S.T. Yau High School Science Award

Research Report

The Team

Name of team member: Derek Aoki
School: Proof School
City, Country: San Francisco, California, United States

Name of supervising teacher: Noah Schweber
Job Title: Instructor
School/Institution: Proof School
City, Country: San Francisco, California, United States

Title of Research Report

Order-Invariant CFO Contains Local-Order-Invariant FO

Date

September 13, 2026

ORDER-INVARIANT CFO CONTAINS LOCAL-ORDER-INVARIANT FO

DEREK AOKI

Abstract. Order-invariant cluster first-order logic (CFO) and local-order-invariant first-order logic are two fragments of order-invariant first-order logic, both motivated by the study of invariant logics on classes of bounded degree. Our main result is that order-invariant CFO contains local-order-invariant FO on arbitrary classes of finite structures, settling the conjecture of Ghasemi and Grange that order-invariant CFO contains plain FO. By combining this with a collapse result of Ghasemi and Grange, we obtain the equivalence of FO, local-order-invariant FO, and order-invariant CFO on classes of finite colored graphs of bounded degree. The proof makes use of Gaifman normal form, which allows our translation from local-order-invariant FO to order-invariant CFO to be effective. Combined with the fixed-parameter tractability of model-checking for order-invariant CFO on bounded-degree colored graph classes, this shows that model-checking for local-order-invariant FO on bounded-degree colored graph classes is fixed-parameter tractable. To generalize our collapse and model-checking results beyond colored graphs, we use a method based on first-order interpretations which also yields a collapse of order-invariant FO to FO when restricted to graph classes of bounded shrub-depth. Finally, we give an application to the database-motivated frozen-witness logic of Abiteboul, Simon, and Vianu by showing that local-order-invariant FO contains deterministic frozen-witness FO, and thus that deterministic frozen-witness FO collapses to FO on classes of finite structures of bounded degree.

2012 ACM Subject Classification: Theory of computation → Finite Model Theory
Key Words and Phrases: Local order invariance, order invariance, cluster first-order logic

Acknowledgements

The author would like to thank Noah Schweber for helpful conversations during the writing of this paper.

Commitments on Academic Honesty and Integrity

We hereby declare that we

1. are fully committed to the principle of honesty, integrity and fair play throughout the competition.
2. actually perform the research work ourselves and thus truly understand the content of the work.
3. observe the common standard of academic integrity adopted by most journals and degree theses.
4. have declared all the assistance and contribution we have received from any personnel, agency, institution, etc. for the research work.
5. undertake to avoid getting in touch with assessment panel members in a way that may lead to direct or indirect conflict of interest.
6. undertake to avoid any interaction with assessment panel members that would undermine the neutrality of the panel member and fairness of the assessment process.
7. observe the safety regulations of the laboratory(ies) where we conduct the experiment(s), if applicable.
8. observe all rules and regulations of the competition.
9. agree that the decision of YHSA is final in all matters related to the competition.

We understand and agree that failure to honour the above commitments may lead to disqualification from the competition and/or removal of reward, if applicable; that any unethical deeds, if found, will be disclosed to the school principal of team member(s) and relevant parties if deemed necessary; and that the decision of YHSA is final and no appeal will be accepted.

Name of team member: Derek Aoki

Signature of team member:



Name of supervising teacher: Noah Schweber

Signature of supervising teacher:



Declaration of Academic Integrity

The participating team declares that the paper submitted is comprised of original research and results obtained under the guidance of the instructor. To the team's best knowledge, the paper does not contain research results, published or not, from a person who is not a team member, except for the content listed in the references and the acknowledgment. If there is any misinformation, we are willing to take all the related responsibilities.

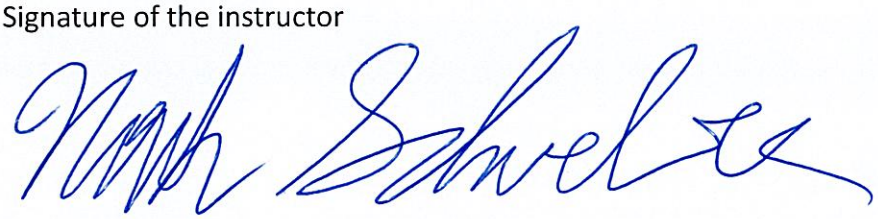
Name of team member: Derek Aoki

Signature of team member



Name of the instructor: Noah Schweber

Signature of the instructor



Date: September 13, 2026

Contents

Acknowledgements 2
1. Introduction 6
2. Preliminaries 8
2.1. Locality 8
2.2. Invariant Logics 8
2.3. Cluster First-Order Logic 9
3. Main Containment Theorem 10
4. Generalizations via Interpretations 14
5. Deterministic Frozen-Witness Logic 18
6. Conclusion 22
Appendix A. Proofs for Section 4 23
References 26

