

Hereditary First-Order Logic: the Tractable Quantifier Prefix Classes

Manuel Bodirsky 

Institut für Algebra, TU Dresden, Germany

Santiago Guzmán-Pro 

Institut für Algebra, TU Dresden, Germany

Abstract

Many computational problems can be modelled as the class of all finite structures $\mathbb{A}$ that satisfy a fixed first-order sentence ϕ *hereditarily*, i.e., we require that every (induced) substructure of $\mathbb{A}$ satisfies ϕ . We call the corresponding computational problem the *hereditary model checking* problem for ϕ , and denote it by $\text{Her}(\phi)$.

We present a complete description of the quantifier prefixes for ϕ such that $\text{Her}(\phi)$ is in P; we show that for every other quantifier prefix there exists a formula ϕ with this prefix such that $\text{Her}(\phi)$ is coNP-complete. Specifically, we show that if Q is of the form $\forall^*\exists\forall^*$ or of the form $\forall^*\exists^*$, then $\text{Her}(\phi)$ can be solved in polynomial time whenever the quantifier prefix of ϕ is Q . Otherwise, Q contains $\exists\exists\forall$ or $\exists\forall\exists$ as a subword, and in this case, there is a first-order formula ϕ whose quantifier prefix is Q and $\text{Her}(\phi)$ is coNP-complete. Moreover, we show that there is no algorithm that decides for a given first-order formula ϕ whether $\text{Her}(\phi)$ is in P (unless $P=NP$).

2012 ACM Subject Classification Theory of computation $\rightarrow$ Complexity theory and logic

Keywords and phrases Quantifier prefix, first-order Logic, Computational Complexity, Polynomial-time algorithm, coNP-completeness

Digital Object Identifier 10.4230/LIPIcs.CSL.2026.6

Related Version *Full Version*: <https://arxiv.org/abs/2411.10860> [3]

Funding Both authors have been funded by the European Research Council (Project POCOCOP, ERC Synergy Grant 101071674). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Acknowledgements  This is a low-co2 research paper: <https://tcs4f.org/low-co2-v1>. This research was developed, written, submitted and presented without the use of air travel.

1 Introduction

Vertex-deletion problems were first studied as optimization problems [17, 18]: Given an input graph $\mathbb{G}$ determine the minimum number of vertices to be deleted so that the remaining (induced) subgraph belongs to a specified class $\mathcal{C}$. Krishnamoorthy and Deo [17] showed that this problem is NP-hard for several natural graph classes $\mathcal{C}$ such as trees, planar, bipartite, hamiltonian, interval, and chordal graphs. Shortly after, Lewis and Yannakakis [18] proved that if $\mathcal{C}$ is a non-trivial hereditary class of finite graphs (i.e., $\mathcal{C}$ is an infinite class of finite graphs which is closed under induced subgraphs and which is not the class of all finite graphs), then the vertex-deletion problem defined by $\mathcal{C}$ is NP-hard.

Given the previous hardness results, it was natural to study vertex-deletion problems from the viewpoint of approximation or of parametrized complexity. Regarding the former, Fujito [13] proposed a unified polynomial-time algorithm that approximates an optimal solution to a vertex-deletion problem defined by a non-trivial hereditary class. Some



© Manuel Bodirsky and Santiago Guzmán-Pro;

licensed under Creative Commons License CC-BY 4.0

34th EACSL Annual Conference on Computer Science Logic (CSL 2026).

Editors: Stefano Guerrini and Barbara König; Article No. 6; pp. 6:1–6:20

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

well-known results include the fixed parameter tractability of the odd-cycle transversal problem [21], and of the feedback vertex set problem [20] – the former corresponds to the vertex-deletion problem for bipartite graphs, and the latter to the vertex-deletion problem for forests.

Recently, the parametrized complexity perspective has been studied systematically for vertex-deletion problems defined by graph classes expressible by some first-order sentence ϕ [1, 12]. In this paper we are interested in the class of graphs $\mathbb{G}$ such that no matter how many vertices are removed from $\mathbb{G}$, the remaining induced subgraph does not satisfy ϕ . Since first-order logic is closed under negations, we prefer the positive phrasing of this problem: Given a graph $\mathbb{G}$, test whether all non-empty induced subgraphs of $\mathbb{G}$ satisfy a specified first-order formula ψ .

Hereditary first-order logic

From now on we work in the setting of relational structures, for which graphs and digraphs are prototypical examples. We use the convention that all structures have a non-empty domain. We begin by introducing *hereditary first-order logic*. A structure $\mathbb{A}$ *hereditarily satisfies* ϕ if every substructure¹ $\mathbb{B}$ of $\mathbb{A}$ satisfies ϕ . We denote by $\text{Her}(\phi)$ the class of finite structures that hereditarily satisfy ϕ . The *hereditary model checking problem* for a fixed sentence ϕ consists of deciding whether an input structure $\mathbb{A}$ belongs to $\text{Her}(\phi)$. Since verifying whether a finite structure $\mathbb{A}$ models a fixed first-order formula ϕ can be done in polynomial time, it follows that $\text{Her}(\phi)$ is in coNP.

We say that a class $\mathcal{C}$ of finite τ -structures is *hereditarily first-order definable* if there is a first-order sentence ϕ such that $\mathcal{C} = \text{Her}(\phi)$. In this case, we also say that $\mathcal{C}$ is in HerFO. Clearly, every class in HerFO is *hereditary*, i.e., closed under taking substructures.

Three simple examples

Clearly, every hereditary class in FO is also in HerFO. The following graph and digraph classes are in HerFO, but not in FO.

► **Example 1 (Forests).** Consider a first-order formula ϕ stating that “there is a loopless vertex of degree at most 1”. Clearly, every forest hereditarily satisfies ϕ . Conversely, suppose that $\mathbb{G}$ is a finite graph that hereditarily satisfies ϕ . Then $\mathbb{G}$ has a loopless vertex v of degree at most 1. We inductively see that the subgraph with vertex set $(G \setminus \{v\})$ is a forest, i.e., has no cycles, and since v has degree at most 1 and does not have a loop, we conclude that $\mathbb{G}$ is a forest.²

► **Example 2 (Chordal graphs).** A graph $\mathbb{G}$ is *chordal* if every cycle of $\mathbb{G}$ contains a chord; equivalently, if $\mathbb{G}$ contains no induced cycle of length $n \geq 4$. Rose [22] proved that a graph $\mathbb{G}$ is chordal if and only if every induced subgraph contains a vertex whose neighbourhood induces a clique. Hence, if ϕ is a first-order sentence stating “there is a (loopless) vertex v such that every two neighbours of v are adjacent”, then $\text{Her}(\phi)$ describes the class of chordal graphs. It is known that membership in this class can be decided in polynomial time [23].

¹ Some authors use the term “induced substructure”; we follow the convention in model theory and omit the adjective ‘induced’.

² This example naturally generalizes to *k-degenerate* graphs, i.e., the class of graphs $\mathbb{G}$ such that every subgraph of $\mathbb{G}$ contains a vertex of degree at most k .

► **Example 3** (Directed acyclic digraphs). The class of acyclic digraphs belongs to HerFO as well. Indeed, it suffices to consider the first-order sentence $\exists x \forall y. \neg E(x, y)$, i.e., the first order sentence that states that there exists a *sink*, i.e., a vertex without outgoing edges. Equivalently, the constraint satisfaction problem for $(\mathbb{Q}, <)$ (denoted by $\text{CSP}(\mathbb{Q}, <)$, see Section 2) is in HerFO. Also this computational problem can be solved in polynomial time (e.g., by depth-first-search).

We are interested in the expressive power of hereditary first-order logic, as well as complexity classification for the hereditary model checking problem. Regarding the former, we will see that HerFO is a particularly natural formalism when it comes to the description of constraint satisfaction problems (for the definition of CSPs, see Section 2). Regarding the latter, we will study complexity classifications based on the quantifier prefix of the fixed first-order sentence ϕ . This is motivated by similar classifications for universal and existential second-order logic [9, 14], which we now review.

Prefix classifications

It is straightforward to observe that $\text{Her}(\phi)$ can be expressed in universal monadic second-order logic.

► **Observation 4.** *Consider a first-order τ -formula $\phi := Q_1 x_1 \dots Q_n x_n. \psi(x_1, \dots, x_n)$ and a finite structure $\mathbb{A}$. If S is a unary predicate not in τ , then $\mathbb{A}$ hereditarily satisfies ϕ if and only if $\mathbb{A}$ models*

$$\forall S \, Q_1 x_1 \dots Q_n x_n \left(\exists z. S(z) \wedge \bigwedge_{i \in U} S(x_i) \implies \bigwedge_{i \in E} S(x_i) \wedge \psi(x_1, \dots, x_n) \right),$$

where U (respectively, E) is the set of indices $i \in \{1, \dots, n\}$ such that Q_i is a universal (respectively, existential) quantifier.

The question whether a given existential second-order (ESO) sentence describes a polynomial-time solvable problem is easily seen to be undecidable (see, e.g., Theorem 1.4.2 in [2]). This motivates quantifier-prefix dichotomy results for existential second-order logic [9, 14]. For instance, the classification for monadic ESO is as follows [14, Figure 1(a)]. For every quantifier prefix $Q \in \{\exists, \forall\}^*$ the following holds:

- either Q is of the form $\exists^* \forall$, and in this case every monadic ESO sentence Φ whose first-order part has a quantifier-free prefix Q is polynomial-time decidable, or
- Q contains $\forall \forall$ or $\forall \exists$ as subwords, and in this case there is an ESO sentence Φ whose first-order part has quantifier prefix Q and deciding Φ is NP-complete.

It follows from this classification, via Observation 4, that $\text{Her}(\phi)$ is in P whenever ϕ is a first-order sentence with a quantifier prefix of the form $\forall^* \exists$. However, as we will soon see, HerFO enjoys a richer polynomial-time solvable fragment.

Our contributions

We first study the expressive power of HerFO in the context of CSPs. We see that HerFO can express finite-domain CSPs not expressible in first-order logic (Example 25), that it can express coNP-complete infinite-domain CSPs (Theorem 8), and that every CSP expressible in HerFO is the CSP of an ω -categorical structure (Remark 6). We then turn to computational aspects of the hereditary model checking problem for fixed sentences ϕ . In particular, we show that if the signature of ϕ is monadic, then $\text{Her}(\phi)$ is solvable in polynomial-time (Proposition 11).

Our first main results are two classifications of the computational complexity of problems in HerFO. The first one concerns sentences with at least one binary predicate (Theorem 20), and the second one concerns sentences ϕ such that $\text{Her}(\phi)$ describes a CSP (Corollary 21). These classifications coincide, and are based on the quantifier prefix $Q \in \{\exists, \forall\}^*$ of the fixed first-order formula:

- either Q is of the form $\forall^*\exists^*$ or the form $\forall^*\exists\forall^*$, and in this case $\text{Her}(\phi)$ is polynomial-time decidable for every first-order formula ϕ with quantifier prefix Q , or
- Q contains $\exists\exists\forall$ or $\exists\forall\exists$ as a subword, and in this case there are first-order sentences with quantifier prefix Q such that $\text{Her}(\phi)$ is coNP-complete.

