a generalised (or, depending on the point of view: restricted) semantics in 'general assignment models' in the sense of [6]. Returning to our opening remarks about 'behaviour' in terms of logic and games, different logics with distinct semantics may be obtained by admitting different *observable configurations* and different *modes of navigation* between these. (For classical modal semantics, think of possible worlds and accessibility relations.) It is in this view, that games and game graphs provide yet another link to bisimulation as the quintessential notion of behavioural equivalence. Bisimulation as the master game equivalence is adaptable to different logics if, instead of the usual structures, we look at the game graphs induced by the semantic games of those other logics. For suitable logics, the associated game graphs formalise the notion of observable configurations (or admissible assignments) and transitions between these (quantification patterns). Thus, levels of bisimulation equivalence between the associated game graphs correspond to levels of Ehrenfeucht–Fraïssé equivalence between the underlying structures, capturing the specific restrictions embodied in the semantics of the logic in question. Some correspondences of this kind are explored at first-order level in [20], and, with much greater generality in mind, in [6], in the terminology of *general assignment models*. In the same vein, suitable abstractions of the associated game graphs (intuitively akin to filtrations or bisimulation quotients) may serve as concise descriptions of structures up to equivalence, or as blue-prints for desired models (quasi-models) towards decidability and complexity arguments.

1.2.2 Bisimulation in Modal Model Theory

The essential observation for a view of bisimulation equivalences as specialisations of corresponding classical first-order Ehrenfeucht–Fraïssé equivalences is the manner in which its back&forth conditions precisely reflect the power of modal quantification. The existential diamond modality $\Diamond_i$, whose semantics in structure $\mathfrak{A}$ is defined in terms of the accessibility relation $R_i^{\mathfrak{A}}$, precisely captures the available moves in the game along R_i-transitions, and the back&forth clauses for R_i reflect potential distinctions w.r.t. properties of nodes accessible from the current nodes through R_i-edges in their respective structures.

On the other hand, the bisimulation games can be taken as the quintessential template for a large class of model-theoretic Ehrenfeucht–Fraïssé style comparison games: if we correctly abstract from the structures at hand a game graph that models the relevant configurations and transitions between them, then levels of bisimulation equivalence correspond to winning strategies for the second player in a game that reflects the expressive power and quantification pattern of some other target logic [20]. In some key examples, the relevant configurations correspond to the *admissible assignments* to first-order variables, and the transitions to their relative accessibility by means of basic quantification steps. In this vein, variations and especially restrictions to the admissible assignments in a first-order framework lead to fragments that can be analysed and understood in terms of bisimulation equivalences between derived game graphs. Among the most pertinent examples are the k-variable fragments FO^k of first-order logic, and the guarded fragment GF of first-order logic. The finite variable fragments FO^k work with a uniform restriction of assignments to size k. This purely quantitative restriction is contrasted in the seminal paper on the guarded fragment [1] by Andréka, van Benthem and Németi with a qualitative restriction of assignments to clusters that are 'guarded' by some relational hyperedge. The new fragment is proposed with a view to a 'dynamic' bounding of the available assignments—it is 'dynamic' in the sense of a position-dependent restriction familiar from modal logics; yet static in the sense of structural analysis. We shall discuss the guarded fragment and the associated ramification of bisimulation in Sect. 1.3. Before that, let us summarise some key features and uses of ordinary, modal bisimulation equivalence, which account for its pivotal role in modal model theory.

The first is a direct corollary of the modal Ehrenfeucht–Fraïssé theorem. If $\varphi \in \text{ML}$ has modal quantifier depth ℓ, then its semantics is invariant under $\sim^\ell$.

The essential feature of *bisimulation invariance* extends to more powerful logics that share the underlying modal quantification pattern, like the modal μ-calculus.

Corollary 1.3 *The semantics of basic modal logic* ML *is invariant under bisimulation equivalence: for* $\varphi \in \text{ML}$, $\mathfrak{A}, a \sim \mathfrak{B}, b \implies \mathfrak{A}, a \models \varphi \Leftrightarrow \mathfrak{B}, b \models \varphi$.

Bisimulation invariance is *the* model-theoretic hallmark of modal logics; in fact so much so, that modal model theory could be equated with *model theory up to bisimulation equivalence*.

1.2.3 Tree Models and Robust Decidability of Modal Logics

The familiar process of tree-unfolding takes a pointed structure $\mathfrak{A}, a$ to a tree structure $\mathfrak{A}^*_a$ with root a, built on the tree of all R_i-labelled paths from a in $\mathfrak{A}$.

Definition 1.4 Let $\mathfrak{A} = (A, (R_i^{\mathfrak{A}}), (P_j^{\mathfrak{A}}), a)$ be a pointed structure (Kripke structure or transition system). Its *tree unfolding* from a is the tree-like structure $\mathfrak{A}^*_a = (A^*_a, (R_i^{\mathfrak{A}^*}), (P_j^{\mathfrak{A}^*}))$ with root a, where A^*_a is the set of edge-labelled paths of the form $w = (a_0, i_0, a_1, \ldots, a_\ell, i_\ell, a_{\ell+1}, \ldots, a_n)$ where $a_0 = a$, i_ℓ such that $e_\ell = (a_\ell, a_{\ell+1}) \in R_{i_\ell}^{\mathfrak{A}}$, with the natural projection

$$\begin{aligned} \pi : A^*_a &\longrightarrow A \\ (a_0, \ldots, a_n) &\longmapsto a_n; \end{aligned}$$

$(w, w') \in R_i^{\mathfrak{A}^*}$ if w' is an extension of w by one R_i-edge, $w = w\widehat{\ }(i, a')$; and $w \in P_j^{\mathfrak{A}^*}$ if $\pi(w) \in P_j^{\mathfrak{A}}$.

Clearly $\mathfrak{A}^*, a \sim \mathfrak{A}, a$. It follows that any bisimulation invariant logic has the *tree model property*. For the finite-depth approximation $\sim^\ell$ of $\sim$, even the truncation $\mathfrak{A}^\ell_a$ to paths of lengths $n \leqslant \ell$ from a satisfies $\mathfrak{A}^\ell, a \sim^\ell \mathfrak{A}, a$. For finite vocabulary (finitely many R_i and P_j), the equivalence relation $\sim^\ell$ has finite index. Therefore, $\mathfrak{A}^\ell_a$ can be pruned so as to retain at most one sibling of each $\sim^\ell$-type among the immediate children of any node, without affecting $\sim^\ell$-types. For basic modal logic, this pruning yields finite tree models.

Corollary 1.5 *Every satisfiable formula* $\varphi \in \mathrm{ML}$ *(of modal quantifier depth* ℓ*) has a finite tree model (of depth* ℓ*).*

These observations are essential for decidability and complexity results for the satisfiability problem, and for what has been called the robust decidability of modal logics. Indeed, it is not just the basic propositional modal logic ML that is decidable for satisfiability. This property is shared by many extensions of ML to much stronger and practically more relevant logics, including linear or branching time temporal logics such as LTL, CTL, CTL*, dynamic logics of programs such as PDL, Parikh's game logic GL and the modal μ-calculus L_μ, the extension of ML by least and greatest fixed points. While basic modal logic ML can be seen as a fragment of first-order logic, this is not the case for these stronger logics; all of them can express properties based on reachability and on other non-local properties that are not first-order. However, it is easy to see that all these logics can be embedded into monadic second-order logic MSO. Among the extensions of modal logics, the modal μ-calculus occupies a special rôle. It encompasses the other logics mentioned (and many more) and it has a clean and interesting model theory. The modal μ-calculus remains decidable in the presence of backward modalities.

The tree model property provides powerful tools for proving decidability and complexity results and for constructing efficient decision procedures. For a quick proof of decidability one can translate formulae of these logics into monadic second-order formulae and invoke Rabin's famous theorem saying that SωS, the monadic theory of the ω-branching tree, is decidable [27]. However, the complexity of monadic logics on infinite trees (and words) is non-elementary. But recall that the proof of Rabin's Theorem is based on tree automata. A much more practical approach for constructing decision procedures for modal logics avoids the detour through monadic second-order logic and directly applies suitable variants of tree automata to modal logics. The theory of finite automata on trees is very well developed, with many different automata models tailored for specific applications, with efficient algorithms for manipulating automata and for reductions between different models, a good understanding of the complexity of the common reasoning tasks for automata (emptiness problems, word problems etc.), and sophisticated optimisation techniques. The tree model property paves the way to make tree automata applicable to the world of modal logics.

The typical complexity level of satisfiability problems for modal logics is EXPTIME. An exception is the basic modal logic ML for which satisfiability is PSPACE-complete. But the addition of rather modest features to ML, for instance a global modality, push up the complexity to EXPTIME; on the other hand, also rather strong extensions of ML such as the modal μ-calculus and even the modal μ-calculus with backward modalities remain EXPTIME-complete. Such results rely on efficient translations of formulae into, say, alternating tree automata, and the EXPTIME-completeness of the emptiness problem for such automata.

1.2.4 Expressive Completeness

As mentioned above, one of the highlights of modal model theory in this sense is the characterisation of basic modal logic as the bisimulation-invariant fragment of first-order logic.

Theorem 1.6 (van Benthem) *For every first-order formula $\varphi(x)$ in a vocabulary of binary relations R_i and unary predicates P_j as above, the following are equivalent:*

(i) *φ is bisimulation invariant.*
(ii) *φ is logically equivalent to a formula of basic modal logic* ML.

In shorthand notation, FO/$\sim\,\equiv$ ML, where the left-hand side suggestively stands for the (syntactically undecidable) collection of bisimulation invariant first-order formulae.

By no means a direct consequence, not even via the finite model property, but rather yet another striking feature of bisimulation equivalence and of modal logic, the same characterisation holds also in the sense of finite model theory:

$$\mathrm{FO}/\!\sim\;\equiv \mathrm{ML}\quad (\mathrm{FMT}).$$

In its basic form this result is due to Rosen [28]; alternative proofs that yield strengthenings and lend themselves to further generalisations have been presented in [19]. We state a few of these generalisations from [11, 18]. *Global bisimulation equivalence*, $\mathfrak{A}, a \sim_{\forall} \mathfrak{B}, b$, refers to a bisimulation relation in which every $a \in A$ is matched to some $b \in B$ and vice versa; *modal logic with a global modality*, ML[∀], is the extension of basic modal logic ML by a global modality, with the full binary relation as its accessibility relation. A *rooted structure* is a structure $\mathfrak{A}, a$ with a single binary accessibility relation R such that every node is reachable on a directed R-path from the root a. *Equivalence structures* are structures that interpret all the binary relations R_i as equivalence relations ($S5$ models).