1. Introduction

A foundational idea in the theory of relational databases is that of representation-independence, introduced by Chandra and Harel [6]: a query should depend only on the isomorphism type of the input database, and not the particular representation. In particular, the order in which the elements of the databases are stored should not affect the output of the query. However, it is natural to want to make use of such information in the representation, so long as the query does not depend on said order. We call such a query, that is, one that uses an order but does not depend on the particular choice of order, order-invariant.

Finite model theory is the study of the queries on databases that can be expressed in formal languages. The seminal result was that of Fagin [11], showing that the queries in NP are exactly the queries which can be expressed in existential second-order logic. Since then, many more bridges of this form have been found. In particular, Immerman and Vardi [18, 27] showed that, in the presence of a linear order, the queries in P are exactly the ones expressible in $\mathrm{FO} + \mathrm{LFP}$, first-order logic with the least fixed point operator. In general, capturing classes below NP with logics makes use of order-invariance or similar invariant auxiliary structure (see the capturings of AC^0 [20], L [19], and others [21]). Thus, understanding exactly how it affects the expressive power of a logic is a central question of finite model theory, and closely related to questions in computational complexity.

It has been proved that there exist properties of finite structures (relational databases) definable in order-invariant FO which are not definable in FO, the first of which was discovered by Gurevich and published in [9]. In other words, requiring the output to not depend on the order of the input does not make the order useless. In fact, it has been shown that the weaker fragment successor-invariant FO, in which formulae can only access the successor relation induced by the order in which the data is stored, can express properties of finite structures that cannot be expressed in FO [26]. Further non-collapse results include those for epsilon-invariant FO and deterministic frozen-witness FO [25], as well as for local-order-invariant FO [4].

Based on this, it is natural to ask whether invariant logics collapse on finite structures or on particular classes of finite structures. For instance, it has been proved that successor-invariant FO and FO express the same queries on classes of bounded degree (roughly, classes of graphs in which each vertex has a bounded number of neighbors) [16], and that order-invariant FO collapses on trees, strings, and hollow trees [5, 17]. The question of whether full order-invariant FO collapses on classes of bounded degree is significant and open. However, a substantial fragment of order-invariant FO, known as order-invariant cluster first-order logic (CFO), has recently been introduced and shown to be contained in FO on classes of colored graphs of bounded degree [15]. A second question concerns the complexity of model-checking for these invariant logics, and it has been shown for instance that model-checking for order-invariant CFO is FPT [15].

An invariant logic of particular interest to us is local-order-invariant logic. Instead of allowing a linear order on an entire graph, we can allow linear orders for the set of neighbors of each vertex. More precisely, a local order on a graph is a ternary relation $\prec (x, y, z)$ such that when x is fixed, $\prec (x, y, z)$ linearly orders the neighbors of x. On structures over arbitrary vocabularies,

a local order is a local order on the Gaifman graph of the structure. Such local orders have been considered in the context of finite model theory prior to being considered in the context of invariant logics [5, 8]. Local-order-invariant FO was first considered in [23], and a number of questions related to this work were asked by Weinstein in [28]. Generalizations of these questions were answered in [4].

Our Contributions

(1) Our main result is that every property expressible in local-order-invariant FO is expressible in order-invariant CFO. Our translation uses Gaifman normal form, as introduced in [12], which makes it effective. Thus, our translation works on arbitrary classes of finite structures. This affirms in particular the conjecture of Ghasemi and Grange in [15] that order-invariant CFO contains FO. An important corollary is that FO and local-order-invariant FO have the same expressiveness on classes of bounded degree. In [4], this result is proved in particular for the full class of structures over a vocabulary σ of degree at most d, and our result generalizes this to arbitrary classes of bounded degree. Another corollary is that, by the effectiveness of the translation to order-invariant CFO, the model-checking problem for local-order-invariant FO is FPT. All of these results are proved first for colored graphs.