Our final main result shows that more fine-grained classifications of the complexity are undecidable: there is no algorithm that tests for a given first-order sentence ϕ whether $\text{Her}(\phi)$ can be solved in polynomial time (Theorem 23). Moreover, we show that this is also the case even if ϕ is restricted to having quantifier-prefix $Q = \exists\exists\exists\forall$ or $Q = \exists\forall\exists$ (Corollary 24). We leave open the decidability of this meta-problem for first-order sentences ϕ with quantifier prefix $\exists\exists\forall$.

Outline

We begin by recalling some basic concepts from finite model theory (Section 2). We then list further examples of problems in HerFO, and prove that certain problems cannot be expressed in HerFO (Section 3); it will also become clear why hereditary model checking is quite natural in the context of constraint satisfaction. The classifications of the complexity of HerFO depending on the allowed quantifier prefix are the main results of Section 4. The undecidability of tractability of HerFO can be found in Section 5. A series of problems that are left open can be found in Section 6. In Appendix B we present a comparison between our complexity classifications, the complexity landscape for the parametrized version of HerFO from [12], and the general model checking problem for UMSO from [14].

2 Preliminaries

We assume basic familiarity with first-order logic and we follow standard notation from model theory, as, e.g., in [15]. We also use standard notions from complexity theory.

(First-order) structures

Given a relational signature τ and a τ -structure $\mathbb{A}$, we denote by $R^{\mathbb{A}}$ the interpretation in $\mathbb{A}$ of a relation symbol $R \in \tau$. Also, we denote relational structures with letters $\mathbb{A}, \mathbb{B}, \mathbb{C}, \dots$, and their domains by $A, B, C, \dots$. In this article, structures have non-empty domains, and we only work with finite signatures τ .

If $\mathbb{A}$ and $\mathbb{B}$ are τ -structures, then a *homomorphism* from $\mathbb{A}$ to $\mathbb{B}$ is a map $f: A \rightarrow B$ such that for all $a_1, \dots, a_k \in A$ and $R \in \tau$ of arity k , if $(a_1, \dots, a_k) \in R^{\mathbb{A}}$, then $R(f(a_1), \dots, f(a_k)) \in R^{\mathbb{B}}$. We write $\mathbb{A} \rightarrow \mathbb{B}$ if there exists a homomorphism from $\mathbb{A}$ to $\mathbb{B}$, and we denote by $\text{CSP}(\mathbb{B})$ the class of finite structures $\mathbb{A}$ such that $\mathbb{A} \rightarrow \mathbb{B}$. Let $\mathcal{C}$ be a class of finite τ -structures. We say that $\mathcal{C}$ is

- *closed under homomorphisms* if for every $\mathbb{B} \in \mathcal{C}$, if $\mathbb{B} \rightarrow \mathbb{A}$, then $\mathbb{A} \in \mathcal{C}$ as well.
- *closed under inverse homomorphisms* if for every $\mathbb{B} \in \mathcal{C}$, if $\mathbb{A} \rightarrow \mathbb{B}$, then $\mathbb{A} \in \mathcal{C}$ as well (i.e., the complement of $\mathcal{C}$ in the class of all finite τ -structures is closed under homomorphisms).

Note that $\text{CSP}(\mathbb{B})$ is closed under inverse homomorphisms.

If $\mathbb{A}$ and $\mathbb{B}$ are τ -structures with disjoint domains A and B , respectively, then the *disjoint union* $\mathbb{A} \uplus \mathbb{B}$ is the τ -structure $\mathbb{C}$ with domain $A \cup B$ and the relation $R^{\mathbb{C}} = R^{\mathbb{A}} \cup R^{\mathbb{B}}$ for every $R \in \tau$. Note that $\text{CSP}(\mathbb{B})$ is *closed under disjoint unions*, i.e., if $\mathbb{A} \in \text{CSP}(\mathbb{B})$ and $\mathbb{B} \in \text{CSP}(\mathbb{B})$, then $\mathbb{A} \uplus \mathbb{B} \in \text{CSP}(\mathbb{B})$. It is well-known that a class of finite τ -structures $\mathcal{C}$ is of the form $\text{CSP}(\mathbb{B})$ for some countably infinite τ -structure $\mathbb{B}$ if and only if $\mathcal{C}$ is closed under inverse homomorphisms and disjoint unions (see, e.g., [2, Lemma 1.1.8]).

For examples we will often use graphs and digraphs, and we will think of these as binary structures with signature $\{E\}$. We follow and adapt standard notions from graph theory [6] to this setting. In particular, given a digraph $\mathbb{D}$, we call $E^{\mathbb{D}}$ the *edge set* of $\mathbb{D}$, and its elements we call *edges*. A *graph* is a digraph whose edge set is a symmetric relation, and an *oriented graph* is a digraph whose edge set is an anti-symmetric relation. Given a positive integer n , we denote by K_n the complete graph on n vertices, and by $\vec{P}_n$ the *directed path* on n vertices. An *oriented path* is an oriented graph whose symmetric closure is a path.

Fragments of first-order logic

A first-order τ -formula ϕ is called *positive* if it does not use the negation symbol $\neg$ (so it just uses the logical symbols $\forall, \exists, \wedge, \vee, =$, variables and symbols from τ). The negation of a positive formula is called *negative*; note that every negative formula is equivalent to a formula in prenex conjunctive normal form where every atomic formula appears in negated form, and all negation symbols are in front of atomic formulas; such formulas will also be called negative.

Consider a quantifier-free formula ψ using only variables, the symbols $\wedge$ and $=$, and symbols from τ , and let ψ' be the formula obtained from ψ by iteratively removing each equality conjunct $x = y$ and substituting each occurrence of the variable y by the variable x . We now consider the structure $\mathbb{C}_\psi$ whose vertex set is the set of variables appearing in ψ' , and the interpretation of $R \in \tau$ consists of those tuples $\bar{x}$ such that ψ' contains the positive literal $R(\bar{x})$. We say that a negative first-order sentence ϕ in prenex conjunctive normal form is *connected* if for every clause φ of the quantifier-free part of ϕ , the structure $\mathbb{C}_{\neg\varphi}$ is connected.

3 Hereditary model checking and CSPs

In this section we see that hereditary model checking naturally arises in the context of constraint satisfaction problems. We do so by providing several examples of CSPs in HerFO. We also include inexpressible examples of HerFO which help build intuition about hereditary model checking.

It is well-known that a universal first-order sentence ϕ in conjunctive normal form describes a CSP if ϕ is negative and connected (see, e.g., Theorem 5.6.2 in [2]). The following observation extends this to not necessarily universal sentences.

► **Observation 5.** *The following statements hold for every first-order sentence ϕ in prenex conjunctive normal form.*

- *If ϕ is negative, then $\text{Her}(\phi)$ is closed under inverse homomorphisms.*
- *If ϕ is negative and connected, then there is a structure $\mathbb{S}$ such that $\text{Her}(\phi) = \text{CSP}(\mathbb{S})$.*

Proof. To prove the first itemized statement, suppose that there is a homomorphism $f: \mathbb{A} \rightarrow \mathbb{B}$ and $\mathbb{B} \in \text{Her}(\phi)$. Let $\mathbb{A}'$ be a substructure of $\mathbb{A}$. The substructure of $\mathbb{B}$ with vertex set $f[A']$ models ϕ . Suppose for contradiction that $\mathbb{A}'$ does not satisfy ϕ . Then it would satisfy the sentence $\neg\phi$, which is equivalent to a positive sentence. Since surjective homomorphisms preserve positive first-order formulas [15, Theorem 2.4.3], we obtain that $\mathbb{A}$ satisfies $\neg\phi$, a contradiction.

It is straightforward to observe that if ϕ is negative and connected, then the class of finite models of ϕ is closed under disjoint unions. Hence, if every substructure of $\mathbb{A}$ and every substructure of $\mathbb{B}$ satisfies ϕ , then every substructure of $\mathbb{A} \uplus \mathbb{B}$ satisfies ϕ , i.e., $\text{Her}(\phi)$ is closed under disjoint unions. By the first claim, $\text{Her}(\phi)$ is also closed under inverse homomorphisms, and it thus follows that there is a structure $\mathbb{S}$ such that $\text{CSP}(\mathbb{S}) = \text{Her}(\phi)$ – for any class $\mathcal{C}$ of τ -structures there exists a τ -structure $\mathbb{S}$ such that $\mathcal{C} = \text{CSP}(\mathbb{S})$ if and only if $\mathcal{C}$ is closed under disjoint unions and inverse homomorphisms [2, Lemma 1.1.8]. ◀

► **Remark 6.** Every CSP in HerFO is the CSP of an ω -categorical structure (see [2] for the many consequences that the model theoretic property of ω -categoricity has for the study of the complexity of the CSP). This follows from a result in [5, Corollary 14], which states that every CSP in monadic second-order logic (MSO) is the CSP for an ω -categorical structure; clearly, HerFO is a fragment of monadic second-order logic (see, e.g., Observation 4). For a concrete application of this remark see Example 26.

► **Example 7.** Let $\tau = \{EQ, N\}$ be the signature where EQ and N are two binary relational symbols which will encode “equal” and “not equal”, respectively. Note that a τ -structure $\mathbb{A}$ belongs to $\text{CSP}(\mathbb{N}, =, \neq)$ if and only if the structure obtained from $\mathbb{A}$ by contracting all connected components of the edge relation $EQ^{\mathbb{A}}$ contains no loop $N(x, x)$. Equivalently, $\mathbb{A} \in \text{CSP}(\mathbb{N}, =, \neq)$ if and only if $N^{\mathbb{A}}$ contains no loop (x, x) and no *contradicting cycle*, i.e., vertices $x_1, \dots, x_n$ such that $EQ(x_i, x_{i+1})$ for all $i \in [n-1]$ and $N(x_1, x_n)$. Despite the fact that the existence of such a cycle is not a first-order property, the problem $\text{CSP}(\mathbb{N}, =, \neq)$ is in HerFO. For a positive integer k we write $d_{EQ}(x) = k$ for the first-order formula stating “ x has exactly k neighbors in the relation EQ (different from x)”, and $d_{EQ}(x) \neq k$ for its negation. Now, consider the formula

$$\phi := \forall x, y \exists z (\neg N(x, y) \vee d_{EQ}(x) \neq 1 \vee d_{EQ}(y) \neq 1 \vee (d_{EQ}(z) \neq 2 \wedge z \notin \{x, y\})).$$

If $\mathbb{A}$ does not hereditarily model ϕ , then there is a subset $\{a_1, \dots, a_n\}$ such that $N(a_1, a_n)$ and every a_i has degree exactly 2 in the relation EQ except for a_1 and a_n , and so $\mathbb{A}$ contains a contradicting cycle. Conversely, if $\mathbb{A}$ has a contradicting cycle one can find a substructure $\mathbb{B}$ of $\mathbb{A}$ that does not model ϕ , namely, any shortest contradicting cycle. Therefore, $\text{CSP}(\mathbb{N}, =, \neq)$ is hereditarily defined by the first-order sentence $\phi \wedge \forall x. \neg N(x, x)$.