Theorem 1.7 *Bisimulation invariant fragments of first-order logic are captured by modal logics over some classes of structures, as follows.*

(i) FO/$\sim \equiv$ ML *over the class of all (finite) structures.*
(ii) FO/$\sim_{\forall} \equiv$ ML[∀] *over the class of all (finite) structures.*
(iii) FO/$\sim \equiv$ ML *over the class of all (finite) equivalence structures.*
(iv) FO/$\sim \equiv$ ML[∀] *over the class of all finite rooted structures.*
(iv) FO/$\sim \equiv$ ML *over the class of all finite irreflexive transitive trees.*

Here (i) is the van Benthem–Rosen characterisation from [5] and [28], respectively; the rest are due to [11, 18].

Several of the finite model theory results above make use of *finite* unfoldings of finite structures that produce *locally* tree-like and fully bisimilar finite models—which is not achievable by tree unfoldings since any globally acyclic bisimilar companion of any cyclic structure is necessarily infinite. Simple combinatorial constructions of finite locally acyclic bisimilar covers of finite graphs for this purpose are presented in [18]. They play a crucial role in the analysis of the expressiveness of first-order formulae that are bisimulation invariant over finite structures. Locally acyclic behaviour suffices due to Gaifman's locality theorem: the semantics of any first-order formula $\varphi(x)$ only depends on certain global multiplicities and the local neighbourhood around x; up to bisimulation, global multiplicities (Gaifman's basic local sentences) can be adjusted comparatively easily even when working in special classes of finite models; what remains is the necessity to control the local neighbourhoods and this is where local tree-likeness is useful.

The van Benthem characterisation of bisimulation invariant first-order logic, as FO/$\sim \equiv$ ML, also has a an exciting extension to its monadic second-order counterpart:

Theorem 1.8 (Janin–Walukiewicz) MSO/$\sim \equiv \mathrm{L}_\mu$, *i.e., for every monadic second-order formula $\varphi(x)$ in a vocabulary of binary relations R_i and unary predicates P_j, the following are equivalent:*

(i) *φ is bisimulation invariant.*
(ii) *φ is logically equivalent to a formula of the μ-calculus* L_μ.

Whether this characterisation holds in the sense of finite model theory, remains one of the great challenges in modal model theory.

1.3 Guarded Bisimulation: A Systematic Lifting to Higher Dimension

The 'dynamic' behaviour of modal logics w.r.t. locally available transitions between single-node assignments is vastly generalised in the setting of guarded logics.

The generalisation manifests itself on various levels: as a liberalisation in the relational type of structures (from graph-like transition systems to relational structures with relations of any arity); a generalisation w.r.t. the restrictions on admissible assignments and quantification patterns (from modal $\Box$ and $\Diamond$ to universal and existential quantification over guarded tuples); a generalisation w.r.t. the relevant notion of bisimulation (from modal to guarded bisimulation); and, in the wake of these generalisations, a shift form graph theory to hypergraph theory as the underlying combinatorial framework.

1.3.1 Guardedness and the Guarded Fragment

With a relational structure $\mathfrak{A} = (A, (R_i^{\mathfrak{A}})_{i \in I})$ with relation symbols R_i of arity r_i, we associate a hypergraph of guarded sets, and a notion of guarded tuples as follows. It will be convenient to use the notation $[\mathbf{a}] := \{a_1, \dots, a_k\}$ to denote the set of components of the tuple $\mathbf{a} = (a_1, \dots, a_k) \in A^k$.

Definition 1.9 A subset $s \subseteq A$ is *guarded* in $\mathfrak{A}$ if s is a singleton set or if there is some tuple $\mathbf{a} \in R_i^{\mathfrak{A}}$ for one of the R_i such that $s \subseteq [\mathbf{a}]$. The *hypergraph of guarded sets* of $\mathfrak{A}$ is the hypergraph $H(\mathfrak{A}) := (A, S[\mathfrak{A}])$ with the set S of all guarded subsets of $\mathfrak{A}$ as the set of hyperedges. A tuple $\mathbf{a} \in A^k$ is a *guarded tuple* if $[\mathbf{a}] \in S(\mathfrak{A})$.

The guarded fragment of first-order logic essentially restricts the relevant assignments of first-order variables to guarded tuples. The actual definition is in terms of the restriction of all quantification by means of an explicit relativisation to some guarded tuples. It thus allows only outermost free variables to be instantiated by unguarded assignments, but for many purposes this does not matter (since outer boolean combinations could be treated separately).

Definition 1.10 For arbitrary relational vocabularies, the guarded fragment GF $\subseteq$ FO is the syntactic fragment of FO generated from atomic formulae by the boolean connectives and quantifications of the form

$$\forall \mathbf{y}\big(\alpha(\mathbf{xy}) \rightarrow \varphi(\mathbf{xy})\big), \quad \text{and, dually,} \quad \exists \mathbf{y}\big(\alpha(\mathbf{xy}) \wedge \varphi(\mathbf{xy})\big),$$

where $\varphi(\mathbf{xy}) \in$ GF has free variables among those listed in $\mathbf{xy}$ and $\alpha(\mathbf{xy})$ is an atomic formula in which all the listed variables occur. The formula α is called the *guard* of this quantification.[1] The semantics of GF is that of FO.

[1] If $\mathbf{xy}$ consists of a single variable symbol z, α can be the equality $z = z$.

The definition generalises the relativised quantification of modal logic, so that it is clear that, w.r.t. expressiveness, ML $\subseteq$ GF $\subseteq$ FO, and in fact even the extension of basic modal logic by global and backward modalities is naturally covered by GF.

1.3.2 Guarded Bisimulation and Model Theory

Just as the model theory of modal logics is governed by (modal) bisimulation equivalence, the nice model-theoretic properties of the guarded fragment are closely related to its invariance under guarded bisimulation equivalence. Guarded bisimulation equivalence $\sim_{\mathsf{g}}$ and its finite approximations $\sim_{\mathsf{g}}^{\ell}$ exactly cover the same station for GF as do $\sim$ and $\sim^{\ell}$ for ML—also w.r.t. their nature as the appropriate specialisations of the first-order framework of back&forth games to the quantification pattern of GF.

The positions of the guarded bisimulation game on structures $\mathfrak{A}$ and $\mathfrak{B}$ are partial isomorphisms between A and B whose domain and image are guarded sets[2]; we use a tuple-based notation $p\colon \mathbf{a} \mapsto \mathbf{b}$ to indicate a partial map from A to B with domain $[\mathbf{a}]$ and image $[\mathbf{b}]$ where $b_i = p(a_i)$. One may also think of a placement of matched pebbles on $\mathbf{a}$ and $\mathbf{b}$; the requirements are that $\mathbf{a}$ and $\mathbf{b}$ are guarded and that $p\colon \mathfrak{A} \restriction [\mathbf{a}] \simeq \mathfrak{B} \restriction [\mathbf{b}]$ is an isomorphism of induced substructures (p a partial isomorphism, $\mathbf{a}$ and $\mathbf{b}$ atom equivalent). Then the available moves for the first player, e.g. on the $\mathfrak{A}$-side, are to guarded tuples $\mathbf{a}'$ together with some specified sub-tuple $\mathbf{a}_0$ of both $\mathbf{a}$ and $\mathbf{a}'$ that stay put—and the response by the second player needs to keep the sub-tuple $\mathbf{b}_0 := p(\mathbf{a}_0)$ fixed and produce an extension $\mathbf{b}'$ such that the new $p'\colon \mathbf{a}' \mapsto \mathbf{b}'$ is again a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

An alternative set-based view has partial isomorphisms between guarded subsets as the positions; the moves correspond to transitions from one guarded subset to another, with a specified (possible empty) subset of their intersection to be respected by the second player's response. This view highlights the hypergraph-theoretic nature, and indeed can be cast as a notion of *hypergraph bisimulation* that additionally needs to respect relational content.

Definition 1.11 For two relational structures $\mathfrak{A}$ and $\mathfrak{B}$ (of the same vocabulary), a set of partial maps Z between $\mathfrak{A}$ and $\mathfrak{B}$ is a *guarded bisimulation* if it satisfies the following, for every $p\colon \mathbf{a} \mapsto \mathbf{b}$ in Z:

(i) (*atom eq.*): $p\colon \mathfrak{A} \restriction \mathbf{a} \simeq \mathfrak{B} \restriction \mathbf{b}$ is a partial isomorphism;
(ii) (*back*): for every guarded tuple $\mathbf{b}'$ of $\mathfrak{B}$ and $\mathbf{b}_0$ with $[\mathbf{b}_0] \subseteq [\mathbf{b}] \cap [\mathbf{b}']$, there is some guarded tuple $\mathbf{a}'$ of $\mathfrak{A}$ and $p'\colon \mathbf{a}' \mapsto \mathbf{b}'$ in Z such that $p'^{-1}(\mathbf{b}_0) = p^{-1}(\mathbf{b}_0)$;
(iii) (*forth*): for every guarded tuple $\mathbf{a}'$ of $\mathfrak{A}$ and $\mathbf{a}_0$ with $[\mathbf{a}_0] \subseteq [\mathbf{a}] \cap [\mathbf{a}']$, there is some guarded tuple $\mathbf{b}'$ of $\mathfrak{B}$ and $p'\colon \mathbf{a}' \mapsto \mathbf{b}'$ in Z such that $p'(\mathbf{a}_0) = p(\mathbf{a}_0)$.

[2] One should except the initial position from the guardedness requirement in order to match the liberal treatment of (outermost) free variables in GF.