(2) To generalize our main result, its corollaries, and the collapse result of [15] from classes of colored graphs of bounded degree to classes of bounded degree over arbitrary vocabularies, we use a certain class of first-order interpretations. We believe that this method of investigating invariant logics via interpretations has been under-explored, as it allows us, among several other things, to straightforwardly combine several results from [7, 10, 24] to obtain the collapse of order-invariant FO to FO when restricted to graph classes of bounded shrub-depth.

(3) It is proved in [4] that epsilon-invariant FO is contained in local-order-invariant FO on classes of bounded degree, but the proof is flawed in that it accounts for neither the extensionality nor the nesting of epsilon terms. However, the basic idea of the proof is usable. Deterministic frozen-witness FO, introduced and studied in [1, 2, 3], is a weaker logic than epsilon-invariant FO. It is of interest in the study of non-determinism in databases. Abiteboul, Simon, and Vianu originally conjectured in [2] that it should collapse to FO, and this conjecture was refuted in [25]. We use Gaifman normal form again to show the containment of deterministic frozen-witness FO in local-order-invariant FO on finite structures, the collapse of deterministic frozen-witness FO on classes of finite structures of bounded degree, and the fixed-parameter tractability of the model-checking problem for deterministic frozen-witness FO on classes of finite structures of bounded degree. This recovers the collapse conjectured by Abiteboul, Simon, and Vianu on classes of bounded degree.

In Section 2, we give the necessary background. In Section 3, we prove our titular containment theorem and its corollaries in the setting of colored graphs. In Section 4, we extend the main containment and related results to arbitrary vocabularies, and prove the collapse for

bounded shrub-depth. In Section 5, we prove the deterministic frozen-witness containment and its corollaries.

2. Preliminaries

From here, σ will be a finite relational vocabulary. We denote by $\mathfrak{F}_\sigma$ the class of finite structures over σ, and $\mathfrak{K}_\sigma$ the class of arbitrary structures over σ.

2.1. **Locality.** For any $X \subseteq A$, we write $A \restriction X$ for the restriction of A to X; i.e., the substructure of A with domain X such that for any tuple $\vec{a}$ of elements of X and any relation $R \in \sigma$ of arity $|\vec{a}|$, we have $R(\vec{a})$ in $A \restriction X$ if and only if $R(\vec{a})$.

Definition 2.1. Let A be in $\mathfrak{K}_\sigma$. We define the Gaifman graph of A, denoted by $\mathcal{G}(A)$, such that the vertices of $\mathcal{G}(A)$ are the elements of A, and such that there are edges between two vertices $a, a' \in A$ if and only if there exist a tuple $\vec{a}$ of elements of A and a relation $R \in \sigma$ of arity $|\vec{a}|$ such that $a, a' \in \vec{a}$ and $R(\vec{a})$ in A. We now recall several topological graph notions as they relate to Gaifman graphs.

(1) For $a, a' \in A$, we define the distance $\delta^A(a, a')$ as the least number of edges in a path in $\mathcal{G}(A)$ from a to a', if such a path exists, and ∞ otherwise.
(2) For $a \in A$, we denote by $B^A_r(a)$ the ball of radius r around a, $\{a' \in A : \delta^A(a, a') \leq r\}$. We denote by $S^A_r(a)$ the sphere of radius r around a, $B^A_r(a) \setminus B^A_{r-1}(a)$.
(3) For $a \in A$, we denote by $N^A_r(a)$ the structure over $\sigma \cup \{c\}$ which is identical to the structure $A \restriction B^A_r(a)$ except that it has the additional constant symbol c which is equal to a. We call this the neighborhood of radius r around a. We say that the r-neighborhood type of $a \in A$, denoted $\mathrm{tp}^A_r(a)$, is the class $\{B \in \mathfrak{K}_\sigma : B \cong N^A_r(a)\}$.

We can generalize the above definitions of balls, spheres, and neighborhoods to arbitrary tuples $\vec{a}$ of elements of A by taking $B^A_r(\vec{a}) = \bigcup_{a \in \vec{a}} B^A_r(a)$.