Some further examples of polynomial-time solvable CSPs in HerFO include a finite-domain CSP in HerFO but not in FO (Example 25), and an infinite-domain CSP in HerFO but not even in Datalog (Example 27).

Hard examples

Here, we present an example of a coNP-complete CSP in HerFO that we will also use in Corollaries 24 and 21; to see another (possibly better-known) one see the full version of this paper.

Let τ consist of two binary symbols E_b and E_r ; we think of a τ -structure is a digraph with blue and red edges. For a positive integer $n \geq 2$ we denote by $\mathbb{T}\mathbb{D}_n$ the structure with vertex set $[n]$ such that $([n], E_b)$ is a directed cycle $1, \dots, n$, and $([n], E_r)$ is a complete (symmetric) graph. Let $\mathcal{T}$ be the set containing the one-element structure with a red loop, and the one-element structure with a blue loop, and all structures $\mathbb{T}\mathbb{D}_n$ for $n \geq 2$. Let $\text{Forb}(\mathcal{T})$ be the class of finite τ -structures $\mathbb{A}$ for which there is no homomorphism $\mathbb{T} \rightarrow \mathbb{A}$ for any $\mathbb{T} \in \mathcal{T}$. Clearly, $\text{Forb}(\mathcal{T})$ is closed under inverse homomorphisms. Since

all structures in $\mathcal{T}$ are connected, $\text{Forb}(\mathcal{T})$ is also closed under disjoint unions, and hence $\text{Forb}(\mathcal{T}) = \text{CSP}(\mathbb{B}_{\mathcal{T}})$ for some τ -structure $\mathbb{B}_{\mathcal{T}}$ (see, e.g., [2, Lemma 1.1.8]).³ We show that $\text{CSP}(\mathbb{B}_{\mathcal{T}})$ is a coNP-complete CSP in HerFO. Consider the τ -sentence

$$\phi_{\mathcal{T}} := \exists x, y \forall z (\neg E_b(z, z) \wedge \neg E_r(z, z) \wedge (\neg E_b(x, z) \vee (x \neq y \wedge \neg E_r(x, y)))).$$

Clearly, no loop satisfies $\phi_{\mathcal{T}}$. We now show that no structure $\mathbb{T}\mathbb{D}_n$, for $n \geq 2$, satisfies $\phi_{\mathcal{T}}$. Consider a pair of vertices $x, y \in \mathbb{T}\mathbb{D}_n$. Firstly, there is some z such that $(x, z) \in E_b^{\mathbb{T}\mathbb{D}_n}$, because the blue edges define a directed cycle in $\mathbb{T}\mathbb{D}_n$. Secondly, if $x \neq y$, then $(x, y) \in E_r^{\mathbb{T}\mathbb{D}_n}$ because the red edges induce a red clique. Hence, $\mathbb{T}\mathbb{D}_n$ does not satisfy the last conjunct of $\phi_{\mathcal{T}}$. Also, since $\phi_{\mathcal{T}}$ is negative and connected, it follows by Observation 5 that if $\mathbb{F} \rightarrow \mathbb{A}$ for some structure $\mathbb{F} \in \mathcal{T}$, then $\mathbb{A}$ does not hereditarily satisfy $\phi_{\mathcal{T}}$. On the other hand, observe that if a τ -structure $\mathbb{A}$ does not hereditarily satisfy $\phi_{\mathcal{T}}$, then $\mathbb{A}$ contains a loop, or there is a subset $A' \subseteq A$ such that $(A', E_r^{\mathbb{A}'})$ is a complete symmetric graph with at least two vertices, and every vertex $x \in A'$ has a blue out-neighbour. So if $\mathbb{A}$ contains no loops, then the shortest directed blue cycle in A' induces a structure isomorphic to $\mathbb{T}\mathbb{D}_n$ for some $n \geq 2$. Therefore, $\text{Her}(\phi_{\mathcal{T}}) = \text{Forb}(\mathcal{T}) = \text{CSP}(\mathbb{B}_{\mathcal{T}})$.

► **Theorem 8.** *$\text{Forb}(\mathcal{T})$ is a coNP-complete CSP hereditarily definable by an $\exists\exists\forall$ - and by an $\exists\forall\exists$ -sentence.*

Proof. It is easy to see that $\phi_{\mathcal{T}}$ is equivalent to the sentence $\phi'_{\mathcal{T}}$ obtained from $\phi_{\mathcal{T}}$ by changing the prefix $\exists x, y \forall z$ to $\exists x \forall z \exists y$. Hence, it follows from the discussion above that $\text{Forb}(\mathcal{T})$ is hereditarily definable by an $\exists\exists\forall$ - and by an $\exists\forall\exists$ -sentence. We now show that $\text{Forb}(\mathcal{T})$ is coNP-complete. Consider an instance ψ of 3SAT with variables V and clauses $C_1, \dots, C_m$, where $C_i = (c_i^1, c_i^2, c_i^3)$ and $c_i^k \in \{v, \neg v\}$ for some $v \in V$. We construct a τ -structure $\mathbb{A}$ with vertices a_i^j for each $i \in [m]$ and $j \in [3]$. The blue edges of $\mathbb{A}$ consist of all pairs (a_i^j, a_{i+1}^k) and (a_m^j, a_1^k) , for $i \in [m-1]$ and $j, k \in [3]$; the red edges of $\mathbb{A}$ correspond to the relation $c_i^k \neq \neg c_j^l$ with $i \neq j$, i.e., $(a_i^k, a_j^l) \in E_r^{\mathbb{A}}$ if and only if the literal c_i^k does not equal the negation of the literal $\neg c_j^l$ – in particular, $E_r^{\mathbb{A}}$ is a symmetric relation without loops.

We claim that ψ is satisfiable if and only if $\mathbb{A} \notin \text{Forb}(\mathcal{T})$. Suppose there is a satisfying assignment for ψ , and consider the vertices $a_i^{k_i}$ where $c_i^{k_i}$ is true in the clause C_i . Then the substructure $\mathbb{A}'$ with domain $a_1^{k_1}, \dots, a_m^{k_m}$ satisfies that every vertex has a blue out-neighbour. Clearly, $c_i^{k_i}$ cannot be the negation of $c_j^{k_j}$, so $(A', E_r^{\mathbb{A}'})$ is a complete red graph. This shows that if ψ is satisfiable, then there is a homomorphism $\mathbb{T}\mathbb{D}_m \rightarrow \mathbb{A}$.

Conversely, suppose that $\mathbb{A} \notin \text{Forb}(\mathcal{T})$. Notice that if there is a substructure $\mathbb{A}'$ of $\mathbb{A}$ that satisfies that every vertex has a blue out-neighbour, then A' contains a vertex $a_i^{k_i}$ for each $i \in [m]$. Moreover, if $(A', E_r^{\mathbb{A}'})$ is a complete graph, then A' contains at most one vertex $a_i^{k_i}$ for each $i \in [m]$. Then the evaluation $f: V \rightarrow \{0, 1\}$ defined by $f(v) = 1$ if there is some clause C_i such that $v = c_i^k$ and $a_i^k \in A'$ shows that ψ is satisfiable. ◀

Inexpressible examples

In this section we study the limitations of the expressive power of HerFO. Clearly, a class $\mathcal{C}$ is first-order definable if and only if the complement of $\mathcal{C}$ is first-order definable. A structure $\mathbb{A}$ is a *minimal obstruction* of a hereditary class $\mathcal{C}$ if $\mathbb{A} \notin \mathcal{C}$ but every proper substructure $\mathbb{A}'$ of $\mathbb{A}$ belongs to $\mathcal{C}$. We show that $\mathcal{C} \in \text{HerFO}$ if and only if the complement of $\mathcal{C}$ contains a first-order definable subclass $\mathcal{F}'$ that contains all minimal obstructions of $\mathcal{C}$ (Lemma 9).

³ The reader familiar with Fraïssé limits [16] may notice that $\mathbb{B}_{\mathcal{T}}$ can be chosen to be a countable *homogeneous* structure, i.e., such that every isomorphism between finite substructures of $\mathbb{B}_{\mathcal{T}}$ can be extended to an automorphism of $\mathbb{B}_{\mathcal{T}}$.

We then apply this observation to show that the class of bipartite graphs is not in HerFO (Example 10). It is well-known that a graph is bipartite if and only if it does not contain an odd cycle, and that the class of odd cycles cannot be expressed by a first-order formula. However, this is not enough to show that the class of bipartite graphs is not in HerFO: there are properties $\mathcal{F}$ that are not first-order definable, but the class of all $\mathcal{F}$ -free structures is in HerFO; see Example 7. To prove that the class of bipartite graphs is not in HerFO, we therefore need the following lemma which we prove in Appendix C.

► **Lemma 9.** *Let $\mathcal{C}$ be a hereditary class of finite τ -structures and let $\mathcal{F}$ be the class of minimal obstructions of $\mathcal{C}$. Then $\mathcal{C}$ is hereditarily first-order definable if and only if there is a first-order sentence ψ such that*

- *if $\mathbb{F} \in \mathcal{F}$, then $\mathbb{F} \models \psi$, and*
- *if $\mathbb{A}$ is a finite τ -structure such that $\mathbb{A} \models \psi$, then there is an embedding $\mathbb{F} \hookrightarrow \mathbb{A}$ for some $\mathbb{F} \in \mathcal{F}$.*

Building on this simple lemma we can now use standard Ehrenfeucht-Fraïssé arguments to show that certain hereditary classes are not in HerFO. If $\mathbb{A}$ and $\mathbb{A}'$ are τ -structures, we write $\mathbb{A} \equiv_k \mathbb{A}'$ if $\mathbb{A}$ and $\mathbb{A}'$ satisfy the same first-order τ -sentences with at most k variables.

► **Example 10 (Bipartite graphs not in HerFO).** It is well known that the minimal obstructions of the class of bipartite graphs are all odd symmetric cycles and the non-symmetric edge. For every positive integer k , there is a large enough odd cycle $\mathbb{C}$ and a large enough even cycle $\mathbb{C}'$ such that $\mathbb{C} \equiv_k \mathbb{C}'$ (this can be shown by an Ehrenfeucht-Fraïssé argument; see, e.g., [8]). Hence, we conclude via Lemma 9 that the class of bipartite graphs is not in HerFO. Moreover, note that $\text{CSP}(K_2)$ is the class of bipartite digraphs. So it also follows from the existence of such cycles $\mathbb{C} \equiv_k \mathbb{C}'$ and Lemma 9 that $\text{CSP}(K_2)$ is not in HerFO.

In the appendix we use Lemma 9 to show that $\text{CSP}(\mathbb{Q}, <, =)$ is not in HerFO (Example 29). Another method for proving that certain CSPs are not expressible in HerFO is using ω -categoricity and Remark 6. We present an example in the appendix (Example 26).