We write $\mathfrak{A}, \mathbf{a} \sim_{\mathsf{g}} \mathfrak{B}, \mathbf{b}$ if there is a guarded bisimulation Z containing $p\colon \mathbf{a} \mapsto \mathbf{b}$. Finite approximations $\sim_{\mathsf{g}}^{\ell}$ are introduced in complete analogy with the modal $\sim$ and $\sim^{\ell}$, and similarly correspond to the existence of winning strategies for ℓ rounds in the guarded bisimulation game. As in the modal case, we introduce $\sim_{\mathsf{g}}^{\omega}$ as the common refinement of the finite levels $\sim_{\mathsf{g}}^{\ell}$.

One obtains the natural variant of the first-order Ehrenfeucht–Fraïssé and Karp theorems for GF. The equivalence relations $\equiv_{\mathrm{GF}}^{\ell}$ and $\equiv_{\mathrm{GF}}$ are introduced as levels of elementary equivalence in GF, where the ℓ in $\equiv_{\mathrm{GF}}^{\ell}$ refers to the nesting depth of guarded quantification (which is typically lower than the first-order quantifier rank, as guarded quantification may quantify over tuples in a single step). The relation $\equiv_{\mathrm{GF}}^{\infty}$ similarly denotes equivalence w.r.t. the infinitary variant of GF, with infinite disjunctions and conjunctions.

Theorem 1.12 (Ehrenfeucht–Fraïssé and Karp theorems for GF) *In restriction to finite relational vocabularies, and for every $\ell \in \mathbb{N}$:*

$$\mathfrak{A}, \mathbf{a} \sim_{\mathsf{g}}^{\ell} \mathfrak{B}, \mathbf{b} \quad \textit{if, and only if,} \quad \mathfrak{A}, \mathbf{a} \equiv_{\mathrm{GF}}^{\ell} \mathfrak{B}, \mathbf{b}.$$

Consequently, in restriction to finite vocabularies $\mathfrak{A}, \mathbf{a} \sim_{\mathsf{g}}^{\omega} \mathfrak{B}, \mathbf{b}$ *if, and only if,* $\mathfrak{A}, \mathbf{a} \equiv_{\mathrm{GF}} \mathfrak{B}, \mathbf{b}$. *Without any restriction on the size of the vocabulary,*

$$\mathfrak{A}, \mathbf{a} \sim_{\mathsf{g}} \mathfrak{B}, \mathbf{b} \quad \textit{if, and only if,} \quad \mathfrak{A}, \mathbf{a} \equiv_{\mathrm{GF}}^{\infty} \mathfrak{B}, \mathbf{b}.$$

1.3.3 Guarded Bisimulation Invariance

The following is an immediate consequence of the guarded Ehrenfeucht–Fraïssé theorem.

Corollary 1.13 *The semantics of $\varphi \in$ GF is invariant under $\sim_{\mathsf{g}}$.*

The expressive completeness assertion in the following characterisation theorem of Andréka–van Benthem–Németi rests on a non-trivial but canonical classical proof by means of compactness and saturation. It provides a beautiful analogue and generalisation of van Benthem's semantic characterisation of ML $\subseteq$ FO, Theorem 1.6.

Theorem 1.14 (Andréka–van Benthem–Németi) *The guarded fragment is semantically characterised as a fragment of first-order logic by its invariance under guarded bisimulation equivalence:* FO$/{\sim_{\mathsf{g}}} \equiv$ GF. *In more detail, for every first-order formula $\varphi(\mathbf{x})$ in a relational vocabulary, the following are equivalent:*

(i) φ is invariant under guarded bisimulation.
(ii) φ is logically equivalent to a formula of GF.

Moreover, a guarded analogue of the Janin–Walukiewicz Theorem (Theorem 1.8) can also be obtained via a natural translation between the guarded and modal worlds. The logics involved are the following: *guarded second-order logic* GSO, which here takes the place of MSO, is the natural restriction of second-order logic that allows to quantify over *sets of guarded tuples*; *guarded fixpoint logic* μGF is the extension of GF by constructors for least and greatest fixed points.

Theorem 1.15 (Grädel–Hirsch–Otto) GSO/$\sim_{\mathsf{g}} \;\equiv\; \mu$GF*, i.e., For every* GSO-*formula* $\varphi(\mathbf{x})$*, the following are equivalent:*

(i) φ *is invariant under guarded bisimulation equivalence.*
(ii) φ *is logically equivalent to a formula of* μGF.

The translations in [14] that directly reduce this assertion to Theorem 1.8 involve an interesting parallelism between modal and guarded tree unfoldings.

Guarded tree unfoldings of relational structures $\mathfrak{A} = (A, (R^{\mathfrak{A}}))$ can be constructed from a tree unfolding of the associated transition system $I(\mathfrak{A}) = (S[\mathfrak{A}] \cup \{\emptyset\}, E)$ where $S[\mathfrak{A}]$ is the set of guarded subsets of $\mathfrak{A}$ and $E = \{(s, s') \colon s \neq s', s = \emptyset \text{ or } s \cap s' \neq \emptyset\}$.[3] From a tree unfolding $I^* := I^*_\emptyset$ of $I(\mathfrak{A})$ from the root node $\emptyset$, with natural projection $\pi \colon I^* \to S(\mathfrak{A}) \cup \{\emptyset\}$ we reconstruct a relational structure

$$\hat{\mathfrak{A}} = (\hat{A}, (R^{\hat{\mathfrak{A}}}))$$

as follows. The universe $\hat{A}$ is the quotient of the disjoint union of copies of sets $\pi(\hat{s}) \subseteq A$,

$$\bigcup_{\hat{s} \in I^*} \{\hat{s}\} \times \pi(\hat{s})$$

w.r.t. the equivalence relation that identifies $a \in \pi(\hat{s}_1)$ with $a \in \pi(\hat{s}_2)$ if, and only if, $\hat{s}_2$ and $\hat{s}_1$ are connected in $I^*_\emptyset$ by a path whose π-projection involves just edges $e = (s, s') \in E$ for which $a \in s \cap s'$. We denote the equivalence class of $(\hat{s}, a)$ for $a \in \pi(\hat{s})$ by $[\hat{s}, a]$, and the set $\{[\hat{s}, a] \colon a \in \pi(\hat{s})\} \subseteq \hat{A}$ by $[\hat{s}]$. The map that sends the equivalence class $[\hat{s}, a]$ of $a \in \hat{s}$ to $a \in A$ is the natural projection associated with the unfolding, for simplicity also denoted $\pi \colon \hat{A} \to A$. Locally, in restriction to every $[\hat{s}] \subseteq \hat{A}$, this projection π is a bijection onto the corresponding guarded subset $s = \pi(\hat{s})$ of $\mathfrak{A}$. Relations R are interpreted in $\hat{\mathfrak{A}}$ such that precisely the sets $[\hat{s}] \subseteq \hat{A}$ are guarded subsets, and such that $\pi \colon \hat{\mathfrak{A}} \to \mathfrak{A}$ is a global relational homomorphism and a local isomorphism in restriction to every subset $[\hat{s}]$.

Definition 1.16 The *guarded tree unfolding* of a relational structure $\mathfrak{A} = (A, (R^{\mathfrak{A}}))$ is the structure $\hat{\mathfrak{A}} = (\hat{A}, (R^{\hat{\mathfrak{A}}}))$ as constructed from a tree unfolding of the intersection graph $I(\mathfrak{A})$ above, together with the natural homomorphic projection $\pi \colon \hat{\mathfrak{A}} \to \mathfrak{A}$, which bijectively associates the guarded subsets $[\hat{s}] \in S(\hat{\mathfrak{A}})$ with their underlying guarded subsets $s = \pi(\hat{s}) \in S[\mathfrak{A}]$.

[3] We attach the empty set as a root to $I(\mathfrak{A})$ and join it to every guarded set to obtain a natural tree unfolding for our purposes, rather than a forest.

It is straightforward to check that the restrictions of the projection homomorphism $\pi : \hat{\mathfrak{A}} \to \mathfrak{A}$ to the guarded subsets of $\hat{\mathfrak{A}}$ form a guarded bisimulation. Therefore, for any guarded subset $[\hat{s}]$ of $\hat{\mathfrak{A}}$ above the guarded subset $s = \pi(\hat{s})$ of $\mathfrak{A}$,

$$\hat{\mathfrak{A}}, [\hat{s}] \sim_{\mathsf{g}} \mathfrak{A}, s,$$

where we allow ourselves to write just the guarded sets $[\hat{s}]$ and s, instead of π-compatible listings of their elements as tuples.

Tree unfoldings as just defined are tree-like also in the sense that their hypergraphs of guarded subsets $S[\hat{\mathfrak{A}}]$ are *acyclic*. There are several equivalent characterisations of the relevant notion of hypergraph acyclicity (also called α-acyclicity in the literature, cf. [4, 9]): in terms of *tree decompositions* that use guarded subsets (hyperedges) as bags; in terms of reducibility by means of reduction steps that allow for

(i) removal of a vertex (from the universe and every hyperedge) provided it is contained in at most one hyperedge, and
(ii) retraction of a hyperedge provided it is fully contained in some other hyperedge;

and in terms of the local criteria of *conformality* and *chordality* for the hypergraph and its associated Gaifman graph.

Definition 1.17 For a hypergraph $H = (A, S)$, define the associated *Gaifman graph* $G(H)$ to have vertex set A and an edge between distinct $a, a' \in A$ precisely if a and a' occur together in some hyperedge $s \in S$.

The hypergraph $H = (A, S)$ is *acyclic* if it is both

(i) *conformal*: each clique in $G(H)$ is contained in a single hyperedge, and
(ii) *chordal*: every cycle in $G(H)$ of length greater than 3 has a chord, i.e., $G(H)$ has no induced subgraphs isomorphic to the k-cycle for $k > 3$.

Since every relational structure $\mathfrak{A}$ is guarded bisimulation equivalent to its guarded tree unfolding, and as GF is invariant under guarded bisimulation equivalence, we find that every satisfiable formula of GF has an acyclic model. This was first stated in [12] as the *generalised tree model property* of GF.

Corollary 1.18 (Grädel) *Every logic that is invariant under guarded bisimulation equivalence has this generalised tree model property: every satisfiable formula has a model whose hypergraph of guarded subsets is acyclic, i.e., a model that admits a tree-decomposition with guarded subsets as bags.*

For a relational vocabulary of width w, this further entails that every satisfiable formula of GF or μGF has a (countable) model of tree width $w - 1$.