Definition 2.2. We define $\mathfrak{F}^d_\sigma$ as the structures of degree at most d over σ, i.e.

$$\mathfrak{F}^d_\sigma = \{A \in \mathfrak{F}_\sigma : |S^A_1(a)| \leq d \text{ for every } a \in \mathcal{G}(A)\}.$$

Similarly, we define $\mathfrak{K}^{<\omega}_\sigma$ as the locally finite structures over σ, i.e.

$$\mathfrak{F}^{<\omega}_\sigma = \{A \in \mathfrak{F}_\sigma : |S^A_1(a)| < \omega \text{ for every } a \in \mathcal{G}(A)\}.$$

2.2. **Invariant Logics.**

Definition 2.3. Let φ be a sentence in the vocabulary $\sigma \cup \{<\}$. We say that φ is order-invariant on $\mathfrak{K} \subseteq \mathfrak{K}_\sigma$ if, for any $A \in \mathfrak{K}$ and any linear orders $<_1, <_2$ on A, we have $(A, <_1) \models \varphi$ if and only if $(A, <_2) \models \varphi$. If φ is order-invariant, we say that $A \models \varphi$ if $(A, <) \models \varphi$ for any linear order $<$.

Definition 2.4. We say that a ternary relation $\prec (x, y, z)$ on a graph A is a local order if, when x is fixed, $\prec (x, y, z)$ is a linear order on $S^A_1(x)$. This definition generalizes to general structures by instead considering the Gaifman graph $\mathcal{G}(A)$. We define local-order-invariance of a sentence φ in $\sigma \cup \{\prec\}$ in the natural way.

The following is the main result of [4]. A generalization is claimed for arbitrary classes of bounded degree and used to prove the claimed collapse of epsilon-invariant FO to FO on classes of bounded degree, but this generalization does not follow from the proof techniques of the paper because the result of [23] used does not apply in general (in particular, it has the additional hypothesis that the query must be invariant on an ambient substructure-closed class).

Theorem 2.5 (Aoki)**.** *Let $d \in \omega$, and let φ be a local-order-invariant sentence over $\mathfrak{F}_\sigma^d$. Then there is a first-order sentence ψ such that for any $A \in \mathfrak{F}_\sigma^d$, $A \models \psi$ if and only if $A \models \varphi$.*

2.3. **Cluster First-Order Logic.** Before describing cluster first-order logic on ordered graphs rigorously, we give a brief informal description. Note that, for the following discussion and definition, the graphs in question will be simple (i.e., undirected and without self-edges). First, we fix an infinite alphabet Σ. The variables of CFO are indexed as x_w^i, with $i \in \omega$ and $w \in \Sigma^*$, and these variables are grouped into clusters based on whether they share the index w. Introducing a variable x_w^j with $j > 0$ via a quantifier is guarded by adjacency with the previous elements of the cluster, in the sense that there must exist some $i < j$ such that $E(x_w^i, x_w^j)$. Also, the relations E, $=$, and $<$ can only be used to compare variables of the same cluster, with the exception that $=$ and $<$ may be used to compare x_w^i with $x_{w\alpha}^0$ for any i. Now for the rigorous definition.

Definition 2.6. Fix an infinite alphabet Σ, and let $S \subseteq \Sigma^* \times \omega$ be a finite set of indices. We say that S is valid if it is downward closed, in the sense that for any $w \in \Sigma^*$, $\alpha \in \Sigma$, and $i \in \omega$,

(1) for every $j < i$, $(w, i) \in S$ implies $(w, j) \in S$, and
(2) $(w\alpha, 0) \in S$ implies $(w, 0) \in S$.

If S is a valid set of indices, we write X_S for the set $\{x_w^i : (w, i) \in S\}$. Letting σ be a vocabulary $\{E, P_1, \ldots, P_n\}$, where E is binary and each P_i is unary, we define $\mathrm{CFO}_S[k]$ by induction on k.

(1) For every $(w, i) \in S$, $j < i$, $l \leq n$, $\alpha \in \Sigma$ such that $(w\alpha, 0) \in S$, and $\sim$ in $\{=, <, >\}$, then $\mathrm{CFO}_S[0]$ contains the atomic formulae $P_l(x_w^i)$, $E(x_w^i, x_w^j)$, $x_w^i \sim x_w^j$, and $x_w^i \sim x_{w\alpha}^0$. Also, $\mathrm{CFO}_S[0]$ is closed under Boolean combinations.
(2) $\mathrm{CFO}_S[k+1]$ contains the formulae
   (a) $\exists x_\varepsilon^0 \varphi$, where $\varphi \in \mathrm{CFO}_{\{(\varepsilon,0)\}}$ (assuming that $S = \emptyset$);
   (b) $\exists x_{w\alpha}^0 \varphi$, where $(w, 0) \in S$, $\alpha \in \Sigma$, $(w\alpha, 0) \notin S$, and $\varphi \in \mathrm{CFO}_{S \cup \{(w\alpha,0)\}}[k]$;
   (c) $\exists x_w^i E(x_w^i, x_w^j) \wedge \varphi$, where $(w, j) \in S$, $i = \min\{l : (w, l) \notin S\}$, and $\varphi \in \mathrm{CFO}_{S \cup \{(w,i)\}}[k]$.