4 Hereditary model checking and quantifier prefixes

A *quantifier prefix* is a word $Q \in \{\exists, \forall\}^*$; we say that a first-order formula ϕ in prenex normal form has quantifier prefix Q if $\phi = Q_1 x_1 \dots Q_n x_n. \psi$ where $Q = Q_1 \dots Q_n$ and ψ is a quantifier-free formula. In this section we present the following dichotomy for quantifier prefixes: for every quantifier prefix $Q \in \{\exists, \forall\}^*$ either

- $\text{Her}(\phi)$ is in P for every first-order formula ϕ with quantifier prefix Q , or
- there exists a first-order formula ϕ with quantifier prefix Q such that $\text{Her}(\phi)$ is coNP-complete.

Moreover, we show that in the former case, $\text{Her}(\phi)$ is expressible by an ESO sentence Ψ whose first-order part is universal.

A relational signature τ is called *monadic* if all relation symbols in τ are monadic. It is easy to see that every problem in HerFO with a monadic signature is in FO, and so polynomial-time solvable. We postpone its simple proof to Appendix C.

► **Proposition 11.** *Let τ be a finite monadic relational signature. For every first-order formula ϕ the class $\text{Her}(\phi)$ is universally definable and hence in P.*

From now on, we only consider the non-monadic case. The key components in the proof of our classification (Theorem 20) are Algorithm 1 (for one of the tractable cases) and the fact that the problem of deciding whether every directed cycle in an input digraph $\mathbb{D}$ induces a symmetric edge is coNP-complete (Theorem 18) and expressible in HerFO (Lemma 17).

The $\forall^*\exists^*$ fragment

In this section we prove that for every $\forall^*\exists^*$ -formula ϕ there is a universal formula ϕ' such that a structure $\mathbb{A}$ hereditarily satisfies ϕ if and only if $\mathbb{A} \models \phi'$.

► **Lemma 12.** *Let ϕ be a $\forall^*\exists^*$ -formula with k universally quantified variables. Then a structure $\mathbb{A}$ hereditarily models ϕ if and only if every k -element substructure of $\mathbb{A}$ models ϕ .*

Proof. We prove the non-trivial (but straightforward) implication. Suppose that every substructure $\mathbb{B}$ of $\mathbb{A}$ with $|B| \leq k$ models ϕ , and let $\mathbb{A}'$ be a substructure of $\mathbb{A}$. If $|A'| \leq k$, then $\mathbb{A}' \models \phi$; otherwise, for a k -tuple $(a_1, \dots, a_k) \in (A')^k$, let $\bar{b}$ be a tuple such that the quantifier-free part ψ of ϕ is true of $(a_1, \dots, a_k, \bar{b})$ in the substructure of $\mathbb{A}'$ with vertex set $\{a_1, \dots, a_k\}$. It follows that $\mathbb{A}' \models \psi(a_1, \dots, a_k, \bar{b})$, and since such a $\bar{b}$ exists for every $\bar{a} \in (A')^k$, we conclude that $\mathbb{A}' \models \phi$, and therefore $\mathbb{A} \in \text{Her}(\phi)$. ◀

► **Corollary 13.** *If ϕ is a $\forall^*\exists^*$ -sentence, then $\text{Her}(\phi)$ is universally definable and hence polynomial-time solvable.*

Proof. If $\phi = \forall x_1, \dots, x_k \exists y_1, \dots, y_l. \psi$ where ψ is quantifier-free, let ϕ' be the formula

$$\forall x_1, \dots, x_k \exists y_1 \in \{x_1, \dots, x_k\}, \dots, y_l \in \{x_1, \dots, x_k\}. \psi.$$

By Lemma 12, a structure hereditarily satisfies ϕ if and only if it satisfies ϕ' . ◀

The $\forall^*\exists\forall^*$ fragment

An *SNP τ -sentence* (short for *strict non-deterministic polynomial-time*) is a sentence of the form

$$\exists R_1, \dots, R_k \forall x_1, \dots, x_n. \psi$$

where ψ is a quantifier free $\tau \cup \{R_1, \dots, R_k\}$ -formula [10, 19]. If a structure $\mathbb{A}$ satisfies the sentence Ψ , we write $\mathbb{A} \models \Psi$. We say that a class of finite τ -structures $\mathcal{C}$ is in *SNP* if there exists an SNP τ -sentence Φ such that $\mathbb{A} \models \Phi$ if and only if $\mathbb{A} \in \mathcal{C}$. We show that if ϕ is a $\forall^*\exists\forall^*$ -formula, then $\text{Her}(\phi)$ is in $\text{SNP} \cap \text{P}$.

In the following we consider a fixed $\forall^*\exists\forall^*$ -formula

$$\phi = \forall x_1, \dots, x_l \exists y \forall x_{l+1}, \dots, x_n. \psi(x_1, \dots, x_l, y, x_{l+1}, \dots, x_n)$$

where ψ is a quantifier-free τ -formula. We expand τ with an $(l+2)$ -ary relation symbol L . We will interpret L as a reflexive linear order with l parameters, i.e., if the first l arguments of L are fixed, then the binary relation defined by the remaining two free variables is a linear order. It is straightforward to observe that there is a universal $\{L\}$ -formula $\text{Lin}(x_1, \dots, x_l)$ such that Lin is true of an l -tuple $\bar{a}$ in an $\{L\}$ -structure $\mathbb{A}$ if and only if the binary relation $L(\bar{a}, x, y)$ defines a reflexive linear order $x \leq_{\bar{a}} y$ on A . Consider now the SNP sentence defined as follows.

$$\begin{aligned} \Phi := \exists L \forall x_1, \dots, x_l, y, x_{l+1}, \dots, x_n. & \text{Lin}(x_1, \dots, x_l) \\ & \wedge \left(\bigwedge_{i \in [n]} L(x_1, \dots, x_l, y, x_i) \right) \Rightarrow \psi(x_1, \dots, x_l, y, x_{l+1}, \dots, x_n). \end{aligned}$$

► **Lemma 14.** *A finite τ -structure $\mathbb{A}$ hereditarily models ϕ if and only if it models Φ .*

Proof. For the easy direction, suppose that $(\mathbb{A}, L)$ models the first-order part of Φ . For all $a_1, \dots, a_l \in A$, let $b \in A$ be the minimum with respect to the linear order $\leq_{a_1, \dots, a_l}$, i.e., the element $b \in A$ such that $L(a_1, \dots, a_l, b, c)$ for all $c \in A$. In particular, for all $a_{l+1}, \dots, a_n \in A$ and $i \in [n]$ the atomic formula $L(a_1, \dots, a_l, b, a_i)$ holds in $\mathbb{A}$, and thus $\mathbb{A} \models \psi(a_1, \dots, a_l, b, a_{l+1}, \dots, a_n)$. Since the first-order part of Φ is universal, every substructure $\mathbb{B}$ of $\mathbb{A}$ also models Φ , and by the previous argument we conclude that $\mathbb{B}$ models ϕ . Hence, $\mathbb{A}$ hereditarily models ϕ .

Conversely, suppose that $\mathbb{A}$ hereditarily satisfies ϕ . For every l -tuple $\bar{a} = (a_1, \dots, a_l)$ of A we define a reflexive linear order $\leq_{\bar{a}}$ such that the expansion

$$(\mathbb{A}, \{(a_1, \dots, a_l, b, c) : b \leq_{(\bar{a})} c\})$$

models the first-order part of Φ . We define elements $b_1, \dots, b_m$ inductively as follows. Let $b_1 \in A$ be any element witnessing that $\mathbb{A}$ satisfies

$$\exists y \forall x_{l+1}, \dots, x_n. \psi(\bar{a}, y, x_{l+1}, \dots, x_n).$$

For $l > 1$, if no b_i is a coordinate of $\bar{a}$ for $i < l$, choose b_l to be any element witnessing that the substructure of $\mathbb{A}$ with domain $A \setminus \{b_1, \dots, b_{l-1}\}$ satisfies $\exists y \forall x_{l+1}, \dots, x_n. \psi(\bar{a}, y, x_{l+1}, \dots, x_n)$ (such a vertex b_l exists since $\mathbb{A}$ hereditarily satisfies ϕ). Otherwise, if some b_i equals some coordinate of $\bar{a}$, then let b_l be an arbitrary element of $A \setminus \{b_1, \dots, b_{l-1}\}$. We define the linear ordering $b_i \leq_{\bar{a}} b_j$ if and only if $i \leq j$, and let $L := \{(\bar{a}, b, c) \in A^{l+2} : b \leq_{\bar{a}} c\}$. It follows from the definition of L and of Lin that $(\mathbb{A}, L) \models \forall a_1, \dots, a_l. \text{Lin}(a_1, \dots, a_l)$.

Suppose that $(\mathbb{A}, L)$ satisfies $\bigwedge_{i \in [n]} L(a_1, \dots, a_l, b, a_i)$ for some $a_1, \dots, a_l, b, a_{l+1}, \dots, a_n \in A$. Let $b_1, \dots, b_m$ be the enumeration of A corresponding to the linear ordering $\leq_{(a_1, \dots, a_l)}$, and suppose that $b = b_l$. Since $(\mathbb{A}, L) \models \bigwedge_{i \in [n]} L(a_1, \dots, a_l, b, a_i)$, i.e., $b \leq_{(a_1, \dots, a_l)} a_i$ for every $i \in [n]$, it must be the case that every a_i belongs to $A \setminus \{b_1, \dots, b_{l-1}\}$. It thus follows from the definition of b_l that the substructure $\mathbb{B}$ of $\mathbb{A}$ with vertex set $A \setminus \{b_1, \dots, b_{l-1}\}$ models $\psi(a_1, \dots, a_l, b_l, a_{l+1}, \dots, a_n)$, and thus $(\mathbb{A}, L) \models \psi(a_1, \dots, a_l, b_l, a_{l+1}, \dots, a_n)$. This shows that $\mathbb{A}$ satisfies Φ . $\blacktriangleleft$

The proof of Lemma 14 suggests a polynomial-time algorithm that on input structure $\mathbb{A}$ finds parameterized linear orderings proving that $\mathbb{A}$ hereditarily satisfies ϕ , or finds a substructure $\mathbb{B}$ of $\mathbb{A}$ that does not model ϕ (which certifies that $\mathbb{A}$ does not hereditarily satisfy ϕ). See Figure 1 for an illustration.

Certifying polynomial-time algorithm. Consider a fixed first-order sentence

$$\phi := \forall x_1, \dots, x_l \exists y. \phi'(x_1, \dots, x_l, y)$$

where ϕ' is a universal formula.