1.3.4 Decidability and Complexity for GF and Its Extensions

As in the case of modal logics, the tree model property for guarded models paves the way to decidability and automata based decision procedures. These do not only work for the guarded fragment GF in its basic form, but also for guarded fixed-point logic μGF and for other variants of guarded logics based on more liberal notions of guarded sets.

Indeed, structures of bounded tree width can be uniformly represented by standard trees, in the graph-theoretic sense, with a bounded set of labels. More precisely, given a tree decomposition of width $k-1$ of a relational τ-structure $\mathfrak{D}$ we fix a set K of $2k$ constants and assign to every element $d \in \mathfrak{D}$ a constant $a_d \in K$ such that distinct elements living at adjacent nodes in the tree decomposition are represented by distinct constants. On the tree T underlying the decomposition of $\mathfrak{D}$ we define monadic predicates $\mathcal{O}_a$ (for $a \in K$) and $R_{\mathbf{a}}$ (for m-ary $R \in \tau$ and $\mathbf{a} \in K^m$) where $\mathcal{O}_a$ is true at those nodes of T where an element represented by a occurs, and $R_{\mathbf{a}}$ is the set of nodes of T where a tuple $(d_1, \ldots, d_m) \in R$ occurs that is represented by $\mathbf{a}$. We thus obtain a tree structure $T(\mathfrak{D})$ which has (beyond the edge relation of the tree) only monadic predicates and which carries all structural information about $\mathfrak{D}$ and its tree decomposition.

On the other hand, a tree T with such monadic relations $\mathcal{O}_a$ and $R_{\mathbf{a}}$ is indeed a tree representation $T(\mathfrak{D})$ for some τ-structure $\mathfrak{D}$ if, and only if, it satisfies certain consistency axioms that turn out to be first-order definable.

There are several options to exploit this for proving decidability and complexity results. The simplest way to prove decidability of guarded fixed-point logic μGF is by an interpretation into SωS, the monadic logic of the countable branching tree. That is, with every formula $\varphi(x_1, \ldots x_m)$ of μGF and every tuple $\mathbf{a} \in K^m$ one can associate a monadic second-order formula $\psi_{\mathbf{a}}(z)$ that describes on the tree structure $\mathcal{T}(\mathfrak{D})$ the same properties of *guarded* tuples that $\varphi(\bar{x})$ does on $\mathfrak{D}$, in the following sense: if $\mathbf{d}$ is a guarded tuple of $\mathfrak{D}$ living at node v of the tree T, and if $\mathbf{a}$ represents $\mathbf{d}$ at v, then

$$\mathfrak{D} \models \varphi(\mathbf{d}) \Longleftrightarrow T(\mathfrak{D}) \models \psi_{\mathbf{a}}(v).$$

On the basis of this translation and of the facts that the consistency axioms for tree representations are first-order, that μGF (and least fixed point logic in general) has the Löwenheim-Skolem property, and that the monadic theory of countable trees is decidable, it is then not difficult to prove that the satisfiability problem for μGF is decidable.

Instead of the reduction to the monadic second-order theory of trees, one can define a similar reduction to the modal μ-calculus with backward modalities. The decidability (and EXPTIME-complexity) of this logic has been established by Vardi [30] by means of two-way alternating automata. To make such a reduction work, one has to observe that the consistency axioms for tree representations can be formulated in this logic (in fact, it is sufficient to use basic modal logic with a global modality and backward modalities) and that least and greatest fixed points in μGF on $\mathfrak{D}$ can be encoded by *simultaneous* modal fixed-point formulae on $T(\mathfrak{D})$.

It should be pointed out that the usual modal μ-calculus, without backward modalities, does not seem to be sufficient for such an approach. Indeed, besides the tree model property, the modal μ-calculus also has the finite model property, while one easily obtains formulae that have only infinite models in μGF and in the μ-calculus with backward modalities.

Finally, the satisfiability problem for guarded fixed-point logic can also be solved by direct application of suitably tailored automata-theoretic methods. The general idea is to associate with every sentence $\psi \in \mu$GF an alternating tree automaton $\mathscr{A}_\psi$ that accepts precisely the (tree descriptions of the) like-tree models of ψ. This reduces the satisfiability problem of ψ to the emptiness problem of the automaton, a problem that is solvable in exponential time with respect to the number of states of the automaton. This was the approach taken in [15] where the decidability of μGF had first been established. Instead of Vardi's two-way automata, Grädel and Walukiewicz use a different variant of alternating automata that work on trees of arbitrary, finite or infinite, degree and do not make use of the orientation of edges. The behaviour of such an automaton on a given tree structure is described by a parity game, and by means of the positional determinacy of these games one can reduce the input trees to trees of bounded branching (and the automata to those used by Vardi for the decidability of the μ-calculus with backward modalities). The size of the automaton $\mathscr{A}_\psi$ is bounded by $|\psi|^{2k \log k}$ where k is the width of ψ. For the following see [15].

Theorem 1.19 (Grädel–Walukiewicz) *The satisfiability problem for* μGF *is decidable, and complete for* 2EXPTIME. *For* μGF-*sentences of bounded width the satisfiability problem is* EXPTIME-*complete.*

It is worth pointing out that the same complexity bounds also hold for GF, the guarded fragment without fixed points [12]. The double exponential complexity of GF and μGF may seem high (and disappointing for practical applications). However, it is not really surprising, since these logics admit predicates of unbounded arity (whereas modal logics are evaluated on graph-like structures). Even a single predicate of arity n on a universe with just two elements admits 2^{2^n} types already at the atomic level, so one cannot really expect lower complexity bounds. In many practical applications, the underlying vocabulary will be fixed and the arity therefore bounded. In such cases the satisfiability problems for GF and μGF are in EXPTIME and thus on the same level as for most modal logics.

Beyond GF and μGF the general approach outlined here also works for other, more general, notions of guarded logics based on more liberal definitions of guardedness. This includes *loosely guarded*, *packed*, or *clique-guarded* logics. While the classical notion of a guarded set means that the entire set is covered by one atomic fact, the most liberal notion, of a clique-guarded set, just requires that any two elements of the set coexist in some atomic fact, which means that the set is a clique in the Gaifman graph of the structure. Most of the algorithmic results on GF and μGF can be extended to the clique-guarded extensions CGF and μCGF (with appropriate modifications, in particular for the notion of bisimulation). For details, see [13].

1.3.5 Guarded Model Constructions

Guarded tree unfoldings provide one example of a specific form of model construction, or in this case: model transformation, that is tailored for the model theoretic analysis of guarded logics. The requirements of acyclicity and finiteness will in general be incompatible; we shall return to the interesting question how much acyclicity can in general be achieved in finite models further below. For a start, however, we consider the finite model property for the guarded fragment, disregarding the issue of acyclicity. The following proof idea stems from [12] and uses a nice combinatorial result, about finite extension properties of partial isomorphisms due to Herwig [16].

Theorem 1.20 (Herwig) *Any finite relational structure* $\mathfrak{A}$ *admits a finite extension* $\bar{\mathfrak{A}} \supseteq \mathfrak{A}$ *(*$\mathfrak{A}$ *becomes an induced substructure of* $\bar{\mathfrak{A}}$*) with the property that every partial isomorphism* $p\colon \mathfrak{A}\restriction \mathrm{dom}(p) \simeq \mathfrak{A}\restriction \mathrm{image}(p)$ *extends to (is induced by) an automorphism* $\bar{p}$ *of* $\bar{\mathfrak{A}}$*.*

It is easy to see that any Herwig extension $\bar{\mathfrak{A}}$ of $\mathfrak{A}$ can be thinned out so that each $R^{\bar{\mathfrak{A}}}$ is generated by the orbit of $R^{\mathfrak{A}}$ under the automorphism group. Let us call such a Herwig extension *special*.

Special Herwig extensions of sufficiently rich finite substructures $\mathfrak{A} \subseteq \mathfrak{B}$ are $\sim^{\ell}_{\mathsf{g}}$-equivalent to $\mathfrak{B}$ itself; this is the core of the finite model property for GF as proved in [12], see Theorem 1.22.

Lemma 1.21 *Let* $\mathfrak{B}$ *be a relational structure,* $\mathfrak{A} = \mathfrak{B}\restriction A$ *an induced finite substructure on a subset* $A \subseteq B$ *that is sufficiently rich to contain, for every guarded tuple* $\mathbf{b}$ *of* $\mathfrak{B}$*, at least one realisation of that* $\sim^{\ell}_{\mathsf{g}}$*-type: there is* $\mathbf{a} \in \mathfrak{A}$ *such that* $\mathfrak{B}, \mathbf{a} \sim^{\ell}_{\mathsf{g}} \mathfrak{B}, \mathbf{b}$*. Then any special Herwig extension* $\bar{\mathfrak{A}} \supseteq \mathfrak{A}$ *is* $\sim^{\ell}_{\mathsf{g}}$*-equivalent to* $\mathfrak{B}$ *in the sense that*

(i) $\bar{\mathfrak{A}} \sim^{\ell}_{\mathsf{g}} \mathfrak{B}$*;*
(ii) for every guarded tuple $\mathbf{a} \in \mathfrak{A}$*:* $\bar{\mathfrak{A}}, \mathbf{a} \sim^{\ell}_{\mathsf{g}} \mathfrak{B}, \mathbf{a}$.

Proof Using the fact that every guarded tuple in $\bar{\mathfrak{A}}$ is in the orbit of some guarded tuple $\mathbf{a}$ of $\mathfrak{A}$ under an automorphism of $\bar{\mathfrak{A}}$ (because $\bar{\mathfrak{A}}$ is special), and that, up to $\sim^{\ell}_{\mathsf{g}}$, every guarded tuple of $\mathfrak{B}$ is represented in $A \subseteq B$, claim (i) directly follows from claim (ii). For claim (ii) it essentially suffices to observe that every back&forth requirement for $\mathbf{a}$ that can be met in $\mathfrak{B}$ can also be met in $\bar{\mathfrak{A}}$, as follows.