► **Lemma 15.** *For every first-order formula $\phi := \forall x_1, \dots, x_l \exists y. \phi'(x_1, \dots, x_l, y)$ where ϕ' is a universal formula $\forall x_{l+1}, \dots, x_n. \psi$ for some quantifier-free formula ψ , and for every τ -structure $\mathbb{A}$ with domain $A = \{a_1, \dots, a_m\}$ the following statements hold.*

- *If Algorithm 1 returns a substructure $\mathbb{A}'$ of $\mathbb{A}$, then $\mathbb{A}' \models \neg\phi$, and $\mathbb{A} \notin \text{Her}(\phi)$.*
- *If Algorithm 1 returns an expansion $(\mathbb{A}, L)$ of $\mathbb{A}$, then $(\mathbb{A}, L)$ satisfies the first-order part of Φ , and $\mathbb{A} \in \text{Her}(\phi)$.*

Proof. To prove the first claim, notice that (by finite induction) at the **if** statement in the **repeat-until** loop of the algorithm, the set A' contains all entries of $\bar{a}$. Hence, if the algorithm returns $\mathbb{A}'$, then $\mathbb{A}'$ does not satisfy $\exists y. \phi'(\bar{a}, y)$ and hence $\mathbb{A}$ does not hereditarily

■ **Algorithm 1** Cert-Her- $\forall^* \exists \forall^*$.

input : a finite τ -structure $\mathbb{A}$ with domain $A = \{a_1, \dots, a_m\}$.
foreach $\bar{a} \in A^l$ **do**
 $S := \emptyset$ and $\leq := \emptyset$.
 repeat
 Let $\mathbb{A}'$ be the substructure of $\mathbb{A}$ with domain $A' := A \setminus S$.
 if $\mathbb{A}'$ satisfies $\forall y. \neg \phi'(\bar{a}, y)$ **then**
 Return $\mathbb{A}'$
 Let $s \in A'$ be such that $\mathbb{A}'$ satisfies $\phi'(\bar{a}, s)$.
 Update $S := S \cup \{s\}$ and $\leq := \leq \cup \{(x, s) \mid x \in S\}$.
 until s equals some coordinate of $\bar{a}$
 Let $\leq_{\bar{a}} := \leq \cup S \times (A \setminus S) \cup \{(a_i, a_j) \in (A \setminus S)^2 \mid i \leq j\}$.
Let $L := \{(\bar{a}, b, c) \in A^{l+2} : \bar{a} \in A^l, b \leq_{\bar{a}} c\}$.
Return $(\mathbb{A}, L)$

satisfy ϕ . Now we argue that the second itemized statement holds. It is straightforward to observe that $\leq_{\bar{a}}$ is a reflexive linear order for every $\bar{a} \in A^l$. Hence, it follows from the definition of L that $(\mathbb{A}, L) \models \forall x_1, \dots, x_l. \text{Lin}(x_1, \dots, x_l)$. To see that $(\mathbb{A}, L)$ models the second conjunct of Φ , let $\bar{b} \in A^l$, $c \in A$, and $b_{l+1}, \dots, b_n \in A$ (so $(\bar{b}, c, b_{l+1}, \dots, b_n)$ is an evaluation of the universally quantified variables of Φ in A). Further, suppose that

$$(\mathbb{A}, L) \models \bigwedge_{i \in [n]} L(b_1, \dots, b_l, c, b_i);$$

otherwise the second conjunct in the definition of Φ is vacuously true for the tuple $(b_1, \dots, b_l, c, b_{l+1}, \dots, b_n)$. By the definition of L , this means that $c \leq_{\bar{b}} b_i$ for every $i \in [n]$. Hence, there is some iteration of the **repeat-until** loop such that $c \in S$. Let S_i be the set S and the end of this iteration i of the loop, and $S_0 := \emptyset$. It follows from the definition of $\leq_{\bar{b}}$ and the assumption that $c \leq_{\bar{b}} b_j$ for each $j \in [n]$, that if $S_i = S_{i-1} \cup \{c\}$, then $(\{b_1, \dots, b_n\} \cap S_{i-1}) = \emptyset$. Since c was added to S in the i -th iteration, it must be the case that in the i -th iteration the **if** statement is not true for $s := c$, i.e., $\mathbb{A}' \models \forall x_1, \dots, x_l. \phi'(x_1, \dots, x_l, c)$. Since $\{b_1, \dots, b_n\} \subseteq A \setminus S_{i-1}$, we conclude that in particular $\mathbb{A}' \models \phi'(\bar{b}, c)$ where the universally quantified variables from ϕ' are interpreted as $(b_{l+1}, \dots, b_n)$. This means that

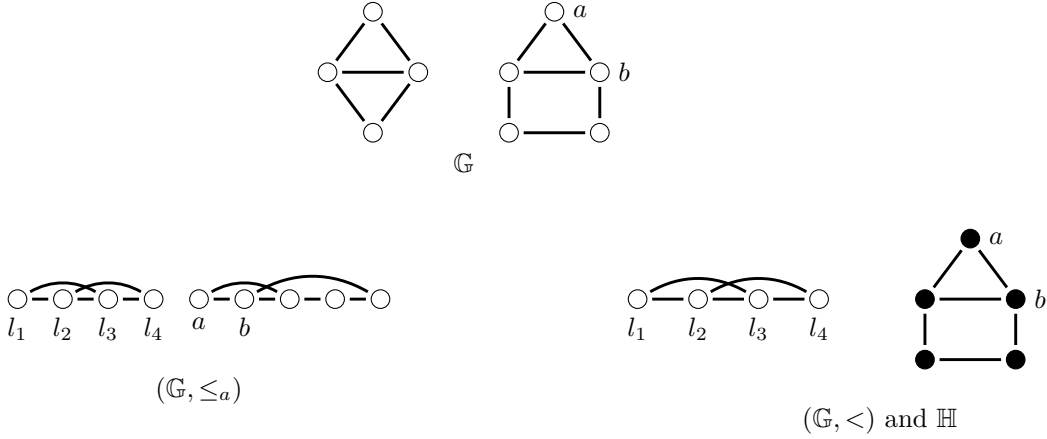
$$(\mathbb{A}, L) \models \left(\bigwedge_{i \in [n]} L(b_1, \dots, b_l, c, b_i) \right) \Rightarrow \psi(b_1, \dots, b_l, c, b_{l+1}, \dots, b_n).$$

Since this is true for any choice of elements $b_1, \dots, b_n$ and c in A , we conclude that $(\mathbb{A}, L)$ satisfies the first-order part of Φ , and by Lemma 14 we conclude that $\mathbb{A} \in \text{Her}(\phi)$. ◀

Clearly, Algorithm 1 runs in polynomial time in the representation size of $\mathbb{A}$. Hence, the following statement is an immediate consequence of Lemma 15.

► **Theorem 16.** *If ϕ is a $\forall^* \exists \forall^*$ -sentence, then there is an SNP sentence Φ such that a finite structure $\mathbb{A}$ satisfies Φ if and only if $\mathbb{A} \in \text{Her}(\phi)$. Moreover, Φ can be efficiently computed from ϕ , and there is a polynomial-time algorithm that either finds an expansion of $\mathbb{A}$ proving that $\mathbb{A} \models \Phi$, or finds a substructure $\mathbb{A}'$ of $\mathbb{A}$ such that $\mathbb{A} \models \neg \phi$.*

Note that this theorem covers the first-order sentences that show that Example 1, Example 2, Example 3, and Example 27 are in HerFO, because they are $\exists \forall^*$ sentences. Hence, the polynomial-time tractability of each of these problems follows from Theorem 16.



■ **Figure 1** Consider a first-order $\{E\}$ -sentence ϕ that states that E is a symmetric relation, and that for every vertex x there is a vertex y not adjacent to x such that the neighbourhood of y induces a clique. Clearly, ϕ can be chosen to be an $\forall\exists\forall^*$ formula. On the top, we depict a graph $\mathbb{G}$ with two distinguished vertices a and b . At the bottom left, we depict the linear order $\leq_a$ (from left to right) of $\mathbb{G}$ obtained via Algorithm 1 (where the linear order between the vertices greater than a can be arbitrary), and that proves that $(\mathbb{G}, a)$ hereditarily satisfies $\phi(a, y, \bar{z})$. At the bottom right, we depict the partial linear order $<$ and the subgraph $\mathbb{H}$ (depicted with black vertices) of $\mathbb{G}$ that Algorithm 1 finds when running the loop for $x = b$, which proves that $(\mathbb{G}, b)$ does not hereditarily satisfy ϕ .

$\exists\forall\exists$ - and $\exists^2\forall$ -sentences

Theorem 8 already provides examples of first-order sentences ϕ and ψ with quantifier prefix $\exists\forall\exists$ and $\exists^2\forall$ such that $\text{Her}(\phi)$ and $\text{Her}(\psi)$ are coNP-complete. Moreover, ϕ and ψ are formulas with a binary signature (with two relation symbols). In this section we provide first-order sentences with the same quantifier prefixes over the signature of digraphs (they only use *one* binary symbol), which is the final ingredient needed to prove Theorem 20.

Consider the class of digraphs $\mathbb{D}$ such that for every directed cycle $d_1, \dots, d_n$ of $\mathbb{D}$ there exist $i, j \in [n]$ such that $d_i d_j$ is a symmetric edge of $\mathbb{D}$, i.e., $(d_i, d_j), (d_j, d_i) \in \mathbb{D}$. In this case, we say that every directed cycle of $\mathbb{D}$ ‘induces a symmetric edge’.

► **Lemma 17.** *Every directed cycle of a finite digraph $\mathbb{D}$ induces a symmetric edge if and only if $\mathbb{D}$ hereditarily satisfies the sentence*

$$\exists x, y \forall a (\neg E(x, a) \vee [E(x, y) \wedge E(y, x)]). \quad (1)$$

Proof. Suppose there is a directed cycle $d_1, \dots, d_n$ of $\mathbb{D}$ that does not induce a symmetric edge. Then the substructure of $\mathbb{D}$ with vertex set $\{d_1, \dots, d_n\}$ satisfies the formula

$$\forall x, y \exists a (E(x, a) \wedge [\neg E(x, y) \vee \neg E(y, x)]),$$

and so $\mathbb{D}$ does not hereditarily satisfy (1).

Conversely, suppose that every directed cycle of $\mathbb{D}$ induces a symmetric edge and let $B \subseteq D$. If B contains a sink b , then $B \models \forall a \neg E(b, a)$, and so B satisfies (1). Otherwise, B contains a directed cycle, and by assumption this directed cycle induces a symmetric edge uv , so by letting $x = u$ and $y = v$ we conclude that B satisfies $\exists x, y (E(x, y) \wedge E(y, x))$. Again, B satisfies (1). ◀

Now, we prove that deciding whether every directed cycle of an input digraph $\mathbb{D}$ induces a symmetric edge is coNP-complete.

► **Theorem 18.** *The problem of deciding whether every directed cycle in an input digraph $\mathbb{D}$ induces a symmetric edge is coNP-complete.*

Proof. Let $\mathcal{C}$ be the class of digraphs described in this theorem, i.e., the class of all finite digraphs such that every directed cycle induces a symmetric edge. Fix $\mathbb{X}$ to be a no-instance to this class, e.g., $\mathbb{X} = \tilde{\mathbb{C}}_3$. We reduce from the coNP-hard CSP in Theorem 8 (deciding membership in $\text{Forb}(\mathcal{T})$) to deciding membership to $\mathcal{C}$. First, on an input $\mathbb{A}$ to $\text{Forb}(\mathcal{T})$, we check whether $\mathbb{A}$ contains a blue loop, a red loop, or $\mathbb{TD}_2$ (i.e., a pair of vertices connected by symmetric pairs of blue and red edges). If so, the reduction returns a $\mathbb{X}$. Otherwise, the reduction returns the following digraph $\mathbb{D} = f(\mathbb{A})$. The vertex set D of $\mathbb{D}$ equals the domain A of $\mathbb{A}$, there are no loops $(u, u) \in E(\mathbb{D})$, and for a pair of distinct vertices u and v we do the following.

- If at least one of (u, v) or (v, u) is not a red edge in $\mathbb{A}$, then we connect u and v by a symmetric pair of edges (u, v) and (v, u) in $E(\mathbb{D})$.
- If $(u, v), (v, u)$ is a symmetric pair of red edges in $\mathbb{A}$, and (u, v) is a blue edge in $\mathbb{A}$, then we connect them by a non-symmetric edge (u, v) in $E(\mathbb{D})$.
- If $(u, v), (v, u)$ is a symmetric pair of red edges in $\mathbb{A}$, and u and v are not connected by blue edges in $\mathbb{A}$, we do not connect u and v by any edge in $\mathbb{D}$.

Clearly, the edge relation $E(\mathbb{D})$ admits a quantifier-free definition in $\mathbb{A}$, and so, if $\mathbb{A}$ is a substructure of $\mathbb{B}$, then $f(\mathbb{A})$ is a substructure of $f(\mathbb{B})$. Clearly, for every $\mathbb{TD}_n$ with $n \geq 3$, the digraph $f(\mathbb{TD}_n)$ is a directed cycle that does not induce any symmetric edge. Conversely, if $f(\mathbb{A})$ is a directed cycle on n vertices, then $\mathbb{A} \cong \mathbb{TD}_n$. It is straightforward to observe that a loopless $\{E_b, E_r\}$ -structure $\mathbb{A}$ with no $\mathbb{TD}_2$ belongs to $\text{Forb}(\mathcal{T})$ if and only if it does not embed any $\mathbb{TD}_n$ for $n \geq 3$. Hence, we obtained a polynomial-time reduction from deciding $\text{Forb}(\mathcal{T})$ to deciding membership to $\mathcal{C}$. Therefore, the coNP-hardness of $\mathcal{C}$ follows from Theorem 8. ◀

► **Corollary 19.** *There are $\exists^2\forall$ and $\exists\forall\exists$ (digraph) formulas ϕ and ψ such that $\text{Her}(\phi)$ and $\text{Her}(\psi)$ are coNP-complete.*

Proof. For ϕ consider the formula from Lemma 17 and for ψ consider the quantifier reordering $\exists x\forall a\exists y$ of ϕ , and notice that ϕ and ψ are logically equivalent. ◀

Prefix classification

The classification for general relational signatures follows from the lemmas proved earlier in this section.

► **Theorem 20.** *Let τ be a relational signature which is not monadic. For every quantifier prefix $Q \in \{\exists, \forall\}^*$ one of the following statements hold:*

- *Q is of the form $\forall^*\exists^*$, or of the form $\forall^*\exists\forall^*$, and in this case $\text{Her}(\phi)$ is in P for every first-order τ -sentence ϕ with quantifier prefix Q , or*
- *Q contains a subword $\exists\exists\forall$ or $\exists\forall\exists$, and in this case there is a first-order τ -sentence ϕ with quantifier prefix Q such that $\text{Her}(\phi)$ is coNP-complete.*

Proof. Clearly, both items describe disjoint and complementary cases. The claim in the first item follows from Corollary 13 and Theorem 16. If τ contains a binary relation symbol, then the claim in the second item follows from Corollary 19. Otherwise, τ must contain a relation R of arity at least three; however, we can use R to model a binary relation, so the claim also holds in this case. ◀

This classification also holds for the intersection of HerFO and CSPs (via Theorem 8).

- **Corollary 21.** *For every quantifier prefix $Q \in \{\exists, \forall\}^*$ one of the following statements hold:*
- *Q is of the form $\forall^* \exists^*$, or of the form $\forall^* \exists \forall^*$, and in this case if $\text{CSP}(\mathbb{A}) = \text{Her}(\phi)$ for some structure $\mathbb{A}$ and some first-order sentence ϕ with quantifier prefix Q , then $\text{CSP}(\mathbb{A})$ is in P, or*
 - *Q contains a subword $\exists \exists \forall$ or $\exists \forall \exists$, and in this case there is a structure $\mathbb{B}$ such that $\text{CSP}(\mathbb{B}) = \text{Her}(\phi)$ for some first-order sentence ϕ with quantifier prefix Q , and $\text{CSP}(\mathbb{B})$ is coNP-complete.*

5 Undecidability of the tractability problem

The *tractability problem* for HerFO asks whether $\text{Her}(\phi)$ can be solved in polynomial time for a given first-order sentence ϕ . In this section we show that if $P \neq NP$, then the tractability problem for HerFO is undecidable. We begin with the following simple observation.

► **Observation 22.** *Consider a first-order τ -sentence $\phi := Q_1 x_1 \dots Q_n x_n \cdot \psi(x_1, \dots, x_n)$. If $\xi(x)$ is a first-order τ -formula, then there is a τ -sentence ϕ_ξ such that a τ -structure $\mathbb{A}$ satisfies ϕ_ξ if and only if there is no element $a \in A$ that satisfies ξ or the substructure of $\mathbb{A}$ with domain $\{a \in A : \mathbb{A} \models \xi(a)\}$ satisfies ϕ . Namely, if $Q_1 = \dots = Q_n = \exists$, then*

$$\phi_\xi := \forall y. \neg \xi(y) \vee \exists x_1, \dots, x_n \left(\bigwedge_{i \in [n]} \xi(x_i) \wedge \psi(x_1, \dots, x_n) \right).$$

Otherwise,

$$\phi_\xi := Q_1 x_1 \dots Q_n x_n \left(\bigwedge_{i \in U} \xi(x_i) \implies \bigwedge_{i \in E} \xi(x_i) \wedge \psi(x_1, \dots, x_n) \right)$$

where U (respectively, E) is the set of indices $i \in \{1, \dots, n\}$ such that Q_i is a universal (respectively, existential) quantifier.

Similarly, if U is a monadic predicate and ϕ is a first-order sentence, then $\phi_U := \phi_{U(x)}$ denotes the relativization of ϕ to the vertices in the set U , and $\phi_{\neg U} := \phi_{\neg U(x)}$ the relativization of ϕ to the complement of U .

► **Theorem 23.** *If $P \neq \text{coNP}$, then it is undecidable to test whether the hereditary model-checking problem for a given first-order sentence ϕ is solvable in polynomial time.*

Proof. We use Trakhtenbrot's theorem, which states that there is no algorithm that decides whether a given first-order formula ϕ has a finite model [24]. We reduce this decision problem to our problem.

Let ϕ be a first-order τ -sentence, let U be a monadic predicate not in τ , and let E a binary predicate also not in τ . Let ψ be a first-order $\{E\}$ -sentence such that the hereditary model-checking problem for ψ is coNP-complete, e.g., ψ can be chosen to be the sentence hereditarily describing the CSP from Corollary 19. Building on the relativizations $\psi_{\neg U}$ and $(\neg \phi)_U$ we define the following first-order $\tau \cup \{U, E\}$ -sentence

$$\chi := (\neg \phi)_U \vee \psi_{\neg U}.$$

We claim that if ϕ does not have a finite model, then $\text{Her}(\chi)$ is polynomial-time solvable, and if ϕ has a finite model, then $\text{Her}(\chi)$ is coNP-complete. We first observe that if ϕ does not have a finite model, then χ is valid on all finite $\tau \cup \{U, E\}$ -structures. To see this, let $\mathbb{A}$ be a $\tau \cup \{U, E\}$ -structure. If $U^{\mathbb{A}} = \emptyset$, then $\mathbb{A}$ satisfies the first disjunct $(\neg \phi)_U$ of χ

(Observation 22), and if $U^{\mathbb{A}} \neq \emptyset$, then the substructure induced by $U^{\mathbb{A}}$ also satisfies $(\neg\phi)_U$, because ϕ has no finite models. In particular, this implies that $\text{Her}(\chi)$ can be solved in polynomial-time, because every instance is a yes-instance.

Now suppose that ϕ has a finite model. Let $\mathbb{S}$ be a minimal model of ϕ , i.e., every proper substructure of $\mathbb{S}$ satisfies $\neg\phi$. We present a polynomial-time reduction from $\text{Her}(\psi)$ to $\text{Her}(\chi)$, which implies that $\text{Her}(\chi)$ is coNP-complete. Given an $\{E\}$ -structure $\mathbb{A}$ we consider the $\tau \cup \{U, E\}$ -structure $\mathbb{B}$ defined as follows:

- the domain B of $\mathbb{B}$ is the disjoint union $A \cup S$,
- the interpretation of U in $\mathbb{B}$ is S ,
- the interpretation of E in $\mathbb{B}$ equals $E^{\mathbb{A}}$, and
- for every $R \in \tau$, the interpretation of R in $\mathbb{B}$ is $R^{\mathbb{S}}$.

The structure $\mathbb{B}$ can be computed from the structure $\mathbb{A}$ in polynomial time (the structure $\mathbb{S}$ is constant). We now show that every substructure of $\mathbb{A}$ satisfies ψ if and only if every substructure of $\mathbb{B}$ satisfies χ . First, note that if a substructure $\mathbb{C}$ of $\mathbb{B}$ does not contain $\mathbb{S}$, then $\mathbb{C}$ satisfies χ . Indeed, if $C \subseteq A$, then there is no element of $\mathbb{C}$ that models $U(x)$, so by Observation 22 it follows that $\mathbb{C} \models (\neg\phi)_U$ and hence $\mathbb{C} \models \chi$. Otherwise, if $\emptyset \neq C \cap S \neq S$, then the substructure $\mathbb{C}'$ of $\mathbb{C}$ with domain $U^{\mathbb{C}}$ satisfies $\neg\phi$, because $\mathbb{C}'$ is a proper substructure of $\mathbb{S}$, and $\mathbb{S}$ is the smallest model of ϕ . Hence, $\mathbb{C} \models (\neg\phi)_U$, and so $\mathbb{C} \models \chi$. These observations imply that every substructure $\mathbb{C}$ of $\mathbb{B}$ satisfies χ if and only if every substructure $\mathbb{D}$ of $\mathbb{B}$ with $S \subseteq D$ satisfies χ . If $S = D$, then $U^{\mathbb{D}} = D$, so no element d of $\mathbb{D}$ satisfies $\neg U(d)$, and similarly as above, we conclude that $\mathbb{D} \models \chi$ because $\mathbb{D} \models \psi_{\neg U}$. Finally, we assume that $S \subseteq D$ and $D \neq S$, and so $D \cap A \neq \emptyset$. Since $S \subseteq D$ and $\mathbb{S} \models \phi$, we have that $\mathbb{D}$ does not satisfy $(\neg\phi)_U$. Hence, $\mathbb{D} \models \chi$ if and only if $\mathbb{D} \models \psi_{\neg U}$, and the latter holds if and only if the substructure of $\mathbb{A}$ with domain $D \cap A$ satisfies ψ . We thus conclude that $\mathbb{B} \in \text{Her}(\chi)$ if and only if $\mathbb{A} \in \text{Her}(\psi)$.