Let $\mathbf{a} \in \mathfrak{A}$ be guarded, $\mathbf{b}$ guarded in $\mathfrak{B}$, and $\mathbf{c}$ a tuple in the intersection $[\mathbf{a}] \cap [\mathbf{b}]$. By the richness assumption on A, there is some $\mathbf{a}' \in \mathfrak{A}$ such that $\mathfrak{B}, \mathbf{a}' \sim^{\ell}_{\mathsf{g}} \mathfrak{B}, \mathbf{b}$. This implies in particular that the tuple $\mathbf{c}'$ in $[\mathbf{a}']$ corresponding to $\mathbf{c}$ in $[\mathbf{a}] \cap [\mathbf{b}]$ is linked to $\mathbf{c}$ by a partial isomorphism p of $\mathfrak{A}$. The automorphism $\bar{p}$ of $\bar{\mathfrak{A}}$ then shows that $\bar{p}(\mathbf{a}')$ overlaps with $\mathbf{a}$ in the tuple $\mathbf{c}$ in $\bar{\mathfrak{A}}$ (just as $\mathbf{b}$ overlaps with $\mathbf{a}$ in $\mathbf{c}$ in $\mathfrak{B}$). By induction on ℓ for claim (ii), i.e. assuming claim (ii) at level $\ell - 1$, we find

$$\bar{\mathfrak{A}}, \bar{p}(\mathbf{a}') \simeq \bar{\mathfrak{A}}, \mathbf{a}' \sim^{\ell-1}_{\mathsf{g}} \mathfrak{B}, \mathbf{a}' \sim^{\ell}_{\mathsf{g}} \mathfrak{B}, \mathbf{b}.$$

This, for all available $\mathbf{b}$ in $\mathfrak{B}$, shows that $\bar{\mathfrak{A}}, \mathbf{a} \sim_{\mathsf{g}}^{\ell} \mathfrak{B}, \mathbf{a}$ as required for (ii) at level ℓ. □

Claim (i) of the lemma directly yields the finite model property for GF, since any $\varphi \in \mathrm{GF}$ of nesting depth ℓ is preserved under $\sim_{\mathsf{g}}^{\ell}$, and since every $\mathfrak{B}, \mathbf{b} \models \varphi$ has a finite substructure $\mathfrak{A} \subseteq \mathfrak{B}$ that contains at least one realisation of each one of the finitely many $\sim_{\mathsf{g}}^{\ell}$-types realised by guarded tuples of $\mathfrak{B}$.

Theorem 1.22 (Grädel) GF *has the finite model property: every satisfiable* $\varphi \in \mathrm{GF}$ *has a finite model.*

Better bounds on the size of small models for a given satisfiable $\varphi \in \mathrm{GF}$ are obtained by a more recent construction in [2], which builds a small model not directly from a given (infinite) model, but from a complete abstract description of the required $\sim_{\mathsf{g}}^{\ell}$-type to be realised.

Proposition 1.23 (Bárány–Gottlob–Otto) *Every satisfiable* $\varphi \in \mathrm{GF}(\sigma)$*, where* σ *is any relational vocabulary of width* w*, has a small finite model whose size can be bounded exponentially in the length of* φ*, for fixed* w*; the dependence on* w*, on the other hand, is doubly exponential.*

The core construction of [2], of which the above really is a technical corollary, yields finite guarded bisimilar covers that are weakly N-acyclic in the sense of the following definitions.

Definition 1.24 A *guarded bisimilar covering* of a relational structure $\mathfrak{A}$ is a homomorphism $\pi : \hat{\mathfrak{A}} \to \mathfrak{A}$ from some relational structure $\hat{\mathfrak{A}}$ (the cover) onto $\mathfrak{A}$, such that the restrictions of π to guarded subsets of $\hat{\mathfrak{A}}$ induce a guarded bisimulation.

Guarded tree unfoldings are natural examples in point; however, we are here mostly interested in coverings of finite $\mathfrak{A}$ by *finite covers* $\hat{\mathfrak{A}}$. The restrictions of the cover homomorphism π to guarded subsets must in particular be partial isomorphisms. The *forth*-property is thus subsumed in the requirement that π is a homomorphism. The *back*-property corresponds to a lifting property familiar from topological or geometric notions of coverings.[4]

Guarded tree unfoldings provide fully acyclic coverings, albeit infinite ones. One useful approximation to acyclicity in finite covers is the following from [2].

Definition 1.25 A covering $\pi : \hat{\mathfrak{A}} \to \mathfrak{A}$ is *weakly* N*-acyclic* if every induced substructure of $\hat{\mathfrak{A}}$ of size up to N is tree-decomposable with bags that project onto guarded subsets of $\mathfrak{A}$ under π.

Proposition 1.26 (Bárány–Gottlob–Otto) *For every* $N \in \mathbb{N}$*, each finite relational* $\mathfrak{A}$ *admits weakly* N*-acyclic coverings by finite structures.*

---

[4] It may be worth to point out that, unlike the finite bisimilar coverings obtained for graph-like structures in [18], the bisimilar coverings of relational structures or of hypergraphs will necessarily be *branched coverings*, and do not provide unique liftings.

An analysis of homomorphisms $h\colon \mathfrak{C} \to \hat{\mathfrak{A}}$, from structures $\mathfrak{C}$ of size up to N into a weakly acyclic cover $\pi\colon \hat{\mathfrak{A}} \to \mathfrak{A}$, shows that $\mathfrak{A}$ must satisfy one of a finite list of potential GF-descriptions of all possible acyclic homomorphic images of $\mathfrak{C}$.[5] If $\mathfrak{A}$ does not satisfy this GF-expressible finite 'disjunction of acyclic conjunctive queries', then $\hat{\mathfrak{A}}$ cannot even admit cyclic homomorphic images of $\mathfrak{C}$. Together with existence of finite, weakly N-acyclic covers, this argument from [2] yields a considerable strengthening of the finite model property for GF, as well as natural applications to database issues regarding conjunctive queries under GF-definable constraints.

For the following, a class $\mathscr{C}$ of σ-structures is said to be defined in terms of *finitely many forbidden homomorphisms* if, for some finite list of finite σ-structures $\mathfrak{C}_1, \ldots, \mathfrak{C}_m$, the class $\mathscr{C}$ consists of precisely those σ-structures $\mathfrak{C}$ that admit no homomorphisms $h\colon \mathfrak{C}_i \to \mathfrak{C}$ for $1 \leqslant i \leqslant m$.

Corollary 1.27 (Bárány–Gottlob–Otto) GF *has the finite model property in restriction to any class $\mathscr{C}$ of relational structures that is defined in terms of finitely many forbidden homomorphisms: for any such class $\mathscr{C}$, φ has a model in $\mathscr{C}$ if, and only if, it has a finite model in $\mathscr{C}$.*

Interestingly, this strengthening of the finite model property for GF can also be obtained from a corresponding strengthening of Herwig's theorem. We briefly present this new alternative proof from [24], which may be of independent systematic interest.[6] The Herwig–Lascar theorem [17] asserts a finite model property for the extension task for partial isomorphisms over classes with finitely many forbidden homomorphisms. An alternative proof of the Herwig–Lascar theorem itself, which is inspired by hypergraph constructions related to the exploration of the finite model theory of GF, see Sect. 1.3.6, can be found in [22, 24, 25].

Theorem 1.28 (Herwig–Lascar) *Let the class of relational structures $\mathscr{C}$ be defined in terms of finitely many forbidden homomorphisms. Suppose that a finite structure $\mathfrak{A} \in \mathscr{C}$ has a possibly infinite extension $\mathfrak{B} \supseteq \mathfrak{A}$ in $\mathscr{C}$ that extends every partial isomorphism of $\mathfrak{A}$ to an automorphism of $\mathfrak{B}$. Then $\mathfrak{A}$ also possesses a finite extension with this property in $\mathscr{C}$.*

Just as Lemma 1.21 links Herwig's theorem to the basic finite model property for GF, the following links the Herwig–Lascar theorem to the stronger finite model property for GF expressed in Corollary 1.27.

A structure $\mathfrak{B}$ is $\sim_{\mathsf{g}}^{\ell}$-*homogeneous* if any guarded tuples $\mathbf{b}, \mathbf{b}'$ in $\mathfrak{B}$ such that $\mathfrak{B}, \mathbf{b} \sim_{\mathsf{g}}^{\ell} \mathfrak{B}, \mathbf{b}'$ are related by an automorphism of $\mathfrak{B}$.

Lemma 1.29 *Let $\mathscr{C}$ be a class of relational structures defined in terms of finitely many forbidden homomorphisms. Let $\mathfrak{B} \in \mathscr{C}$ be $\sim_{\mathsf{g}}^{\ell}$-homogeneous. Let $\mathfrak{B}'$ be the*

---

[5] Caveat: $\pi(h(\mathfrak{C})) \subseteq \mathfrak{A}$ need not itself be acyclic.

[6] It should be noted that this stand-alone argument does not support the complexity bounds that flow from the more constructive proof of Corollary 1.27 in [2].

expansion of $\mathfrak{B}$ *by a new relation for each one of the finitely many* $\sim_{\mathrm{g}}^{\ell}$*-types realised in* $\mathfrak{B}$*. Let* $\mathfrak{A}' = \mathfrak{B}' \upharpoonright A$ *be large enough to contain, for every guarded tuple* $\mathbf{b}$ *of* $\mathfrak{B}$*, at least one realisation of that* $\sim_{\mathrm{g}}^{\ell}$*-type.*

Then $\mathfrak{A}'$ *has a special Herwig extension* $\bar{\mathfrak{A}}' \supseteq \mathfrak{A}'$ *in* $\mathscr{C}$ *that is* $\sim_{\mathrm{g}}$*-equivalent to* $\mathfrak{B}'$ *in the sense that* $\bar{\mathfrak{A}}' \sim_{\mathrm{g}} \mathfrak{B}'$ *and* $\bar{\mathfrak{A}}', \mathbf{a} \sim_{\mathrm{g}} \mathfrak{B}', \mathbf{a}$ *for every guarded tuple* $\mathbf{a} \in \mathfrak{A}$*.*

Proof In view of Lemma 1.21 and Theorem 1.28 it suffices to show that the extension task for $\mathfrak{A}'$ has some, possibly infinite, solution in $\mathscr{C}$. But $\mathfrak{B}'$, being homogeneous, is such an infinite solution. □

Proof (of Corollary 1.27) Let $\mathscr{C}$ be defined by the condition that there are no homomorphic images of the finite structures $\mathfrak{C}_1, \ldots, \mathfrak{C}_m$. The class $\mathscr{C}_0 \supseteq \mathscr{C}$ of structures that admit no acyclically embedded homomorphic images of the $\mathfrak{C}_i$ is definable in GF by some $\gamma \in$ GF of guarded nesting depth ℓ, for some ℓ. To find finite models of $\varphi \in$ GF in $\mathscr{C}$, we moreover choose ℓ greater or equal to the nesting depth of φ. If φ has an infinite model in $\mathscr{C}$, then a $\sim_{\mathrm{g}}^{\ell}$-homogeneous infinite model $\mathfrak{B}$ of φ in $\mathscr{C}$ can be obtained as a suitable regular tree-like model of $\varphi \wedge \gamma$ (which in turn could be obtained from an arbitrary finite model of $\varphi \wedge \gamma$). An application of the lemma then yields a finite model in $\mathscr{C}$. □

Beside the notion of weakly N-acyclic coverings from [2], there is the stronger notion of N-acyclic coverings from [21], which rules out any small cyclic substructures in the cover. This yields an even stronger finite model property for GF and is essential for an expressive completeness proof for GF in finite model theory, as sketched in the next section. More canonical constructions of N-acyclic coverings and related hypergraph constructions have recently been explored in [22, 25]. But unlike the case of weakly N-acyclic covers, the known constructions of fully N-acyclic finite covers do not provide feasible size bounds.