All together this shows that if $P \neq \text{coNP}$, then ϕ has a finite model if and only if $\text{Her}(\chi)$ is not solvable in polynomial time. ◀

We say that two quantifier prefixes are *dual* to each other if one can be obtained from the other by exchanging the symbols $\forall$ and $\exists$.

► **Corollary 24.** *Consider a quantifier prefix $Q \in \{\exists, \forall\}^*$, and assume $P \neq \text{NP}$. If the finite satisfiability problem for first-order sentences whose quantifier prefix Q is undecidable, then the tractability problem for HerFO is undecidable for first-order sentences whose quantifier prefix is dual to Q . In particular, we obtain that the tractability problem for HerFO is undecidable if the quantifier prefix of the input may contain $\exists\forall\exists$ or $\exists^3\forall$ as subwords.*

Proof. Denote by Q' the dual of Q . It follows from [7, Theorem 3.0.1] that if the finite satisfiability problem for first-order sentences with quantifier prefix Q is undecidable, then Q contains $\forall\exists\forall$ or $\forall^3\exists$ as a subword. Dually, Q' contains $\exists\forall\exists$ or $\exists^3\forall$ as a subword, and so, by Theorem 8, there is a FO sentence ψ with quantifier prefix Q' such that $\text{Her}(\psi)$ is coNP-complete. Now, we can reduce from the finite satisfiability, because the finite satisfiability problem for the fragments of FO with quantifier prefix $Q = \forall\exists\forall$ or $Q = \forall^3\exists$ is undecidable [7, Theorem 3.0.1]. ◀

6 Conclusion and Open Problems

We introduced the hereditary first-order model checking problem HerFO , and presented a complexity classification for HerFO based on allowed quantifier prefixes. A number of open problems are left for future research.

1. Are there first-order sentences ϕ such that $\text{Her}(\phi)$ is coNP-intermediate (assuming $\text{P} \neq \text{NP}$)? We conjecture that there are.
2. Is every finite-domain CSP which is in HerFO also in P? If $\text{coNP} \neq \text{NP}$, then it follows from the P vs. NP-complete finite-domain CSP dichotomy that this answer has a positive question. We ask for a proof which does not use any complexity-theoretic assumptions.
3. Which finite-domain CSPs are in HerFO?
4. Is every CSP in HerFO also of the form $\text{Her}(\phi)$ for some negative connected sentence ϕ ?
5. Is it true that the tractability problem for HerFO is undecidable even for first-order sentences with quantifier prefix $\exists\exists\forall$ (assuming $\text{P} \neq \text{NP}$)? This is the only quantifier-prefix for which the decidability of the tractability problem for HerFO remains open (see Corollary 24 and the first item of Theorem 20).

References

- 1 Max Bannach, Florian Chudigiewitsch, and Till Tantau. On the descriptive complexity of vertex deletion problems. In Rastislav Kráľovic and Antonín Kucera, editors, *49th International Symposium on Mathematical Foundations of Computer Science, MFCS 2024, August 26-30, 2024, Bratislava, Slovakia*, volume 306 of *LIPIcs*, pages 17:1–17:14, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.MFCS.2024.17.
- 2 Manuel Bodirsky. *Complexity of Infinite-Domain Constraint Satisfaction*. Lecture Notes in Logic (52). Cambridge University Press, Cambridge, United Kingdom; New York, NY, 2021. doi:10.1017/9781107337534.
- 3 Manuel Bodirsky and Santiago Guzmán-Pro. Hereditary First-Order Logic: the tractable quantifier prefix classes. Preprint arXiv:2411.10860, 2024. doi:10.48550/arXiv.2411.10860.
- 4 Manuel Bodirsky and Jan Kára. A fast algorithm and datalog inexpressibility for temporal reasoning. *ACM Trans. Comput. Log.*, 11(3):15:1–15:21, 2010. doi:10.1145/1740582.1740583.
- 5 Manuel Bodirsky, Simon Knäuer, and Sebastian Rudolph. Datalog-expressibility for monadic and guarded second-order logic. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 120:1–120:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. Preprint available at <https://arxiv.org/abs/2010.05677>. doi:10.4230/LIPIcs.ICALP.2021.120.
- 6 J. Adrian Bondy and Uppaluri S. R. Murty. *Graph Theory*. Graduate Texts in Mathematics. Springer, Berlin, 2008. doi:10.1007/978-1-84628-970-5.
- 7 Egon Börger, Erich Grädel, and Yuri Gurevich. *The Classical Decision Problem*. Perspectives in Mathematical Logic. Springer, 1997.
- 8 Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite model theory*. Perspectives in Mathematical Logic. Springer, Berlin, Heidelberg, New York, 1995. Second edition.
- 9 Thomas Eiter, Georg Gottlob, and Thomas Schwentick. The model checking problem for prefix classes of second-order logic: A survey. In Andreas Blass, Nachum Dershowitz, and Wolfgang Reisig, editors, *Fields of Logic and Computation, Essays Dedicated to Yuri Gurevich on the Occasion of His 70th Birthday*, volume 6300 of *Lecture Notes in Computer Science*, pages 227–250, Berlin, Heidelberg, 2010. Springer. doi:10.1007/978-3-642-15025-8_13.
- 10 Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998. doi:10.1137/S0097539794266766.
- 11 Tomás Feder and Moshe Y. Vardi. Homomorphism closed vs. existential positive. In *18th IEEE Symposium on Logic in Computer Science (LICS 2003), 22-25 June 2003, Ottawa, Canada, Proceedings*, pages 311–320. IEEE Computer Society, 2003. doi:10.1109/LICS.2003.1210071.
- 12 Fedor V. Fomin, Petr A. Golovach, and Dimitrios M. Thilikos. On the parameterized complexity of graph modification to first-order logic properties. *Theory Comput. Syst.*, 64(2):251–271, 2020. doi:10.1007/s00224-019-09938-8.

- 13 Toshihiro Fujito. A unified approximation algorithm for node-deletion problems. *Discret. Appl. Math.*, 86(2-3):213–231, 1998. doi:10.1016/S0166-218X(98)00035-3.
- 14 Georg Gottlob, Phokion G. Kolaitis, and Thomas Schwentick. Existential second-order logic over graphs: Charting the tractability frontier. *J. ACM*, 51(2):312–362, 2004. doi:10.1145/972639.972646.
- 15 Wilfrid Hodges. *Model theory*, volume 42 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, Cambridge, 1993.
- 16 Wilfrid Hodges. *A shorter model theory*. Cambridge University Press, Cambridge, 1997.
- 17 Mukkai S. Krishnamoorthy and Narsingh Deo. Node-deletion np-complete problems. *SIAM J. Comput.*, 8(4):619–625, 1979. doi:10.1137/0208049.
- 18 John M. Lewis and Mihalis Yannakakis. The node-deletion problem for hereditary properties is np-complete. *J. Comput. Syst. Sci.*, 20(2):219–230, 1980. doi:10.1016/0022-0000(80)90060-4.
- 19 Christos H. Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. Syst. Sci.*, 43(3):425–440, 1991. doi:10.1016/0022-0000(91)90023-X.
- 20 Venkatesh Raman, Saket Saurabh, and C. R. Subramanian. Faster fixed parameter tractable algorithms for finding feedback vertex sets. *ACM Trans. Algorithms*, 2(3):403–415, July 2006. doi:10.1145/1159892.1159898.
- 21 Bruce A. Reed, Kaleigh Smith, and Adrian Vetta. Finding odd cycle transversals. *Oper. Res. Lett.*, 32(4):299–301, 2004. doi:10.1016/j.orl.2003.10.009.
- 22 Donald J. Rose. A note on consistent ordering and zero circulation. *J. ACM*, 18(4):573–575, 1971. doi:10.1145/321662.321671.
- 23 Donald J. Rose, Robert Endre Tarjan, and George S. Lueker. Algorithmic aspects of vertex elimination on graphs. *SIAM J. Comput.*, 5(2):266–283, 1976. doi:10.1137/0205021.
- 24 Boris A. Trakhtenbrot. Understanding basic automata theory in the continuous time setting. *Fundam. Informaticae*, 62(1):69–121, 2004. (in Russian). URL: <http://content.iospress.com/articles/fundamenta-informaticae/fi62-1-04>.

A Examples

Polynomial-time solvable examples

Example 3 shows that $\text{CSP}(\mathbb{Q}, <)$ is in HerFO but not in FO. There are also finite-domain CSPs that are in HerFO, but not in FO, as the following examples show. First, recall that the *algebraic length* of an oriented walk P is the absolute value of the difference between the number of forward edges and the number of backward edges in P . It is well-known (and straightforward to observe) that a digraph D homomorphically maps to the directed path of length 2 if and only if every oriented walk in D has algebraic length at most 2.

► **Example 25.** The problem $\text{CSP}(\vec{P}_3)$ is in HerFO. Consider a first-order formula ϕ saying that there is a loop, or there are (directed) edges (x, y) and (a, b) such that the following hold:

- y has exactly one out-neighbour different from b , and x is the unique in-neighbour of y ;
- a has exactly one in-neighbour different from x , and b is the unique out-neighbour of a ;
- every $z \notin \{x, y, a, b\}$ either
 - has no in-neighbours and exactly two out-neighbours, and they are different from b , or
 - has no out-neighbours and exactly two in-neighbours, and they are different from x .

Suppose that a loopless digraph $\mathbb{D}$ satisfies ϕ , and let $a, b, x, y \in D$ be witnesses that show that ϕ holds in $\mathbb{D}$. By using the first bullet, we find $z_1 \in D \setminus \{b\}$ such that $(y, z_1) \in E(\mathbb{D})$, and by using the third point iteratively, we find an alternating oriented path $z_1, z_2, \dots, z_k$

for some $k \geq 3$, i.e., $(z_2, z_1), (z_2, z_3), (z_4, z_3), \dots \in E(\mathbb{D})$, and $z_k \in \{a, b, x, y\}$. We claim that we can find an oriented walk of the form

$$x \rightarrow y \rightarrow z_1 \leftarrow z_2 \rightarrow \dots \leftarrow z_k, \text{ where } z_k = b, \text{ or of the form}$$

$$x \rightarrow y \rightarrow z_1 \leftarrow z_2 \rightarrow \dots \rightarrow z_k, \text{ where } z_k \in \{x, a\}.$$

We may choose z_k so that $z_k \in \{a, b, x, y\}$ and $z_i \notin \{a, b, x, y\}$ for $i \in [k-1]$. We first consider the case where z_k is an in-neighbour of z_{k-1} . In this case, we know by the first item that $z_k \neq y$, because $k \geq 3$ and y has exactly one out-neighbour different from b , namely $z_1 \neq z_{k-1}$. We also know that $z_k \neq a$ because the unique out-neighbour of a is $b \neq z_{k-1}$ (second bullet). Finally, $z_k \neq x$, because z_{k-1} has exactly two in-neighbours, and they are different from x , and so, $z_k = b$. In the case where z_k is an out-neighbour of z_{k-1} , then $z_k \neq y$ because the unique in-neighbour of y is $x \neq z_{k-1}$. Similarly as before, we see that $z_k \neq b$ by applying the third bullet to z_{k-1} .

Using the existence of such paths we can easily find an oriented path of algebraic length three in $\mathbb{D}$. Indeed, in the case where the path ends with an edge $z_{k-1} \leftarrow z_k$, and $z_k = b$, we use the unique in-neighbour a' of a to construct a path of length three a', a, b, z_{k-1} . Now, consider the case when the path ends with an edge $z_{k-1} \rightarrow z_k$. If $z_k = x$, then z_{k-1}, x, y, z_1 is a directed walk of length three, and if $z_k = a$, then $x, y, z_1, \dots, z_{k-1}, a, b$ is an oriented walk of algebraic length three. Hence, if $\mathbb{D} \notin \text{Her}(\neg\phi)$, then either $\mathbb{D}$ contains a loop or a walk of algebraic length 3, and so $\mathbb{D} \not\models \vec{P}_3$.

Conversely, if $\mathbb{D} \not\models \vec{P}_3$, then either $\mathbb{D}$ contains a loop or it has an oriented path of algebraic length 3. In the former case, $\mathbb{D}$ clearly satisfies ϕ , and in the latter, by choosing the shortest such path $v_1, \dots, v_n$ we find a substructure of $\mathbb{D}$ that models ϕ ; namely, the substructure with vertex set $\{v_1, \dots, v_n\}$. Therefore, if $\mathbb{D} \not\models \vec{P}_3$, then $\mathbb{D} \notin \text{Her}(\neg\phi)$.

Inexpressible examples and ω -categoricity

A digraph $\mathbb{D}$ is *balanced* if every oriented cycle in $\mathbb{D}$ has the same number of forward and of backward edges. Clearly, the class of balanced digraphs is hereditary, and so, it could a priori be expressible in HerFO. Here, we use the notion of ω -categoricity (Remark 6) to prove that the class of balanced digraphs is not expressible in HerFO.

► **Example 26.** It is folklore that a digraph $\mathbb{D}$ maps homomorphically to some directed path if and only if $\mathbb{D}$ is balanced. Equivalently, if succ is the successor relation in $\mathbb{Z}$, then $\mathbb{D} \rightarrow (\mathbb{Z}, \text{succ})$ if and only if $\mathbb{D}$ is a balanced digraph. Hence, the class of balanced digraphs is in HerFO iff $\text{CSP}(\mathbb{Z}, \text{succ})$ is in HerFO. It is known that there is no ω -categorical structure $\mathbb{A}$ whose CSP equals the CSP of $(\mathbb{Z}, \text{succ})$ [2, Proposition 5.8.2]. Therefore, $\text{CSP}(\mathbb{Q}, \text{succ})$ (equivalently, the class of balanced digraphs) is not expressible in HerFO.

HerFO and Datalog

Datalog can be seen as the subclass of SNP where we require that the first-order part ψ of the SNP sentence

- is *Horn*, i.e., written in conjunctive normal form such that each clause contains at most one positive literal, and
- is such that every positive literal in ψ is existentially quantified.

A class $\mathcal{C}$ of finite models is *in Datalog* if there exists an SNP sentence Φ of the form described above such that a finite structure is in $\mathcal{C}$ if and only if it does *not* satisfy Φ . Note that if a class is in Datalog, then it is closed under homomorphisms⁴ and in P.

A simple example of a CSP which is solved by a Datalog program is $\text{CSP}(K_2)$. In Example 10 we showed that this class is not in HerFO. There are also CSPs that are in HerFO and in P, but not in Datalog.

► **Example 27.** For any positive integer k consider a $(k+1)$ -ary relation symbol R . The problem $\text{CSP}(\mathbb{Q}, \{(x, y_1, \dots, y_k) : x < \max\{y_1, \dots, y_k\}\})$ is in P, but not in Datalog [4]. Notice that the problem is in HerFO: consider the formula

$$\exists x \forall y_1, \dots, y_k. \neg R(x, y_1, \dots, y_k).$$

► **Corollary 28.** *The classes of CSPs in Datalog and in $\text{P} \cap \text{HerFO}$ are incomparable.*

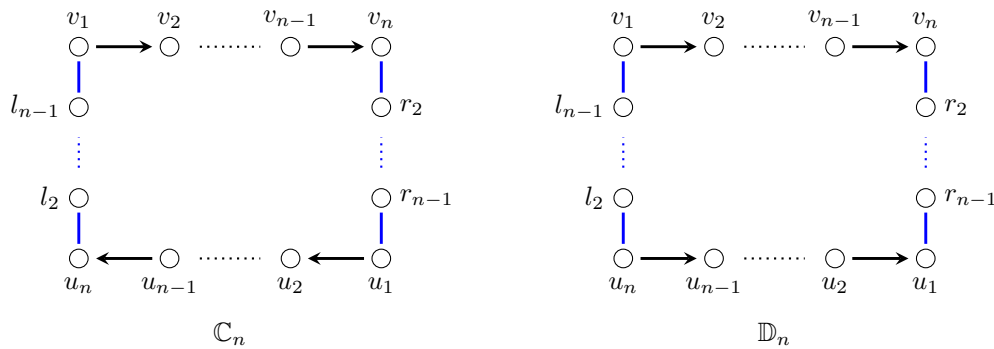
HerFO and pp definitions

A *primitive positive formula* is an existential positive formula whose quantifier-free part only uses variables and symbols from $\{\wedge, =\} \cup \tau$ (i.e., disjunction is forbidden). Given a structure $\mathbb{A}$ we say that a relation $R \subseteq A^r$ is *primitively positively definable* (pp-definable) in $\mathbb{A}$ if there is a primitive positive formula $\phi(x_1, \dots, x_r)$ such that $\bar{a} \in R$ if and only if $\mathbb{A} \models \phi(\bar{a})$. A class of structures $\mathcal{C}$ is *preserved* by primitive positive definitions if for every $\mathbb{A} \in \mathcal{C}$, and every pp-definable relation R in $\mathbb{A}$, the structure $(\mathbb{A}, R)$ belongs to $\mathcal{C}$. In Example 3 we showed that $\text{CSP}(\mathbb{Q}, <)$ is expressible in HerFO. The following example shows that $\text{CSP}(\mathbb{Q}, <, =)$ is not expressible in HerFO, and so, the class of structures whose CSP in HerFO is not preserved by primitive positive definitions.

► **Example 29.** The problem $\text{CSP}(\mathbb{Q}, <, =)$ is not hereditarily definable. Consider the family of structures $\mathbb{C}_n$ and $\mathbb{D}_n$ illustrated below (where undirected blue edges represent the (symmetric) binary relation $=$, and directed black edges represent the binary relation $<$). It is not hard to see that for every positive integer n ,

- $\mathbb{C}_n$ is a minimal obstruction of $\text{CSP}(\mathbb{Q}, <, =)$, and
- $\mathbb{D}_n \in \text{CSP}(\mathbb{Q}, <, =)$.

Now, for every $k \in \mathbb{N}$ one can choose n, m appropriately so that $\mathbb{C}_n \equiv_k \mathbb{D}_m$, and it thus follows by Lemma 9 that this CSP is not expressible in HerFO.



⁴ Our definition is standard, but different from the terminology of Feder and Vardi [10, 11], who defined that $\text{CSP}(\mathbb{B})$ is in Datalog if its complement is in Datalog in the standard sense.

B Comparison of complexity classifications

As we discussed in the introduction, for every first-order sentence ϕ , there is a UMSO sentence Φ such that $\text{Her}(\phi)$ is described by Ψ , and the first-order part of Φ has the same quantifier prefix as ϕ . Hence, if every problem expressible by a UMSO sentence with quantifier prefix Q can be solved in polynomial time, then $\text{Her}(\phi)$ can be solved in polynomial time for ϕ with quantifier prefix Q .

We also discussed the *vertex deletion problem* (parametrized version of HerFO): for a fixed first-order sentence ϕ , decide whether an input structure $\mathbb{A}$ satisfies ϕ after deleting at most k vertices. In Table 1, we present a comparison between our complexity classification based on the quantifier prefix and similar known classifications for UMSO and for the parametrized version. We restrict these classification to digraphs, because, as far as we are aware, complexity classifications for the vertex deletion problem have only been considered for digraphs.

■ **Table 1** Complexity landscape for the model checking problems for UMSO, for HerFO, and for the vertex deletion problem (all restricted to digraphs). In the HerFO column, “Always in P” means that $\text{Her}(\phi)$ is polynomial-time solvable whenever ϕ has quantifier prefix Q ; in the UMSO column, it means that the model checking problem can be solved in polynomial time for every UMSO sentence Φ whose first-order part has quantifier prefix Q . “Hard” in the HerFO and UMSO columns means that the corresponding fragments can express coNP-complete problems. On the last column, FPT means that for every first-order sentence ϕ with quantifier prefix Q , the vertex deletion problem to ϕ is fixed-parameter tractable; and “hard” means that there are W[2]-hard problems in that fragment.

Quantifier prefix Q	HerFO (Theorem 20)	UMSO [14, Figure 1(a)]	Vertex deletion [12, Theorem 1]
$\forall^* \exists \exists$	Always in P	Always in P	Always FPT
$\forall^* \exists^*$	Always in P	Hard	Always FPT
$\forall^* \exists \forall^*$	Always in P	Hard	Hard
$\exists \exists \forall$	Hard	Hard	Always FPT
$\exists \forall \exists$	Hard	Hard	Always FPT
$\forall^* \exists^* \forall^*$	Hard	Hard	Hard

C Two missing proofs

Proof of Lemma 9. Suppose that $\mathcal{C} = \text{Her}(\phi)$ for some first-order formula ϕ . We claim that $\psi := \neg\phi$ satisfies both itemized statements. Firstly, note that if $\mathbb{F} \in \mathcal{F}$, then $\mathbb{F} \notin \text{Her}(\phi)$, so some substructure of $\mathbb{F}$ does not satisfy ϕ . Moreover, all substructures of $\mathbb{F}$ belong to $\text{Her}(\phi)$, so we have that $\mathbb{F}$ itself does not satisfy ϕ , and hence satisfies ψ . If $\mathbb{A}$ is a finite τ -structure such that $\mathbb{A} \models \psi$, then $\mathbb{A} \notin \mathcal{C}$, and hence there exists $\mathbb{F} \in \mathcal{F}$ which embeds into $\mathbb{A}$. This shows the forward implication of the statement. Conversely, if there exists a formula ψ that satisfies both items of the statement, then it is similarly straightforward to show that $\mathcal{C} = \text{Her}(\neg\psi)$. ◀

Proof of Proposition 11. We claim that $\mathbb{A}$ hereditarily models ϕ if and only if every substructure $\mathbb{B}$ of $\mathbb{A}$ with at most $2^{|\tau|}$ many elements satisfies ϕ . One direction follows from definition of hereditary satisfiability. For the converse implication, let $\mathbb{B}$ be a substructure of $\mathbb{A}$ and consider a minimal subset $C \subseteq B$ such that for every $b \in B$ there is a $c \in C$ such that $U(b) \Leftrightarrow U(c)$ for all $U \in \tau$. Clearly, $|C| \leq 2^{|\tau|}$ and it is straightforward to observe that $\mathbb{B} \models \phi$ if and only if $\mathbb{C} \models \phi$, and the claim follows. ◀