Definition 1.30 A guarded bisimilar covering $\pi : \hat{\mathfrak{A}} \to \mathfrak{A}$ is N*-acyclic* if every induced substructure of size up to N of the cover $\hat{\mathfrak{A}}$ is acyclic.

Proposition 1.31 (Otto) *For every* $N \in \mathbb{N}$*, each finite relational* $\mathfrak{A}$ *admits* N*-acyclic coverings by finite structures.*

Corollary 1.32 (Otto) GF *has the finite model property in restriction to any class* $\mathscr{C}$ *of relational structures that is defined in terms of finitely many forbidden cyclic substructures.*

1.3.6 Expressive Completeness

The N-acyclic finite guarded bisimilar covers of Proposition 1.31 are also essential for the proof of the finite model theory version of Theorem 1.14. The issue at stake is the expressive completeness assertion, that a first-order definable property of guarded tuples in (finite) relational structures is expressible in GF (over all finite structures) if it is closed under guarded bisimulation equivalence (among finite structures). For both, the classical and the finite model theory reading, the Ehrenfeucht–Fraïssé theorem for GF shows that it suffices to prove the following, which may be read as a compactness property for $(\sim_{\mathsf{g}}^{\ell})_{\ell\in\mathbb{N}}$ versus $\sim_{\mathsf{g}}$: for any $\varphi(\mathbf{x}) \in \mathrm{FO}$ (in an explicitly guarded tuple $\mathbf{x}$ of free variables),

$$(*) \quad \begin{cases} \varphi(\mathbf{x}) \quad \text{invariant under } \sim_{\mathsf{g}} \quad \Rightarrow \\ \varphi(\mathbf{x}) \quad \text{invariant under } \sim_{\mathsf{g}}^{\ell} \quad \text{for some } \ell \in \mathbb{N}. \end{cases}$$

The classical proof typically achieves this through

(i) a compactness argument that reduces $(*)$ to: invariance under $\sim_{\mathsf{g}}$ implies invariance under $\sim_{\mathsf{g}}^{\omega}$ (i.e., $\equiv_{\mathrm{GF}}$); and
(ii) a proof of claim (i) through an upgrading argument involving saturated models: for $\mathfrak{A} \equiv_{\mathrm{GF}} \mathfrak{B}$ there are $\mathfrak{A}^* \equiv_{\mathrm{FO}} \mathfrak{A}$ and $\mathfrak{B}^* \equiv_{\mathrm{FO}} \mathfrak{B}$ for which (by saturation) $\mathfrak{A}^* \equiv_{\mathrm{GF}} \mathfrak{B}^*$ implies $\mathfrak{A}^* \sim_{\mathsf{g}} \mathfrak{B}^*$; the claim is then apparent from this diagram:

$$\begin{array}{ccc} \mathfrak{A} & \text{———} \equiv_{\mathrm{GF}} \text{———} & \mathfrak{B} \\ | & & | \\ \equiv_{\mathrm{FO}} & & \equiv_{\mathrm{FO}} \\ | & & | \\ \mathfrak{A}^* & \text{———} \sim_{\mathrm{g}} \text{———} & \mathfrak{B}^* \end{array}$$

For the finite model theory version, a passage through the necessarily infinite companion structures, which are involved in both parts of this classical argument, is not supported by the assumptions.

Instead, the upgrading needs to be based on a more constructive approach to model transformations, and focuses on a concrete level ℓ in $(*)$ that is determined by the width of the vocabulary and the quantifier rank q of the given φ. It follows this pattern:

$$\begin{array}{ccc} \mathfrak{A} & \text{———} \equiv_{\mathrm{GF}}^{\ell} \text{———} & \mathfrak{B} \\ | & & | \\ \sim_{\mathrm{g}} & & \sim_{\mathrm{g}} \\ | & & | \\ \hat{\mathfrak{A}} & \text{———} \equiv_{\mathrm{FO}}^{q} \text{———} & \hat{\mathfrak{B}} \end{array}$$

Here $\hat{\mathfrak{A}}$ and $\hat{\mathfrak{B}}$ are obtained as (finite) guarded bisimilar covers of $\mathfrak{A}$ and $\mathfrak{B}$, respectively, that need to be sufficiently acyclic and finitely saturated w.r.t. multiplicities: a certain level of N-acyclicity is necessary because $\hat{\mathfrak{A}}$ and $\hat{\mathfrak{B}}$ may necessarily have cycles, and differences w.r.t. short cycles would be FO-expressible at low quantifier rank; similarly for differences w.r.t. small branching degrees between relational hyperedges, which can also not be controlled in GF.

Technically rather intricate arguments in [21] use Proposition 1.31 as a starting point to provide companions $\hat{\mathfrak{A}}$ and $\hat{\mathfrak{B}}$ that support this proof idea.

Theorem 1.33 (Otto) $\mathrm{FO}/{\sim_\mathsf{g}} \equiv \mathrm{GF}$, *also in the sense of finite model theory: For every first-order formula* $\varphi(\mathbf{x})$ *in a relational vocabulary, the follwong are equivalent:*

(i) φ *is invariant under guarded bisimulation among finite structures.*
(ii) φ *is logically equivalent over all finite structures to a formula of* GF.

1.4 Guarded Negation Bisimulation

One natural decidable fragment of first-order logic that stands out because of its considerable algorithmic importance, is the positive existential fragment: $\exists\mathrm{posFO} \subseteq \mathrm{FO}$ is generated from atomic formulae by conjunction, disjunction and existential quantification. It is semantically characterised, as a fragment of FO, by preservation under homomorphisms. This characterisation is known as the Lyndon–Tarski theorem in classical model theory; for finite model theory, it was proved by Rossman in [29], with characteristically different techniques that also shed new light on the classical version. Any $\exists$posFO-formula can be equivalently re-written as a disjunction over existentially quantified conjunctions of atoms—so that it corresponds, in database terminology, to a *union of conjunctive queries*. And a conjunctive query asserts the existence of a homomorphism: consider a conjunctive query $\varphi = \varphi(\mathbf{x}) = \exists \mathbf{y} \bigwedge_\mathbf{i} \alpha_\mathbf{i}(\mathbf{z_i})$ with relational atoms $\alpha_i(\mathbf{z}_i)$ for tuples of variables $\mathbf{z}_i$ from $[\mathbf{xy}]$. With the template $\bigwedge_i \alpha_i(\mathbf{z}_i)$ associate a relational structure $\mathfrak{C}_\varphi$ whose universe is the set of variables $[\mathbf{xy}]$, and whose relations are interpreted by putting $\mathbf{z}_i$ into the relation involved in the atom α_i. Then $\mathfrak{A}, \mathbf{a} \models \varphi$ if, and only if, there is a homomorphism $h \colon \mathfrak{C}_\varphi \to \mathfrak{A}$ that maps $\mathbf{x}$ to $\mathbf{a}$. Interestingly, φ can equivalently be expressed in GF (i.e., is invariant under guarded bisimulation equivalence) if, and only if, $\mathfrak{C}_\varphi$ is acyclic.

$\exists\mathrm{posFO} \subseteq \mathrm{FO}$ or the formalism of (unions of) conjunctive queries are closed under nesting, but closure under (unconstrained) negation generates all of relational FO and becomes undecidable for satisfiability. The guarded fragment $\mathrm{GF} \subseteq \mathrm{FO}$, on the other hand, is closed under negation, but not under (unconstrained) nesting.

The introduction of the *guarded negation fragment* $\mathrm{GN} \subseteq \mathrm{FO}$ in [3] combines the innocuous ingredients in GF and $\exists$posFO with the natural constraints to produce a common extension of GF and $\exists$posFO that retains many of the good features, most notably decidability.

We follow the pattern of the treatment so far and put the appropriate notions of back&forth equivalence centre-stage. The characteristic feature is the interleaving of (local, and possibly size-bounded) homomorphisms with modal or guarded bisimulation.

1.4.1 Homomorphisms and Bisimulation

We start with a back&forth equivalence that interleaves homomorphisms with modal bisimulation; this will provide the Ehrenfeucht–Fraïssé notion and semantic characterisation of the *unary negation fragment* UN $\subseteq$ FO of [10], a modal precursor to the *guarded negation fragment* GN $\subseteq$ FO of [3].

A *unary negation bisimulation* relation between relational structures $\mathfrak{A}$ and $\mathfrak{B}$ is a set $Z \subseteq A \times B$ of positions, which are just pairs of related vertices in $\mathfrak{A}$ and $\mathfrak{B}$ as in modal bisimulation, subject to atom equivalence and more complex back&forth conditions involving homomorphisms. For all $(a, b) \in Z$:

(i) (*atom eq.*): $\mathfrak{A} \restriction \{a\} \simeq \mathfrak{B} \restriction \{b\}$;
(ii) (hom-*back*): for every $B_0 \subseteq B$ there is a homomorphism $h \colon \mathfrak{B} \restriction B_0 \to \mathfrak{A}$ such that $(h(b), b) \in Z$ for all $b \in B_0$, and $h(b) = a$ if $b \in B_0$;
(iii) (hom-*forth*): for every $A_0 \subseteq A$ there is a homomorphism $h \colon \mathfrak{A} \restriction A_0 \to \mathfrak{B}$ such that $(a, h(a)) \in Z$ for all $a \in A_0$, and $h(a) = b$ if $a \in A_0$.

We write $\mathfrak{A}, a \sim_{\mathsf{hom}} \mathfrak{B}, b$ if $(a, b) \in Z$ for some unary negation bisimulation relation Z between $\mathfrak{A}$ and $\mathfrak{B}$; $\mathfrak{A}, a \sim^{\ell}_{\mathsf{hom}} \mathfrak{B}, b$ for the finite approximation corresponding to a strategy for the second player for ℓ rounds in the natural bisimulation game associated with this back&forth scenario.

A generalisation of this idea leads from an equivalence between individual elements (as in modal bisimulation) to an equivalence based on guarded tuples (as in guarded bisimulation), similarly interleaving bisimulation with local homomorphisms: this is the notion of *guarded negation bisimulation* equivalence from [3].

A *guarded negation bisimulation* relation between relational structures $\mathfrak{A}$ and $\mathfrak{B}$ is a set Z of partial isomorphisms $\rho \colon \mathbf{a} \mapsto \mathbf{b}$ between guarded tuples or subsets, such that, for all $\rho \colon \mathbf{a} \mapsto \mathbf{b}$ in Z:

(i) (*atom eq.*): $\rho \colon \mathfrak{A} \restriction \mathbf{a} \simeq \mathfrak{B} \restriction \mathbf{b}$ (isomorphism of guarded substructures);
(ii) (hom-*back*): for all $B_0 \subseteq B$ there is a homomorphism $h \colon \mathfrak{B} \restriction B_0 \to \mathfrak{A}$ that is compatible with the restriction of ρ^{-1} to B_0, and such that $\rho' \colon h(\mathbf{b}') \mapsto \mathbf{b}'$ is in Z for all guarded tuples $\mathbf{b}'$ from B_0;
(iii) (hom-*forth*): for all $A_0 \subseteq A$ there is a homomorphism $h \colon \mathfrak{A} \restriction A_0 \to \mathfrak{B}$ that is compatible with the restriction of ρ to A_0, and such that $\rho' \colon \mathbf{a}' \mapsto h(\mathbf{a}')$ is in Z for all guarded tuples $\mathbf{a}'$ from A_0.

We write $\mathfrak{A}, \mathbf{a} \sim_{\mathsf{ghom}} \mathfrak{B}, \mathbf{b}$ and $\mathfrak{A}, \mathbf{a} \sim^{\ell}_{\mathsf{ghom}} \mathfrak{B}, \mathbf{b}$ to denote guarded bisimulation equivalence and its finite approximations.

Simple size-bounded versions of $\sim_{\mathsf{hom}}$ and $\sim_{\mathsf{ghom}}$ and their finite approximations are technically useful: we restrict conditions (hom-*back*) and (hom-*forth*) to subsets $B_0 \subseteq B$ and $A_0 \subseteq A$ of size up to k, for some fixed $k \in \mathbb{N}$. We write e.g. $\mathfrak{A}, \mathbf{a} \sim_{\mathsf{ghom};k} \mathfrak{B}, \mathbf{b}$ and $\mathfrak{A}, \mathbf{a} \sim^{\ell}_{\mathsf{ghom};k} \mathfrak{B}, \mathbf{b}$ in connection with this restricted notion of *k-bounded guarded negation bisimulation*, and similarly, e.g., $\mathfrak{A}, a \sim_{\mathsf{hom};k} \mathfrak{B}, b$ for a corresponding notion of *k-bounded unary negation bisimulation*.

We discuss briefly the extensions of modal logic and the guarded fragment that are obtained by closure of the existential positive fragment of FO under negation in suitably restricted settings:

- negation of 'unary' formulae in a single free variable for the unary negation fragment [10];
- negation of 'guarded' formulae in an explicitly guarded tuple of free variables for the guarded negation fragment [3].

Definition 1.34 The formulae of the *unary negation fragment* UN $\subseteq$ FO are generated from the atomic formulae by positive boolean connectives, existential quantification, and negation on formulae in at most one free variable.

It is obvious that, for suitable modal vocabularies, ML $\subseteq$ UN and that generally $\exists$posFO $\subseteq$ UN; both inclusions are easily seen to be strict (for non-trivial vocabularies). It turns out that formulae of UN (in at most a single free variable) are preserved under unary negation bisimulation, and in fact this property characterises the unary negation fragment as a fragment of FO, classically. See [10] for this and many related model-theoretic results, also regarding the fixpoint extension of UN and including decidability for satisfiability and finite satisfiability.

Definition 1.35 The formulae of the *guarded negation fragment* GN $\subseteq$ FO are generated from the atomic formulae by positive boolean connectives, existential quantification, and negation on formulae in an explicitly guarded tuple of free variables.

It is not hard to see that UN $\subseteq$ GN and GF $\subseteq$ GN, and that these inclusions are strict in general. Formulae of GN (in an explicitly guarded tuple of free variables) are preserved under guarded negation bisimulation equivalence; this preservation property also characterises GN as a fragment of FO, in the sense of classical model theory, as shown in [3].

For useful Ehrenfeucht–Fraïssé correspondences, which rely on the natural notion of nesting depth in GN and UN and induce equivalence relations of finite index, we need to bound the size of the existential quantifications (conjunctive queries) by some width parameter. For the games and bisimulation notions this restriction leads to the size bounded equivalences like $\sim^{\ell}_{\mathsf{ghom};\mathsf{k}}$. For the logics, we correspondingly let GN$[k]$ $\subseteq$ GN stand for those formulae that can be generated with existential quantifications over up to k variables at a time. To avoid pathologies, we shall always assume that k is no less than the width of the vocabulary.

It is then not hard to see that equivalence w.r.t. GN$[k]$ up to nesting depth ℓ and $\sim^{\ell}_{\mathsf{ghom};\mathsf{k}}$ are related in an Ehrenfeucht–Fraïssé correspondence. The theorem gives an indicative example; its variants for UN and also for infinitary versions of UN and GN in the style of Karp theorems are straightforward.

Theorem 1.36 (Ehrenfeucht–Fraïssé for GN[k]) *In restriction to finite relational vocabularies, fixed $k \in \mathbb{N}$, and for every $\ell \in \mathbb{N}$:*

$$\mathfrak{A}, \mathbf{a} \sim^{\ell}_{\mathsf{ghom};\mathsf{k}} \mathfrak{B}, \mathbf{b} \quad \textit{if, and only if,} \quad \mathfrak{A}, \mathbf{a} \equiv^{\ell}_{\mathrm{GN}[k]} \mathfrak{B}, \mathbf{b}.$$

1.4.2 Towards a (Finite) Model Theory of Guarded Negation

We summarise some key techniques and a few further results for the model theory of GN and GN$[k]$, especially pertaining to the finite model property and to the expressive completeness concern in finite model theory. We concentrate on guarded negation rather than unary negation, since this is the richer of the two settings; technically it is, moreover, more directly related to one of our main themes, viz., to the interesting passage from graph-like structures to general relational structures with an emphasis on the hypergraph of guarded subsets.

Theorem 1.37 (Bárány–ten Cate–Segoufin) GN *has the finite model property.*

The argument from [3] is based on a reduction from GN-satisfiability to satisfiability of GF under constraints imposed by *forbidden homomorphisms*, and thus, essentially, a reduction to Corollary 1.27.

The semantics of a formula $\varphi(\mathbf{x}) \in$ GN (in explicitly guarded free variables $\mathbf{x}$) can be translated into a collection of auxiliary specifications that subject certain guarded tuples $\mathbf{a}$ in a prospective model $\mathfrak{A}$ to *positive* or *negative* requirements w.r.t. homomorphisms:

- (pos. hom.) requiring the existence of a homomorphism $h\colon \mathfrak{C}, \mathbf{c} \to \mathfrak{A}, \mathbf{a}$, for certain finite templates $\mathfrak{C}, \mathbf{c}$;
- (neg. hom.) ruling out the existence of any homomorphism $h\colon \mathfrak{C}, \mathbf{c} \to \mathfrak{A}, \mathbf{a}$, for certain finite templates $\mathfrak{C}, \mathbf{c}$.

In both cases, the templates $\mathfrak{C}, \mathbf{c}$ are abstracted from the underlying conjunctive queries or positive existential parts (in a suitable normal form). A standard process of relational Skolemisation thus translates $\varphi(\mathbf{x})$ into a positive boolean combination of requirements of the form (pos. hom.) and (neg. hom.) for all tuples in certain (auxiliary) relations. A further crude Skolemisation step serves to provide realisations of positive requirements in image substructures that are guarded as a whole by new auxiliary relations; this puts all (pos. hom.) requirements into GF, and leaves just the negative requirements of the form (neg. hom.) to cope with. But this is precisely the situation in which Corollary 1.27 yields finite models whenever there are any models.

The requirements for an expressive completeness proof for GN$[k]$ in relation to all $\sim_{\mathsf{ghom};k}$-invariant FO-definable properties (of guarded tuples), which is meant to work in finite model theory, are considerable higher. The basic idea again is to use an upgrading through $\sim_{\mathsf{ghom};k}$-compatible model transformations that work in finite structures. I.e., we want to follow this pattern, presented without the guarded parameter tuples:

$$\begin{array}{ccc}
\mathfrak{A} & \text{———} \sim^{\ell}_{\mathsf{ghom};k} \text{———} & \mathfrak{B} \\
| & & | \\
\sim_{\mathsf{ghom};k} & & \sim_{\mathsf{ghom};k} \\
| & & | \\
\hat{\mathfrak{A}} & \text{———} \equiv^{q}_{\mathrm{FO}} \text{———} & \hat{\mathfrak{B}}
\end{array}$$

More precisely, given some first-order φ of quantifier rank q that is invariant under $\sim_{\mathsf{ghom;k}}$, and finite structures $\mathfrak{A}$ and $\mathfrak{B}$ that are $\sim^{\ell}_{\mathsf{ghom;k}}$-equivalent for sufficiently high level ℓ, we need to provide finite $\sim_{\mathsf{ghom;k}}$-equivalent companion structures $\hat{\mathfrak{A}}$ and $\hat{\mathfrak{B}}$ for which $\sim^{\ell}_{\mathsf{ghom;k}}$-equivalence implies $\equiv^{q}_{\mathrm{FO}}$-equivalence, so that

$$\hat{\mathfrak{A}} \models \varphi \quad \text{iff} \quad \hat{\mathfrak{B}} \models \varphi.$$

If this can generally be achieved, for a uniform level ℓ that only depends on φ, then the diagram shows that φ is preserved under $\sim^{\ell}_{\mathsf{ghom;k}}$, and by the Ehrenfeucht–Fraïssé theorem for GN[k], Theorem 1.36, is equivalently expressible in GN[k].

The crucial features with respect to which $\hat{\mathfrak{A}}$ and $\hat{\mathfrak{B}}$ need to agree, even though these features are *not* GN-definable are

- presence of small cyclic configurations other than those explicitly ruled out by (neg. hom.) assertions;
- multiplicities (up to a threshold) and isomorphism types of realisations of (pos. hom.) assertions.

That $\mathfrak{A}$ and $\mathfrak{B}$ agree w.r.t. the relevant (pos. hom.) and (neg. hom.) assertions follows from their $\sim^{\ell}_{\mathsf{ghom;k}}$-equivalence. Then agreement w.r.t. to the above features is relatively easy to achieve in infinite tree unfoldings of $\mathfrak{A}$ and $\mathfrak{B}$ that are simultaneously saturated w.r.t. *all* admissible isomorphism types of the relevant (pos. hom.) assertions. Relational Skolemisation and an application of the finite model property for GN, Theorem 1.37, yield finite companions $\hat{\mathfrak{A}}'_0$ and $\hat{\mathfrak{B}}'_0$. These further admit finite coverings by suitable $\hat{\mathfrak{A}}'$ and $\hat{\mathfrak{B}}'$ whose degree of acyclicity and saturation w.r.t. small multiplicities show them to be equivalent in the sense of $\equiv^{q}_{\mathrm{FO}}$ (this last part of the argument is as for Theorem 1.33). This yields the following result from [23].

Theorem 1.38 (Otto) $\mathrm{FO}/{\sim_{\mathsf{ghom;k}}} \equiv \mathrm{GN}[k]$, *classically and in the sense of finite model theory.*

1.5 Summary

We have seen that bisimulation equivalence is a very flexible and powerful concept for the analysis of many logics. In its classical form it is one of the crucial tools in the study of modal logics, and its generalisations to various forms of guarded bisimulation provide indispensable methods for understanding the expressive power as well as the model-theoretic and algorithmic properties of more and more powerful variants of guarded logics.

First of all, an appropriate notion of bisimulation for a logic L characterises semantic invariance and logical indistinguishability: bisimilar nodes or tuples in two structures cannot be distinguished by formulae of L. In this sense, bisimulation is closely related to the characterisation of elementary equivalence via

Ehrenfeucht-Fraïssé games, and bisimulation games can indeed be viewed as special cases of these. The specific form of a bisimulation depends mostly on the nature of the quantification patterns that the associated logic provides. In game theoretic terms, the restrictions on the permitted forms of quantification are reflected by the rules in the associated bisimulation game. In modal and guarded bisimulation games the configurations at any position in a play are restricted in the sense that they may only contain elements that are, in a sense, 'close together'. As a consequence, bisimulation permits us to control the complexity of model constructions and leads to results about model-theoretic properties of modal and guarded logics such as the tree model property of modal logics and the fact that satisfiable guarded formulae have models of bounded tree width. While such results are usually not too difficult to establish for infinite models, corresponding constructions for finite models may be quite challenging and require intricate combinatorial arguments and sophisticated mathematical techniques.

A further highlight of the bisimulation-based analysis of logics are the characterisation theorems that provide, inside a classical level of logical expressiveness such as first-order or monadic second-order definability, a sort of converse of bisimulation invariance. Typically such characterisation theorems state that a modal or guarded logic is not only invariant under bisimulation, but is in fact (up to logical equivalence) precisely the bisimulation invariant part of that level. Again such theorems are, by means of compactness and model-theoretic notions such as saturation or by automata-theoretic methods, better understood and easier to prove for arbitrary (i.e. finite or infinite) models, and much more challenging, and in some cases open, on finite structures.

A related issue that we have not treated here concerns Lindström characterisations of modal and guarded logics. It is shown in [7, 8] that no logic that is bisimulation invariant, compact, and closed under relativisation can properly extend the basic modal logic ML. In this proof, a crucial role is played by a locality criterion (which is implied by compactness and relativisation for any bisimulation closed logic) saying that the truth of a formula at a given node only depends on a neighbourhood of points reachable in a bounded number of steps. For guarded logics, and even for modal logics with a global modality no such locality criterion is available. To obtain Lindström characterisations for GF and ML[∀], Otto and Piro [26] use instead the Tarski Union Property saying that the union of any elementary chain is itself an elementary extension of each structure in the chain. They show that ML[∀] and GF are the maximal compact logics that satisfy the Tarski Union Property and the corresponding bisimulation invariance. It is open whether there are Lindström characterisations of these logics that are not based on the Tarski Union Property but, say, on compactness and relativisation.

Finally the bisimulation-based analysis of modal and guarded logics also leads to important insights concerning their algorithmic properties. Since satisfiable formulae always admit simple models, for instance tree-like ones, and since modal and guarded logics, including the fixed-point variants such as the modal μ-calculus and the guarded fixed-point logic μGF can be embedded or interpreted in monadic second-order logic on trees, powerful automata theoretic methods become available for checking satisfiability and for evaluating formulae. It still remains to determine where the limits are for fragments of first-order logic (and fixed-point logic or even second-order logic) that are invariant under a suitable notion of (guarded) bisimulation that is sufficient to ensure similar model-theoretic and algorithmic properties as those that have been established for modal and guarded logic. In particular, can we find in this way stronger decidable fragments of first-order logic, fixed-point logic and second-order logic than those known so far?

References

1. Andréka H, van Benthem J, Németi I (1998) Modal languages and bounded fragments of predicate logic. J Philoso Logic 27:217–274
2. Bárány V, Gottlob G, Otto M (2014) Querying the guarded fragment. Logical Methods Comput Sci (to appear)
3. Bárány V, ten Cate B, Segoufin L (2011) Guarded negation. In: Proceedings of ICALP, pp 356–367
4. Beeri C, Fagin R, Maier D, Yannakakis M (1983) On the desirability of acyclic database schemes. J ACM 30:497–513
5. van Benthem J (1983) Modal logic and classical logic. Bibliopolis, Napoli
6. van Benthem J (2005) Guards, bounds, and generalized semantics. J Logic Lang Inform 14(3):263–279
7. van Benthem J (2007) A new modal Lindström theorem. Log Univers 1:125–138
8. van Benthem J, ten Cate B, Väänänen J (2007) Lindström theorems for fragments of first-order logic. In: Proceedings of 22nd IEEE symposium on logic in computer science, LICS 2007, pp 280–292
9. Berge C (1973) Graphs and hypergraphs. North-Holland, Amsterdam
10. ten Cate B, Segoufin L (2011) Unary negation. In: Proceedings of STACS, pp 344–355
11. Dawar A, Otto M (2009) Modal characterisation theorems over special classes of frames. Ann Pure Appl Logic 161:1–42
12. Grädel E (1999) On the restraining power of guards. J Symbolic Logic 64:1719–1742
13. Grädel E (2002) Guarded fixed point logics and the monadic theory of countable trees. Theoret Comput Sci 288:129–152
14. Grädel E, Hirsch C, Otto M (2002) Back and forth between guarded and modal logics. ACM Trans Comput Logics 3:418–463
15. Grädel E, Walukiewicz I (1999) Guarded fixed point logic. In: Proceedings of 14th IEEE symposium on logic in computer science, LICS 1999, pp 45–54
16. Herwig B (1995) Extending partial isomorphisms on finite structures. Combinatorica 15:365–371
17. Herwig B, Lascar D (2000) Extending partial isomorphisms and the profinite topology on free groups. Trans AMS 352:1985–2021
18. Otto M (2004) Modal and guarded characterisation theorems over finite transition systems. Ann Pure Appl Logic 164(12):1418–1453
19. Otto M (2006) Bisimulation invariance and finite models. In: Colloquium logicum 2002. Lecture notes in logic, pp 276–298, ASL
20. Otto M (2011) Model theoretic methods for fragments of FO and special classes of (finite) structures. In: Esparza J, Michaux C, Steinhorn C (eds) Finite and algorithmic model theory, volume 379 of LMS lecture notes, pp 271–341. CUP
21. Otto M (2012a) Highly acyclic groups, hypergraph covers and the guarded fragment. J ACM 59:1
22. Otto M (2012b) On groupoids and hypergraphs. arXiv:1211.5656 (preprint)

23. Otto M (2013a) Expressive completeness through logically tractable models. Ann Pure Appl Logic 164(12):1418–1453
24. Otto M (2013b) Groupoids, hypergraphs and symmetries in finite models. In: Proceedings of 28th IEEE symposium on logic in computer science, LICS 2013
25. Otto M (2014) Finite groupoids, finite coverings and symmetries in finite structures. arXiv:1404.4599 (preprint)
26. Otto M, Piro R (2008) A Lindström characterisation of the guarded fragment and of modal logic with a global modality. In: Areces C, Goldblatt R (eds) Advances in modal logic 7, pp 273–288
27. Rabin M (1969) Decidability of second-order theories and automata on infinite trees. Trans AMS 141:1–35
28. Rosen E (1997) Modal logic over finite structures. J Logic Lang Inform 6:427–439
29. Rossman B (2008) Homomorphism preservation theorems. J ACM 55:1–53
30. Vardi M (1998) Reasoning about the past with two-way automata. In: Automata, languages and programming ICALP 98. Lecture notes in computer science, vol 1443. Springer, pp 628–641