   Also, $\mathrm{CFO}_S[k+1]$ is closed under Boolean combinations.

We write $\mathrm{CFO}_S = \bigcup_{k \in \omega} \mathrm{CFO}_S[k]$, and we write CFO for $\mathrm{CFO}_\emptyset$. Because CFO_S is a fragment of first-order logic, its semantics are very natural. That is, for $\varphi \in \mathrm{CFO}_S$, $A \in \mathfrak{F}_\sigma$, a linear order $<$ on A, and a valuation $\nu : X_S \to A$, we have $(A, <, \nu) \models \varphi$ (or $(A, <) \models \varphi$ if $S = \emptyset$) if this holds when φ is treated as a first-order formula with free variables in X_S. Order-invariance of a CFO sentence is defined in the obvious way.

Though CFO was defined originally in [15] only for (ordered) colored graphs, there is a natural extension to (ordered) structures over any vocabulary. To do this, we first note the following straightforward lemma.

Lemma 2.7. *For every vocabulary σ, there exists a formula $E_G(x, y)$ such that for any a, a', we have $E_G(a, a')$ in A if and only if a and a' are adjacent in $\mathcal{G}(A)$.*

Definition 2.8. Define CFO over an arbitrary vocabulary σ identically to CFO over colored graphs, except

(1) $\mathrm{CFO}_S[0]$ contains $R(\vec{x})$ where the elements of $\vec{x}$ are all of the form x^i_w for some w with $(w, i) \in S$, and
(2) we guard cluster quantification by $\exists x^i_w E_G(x^i_w, x^j_w) \wedge \varphi$, where E_G is defined as above.

Theorem 2.9 (Ghasemi–Grange)**.** *Let σ be a colored graph vocabulary as above, and let $\mathfrak{K}$ be a class of finite colored graphs over σ of degree at most d. Let φ be an order-invariant* CFO *sentence on $\mathfrak{K}$. Then φ is equivalent over $\mathfrak{K}$ to some first-order sentence ψ in the vocabulary σ. In other words, order-invariant* CFO *is contained in* FO *on colored graph classes of bounded degree.*

The following was explicitly conjectured in [15].

Conjecture 2.10 (Ghasemi–Grange)**.** *Let $\mathfrak{K}$ be a class of finite colored graphs over σ, and let φ be a first-order sentence. Then φ is equivalent on $\mathfrak{K}$ to some order-invariant* CFO *sentence ψ. In other words,* FO *is contained in order-invariant* CFO *on arbitrary classes of finite colored graphs.*

3. Main Containment Theorem

In this section, we will only work with colored simple graphs over a vocabulary σ.

We will now introduce Gaifman normal form and several relevant definitions. While it is true that in this subsection we are only working with colored simple graphs, the following applies to an arbitrary relational vocabulary.

For $r \in \omega$ and a first-order formula $\varphi(\vec{x})$, we write $\varphi^{(r)}$ for the formula which is identical to φ except the quantifiers are bounded to $B^A_r(\vec{x})$. We call $\varphi^{(r)}$ an r-local formula. As a consequence of Lemma 2.7, there exist formulae which express $\delta(x, y) \leq r$ and $\delta(x, y) > r$ for every r. So we can instead phrase $\varphi^{(r)}$ as the formula that results when $\exists y \psi$ is replaced by $\exists y(\bigvee_i \delta(x_i, y) \leq r \wedge \psi)$, and similarly for universal quantifiers. We define basic local sentences as sentences of the form

$$\exists x_1, \ldots, x_n \left(\bigwedge_i \varphi^{(r)}(x_i) \wedge \bigwedge_{i<j} \delta(x_i, x_j) > 2r \right).$$

In other words, there is an r-independent set of size n, with each element satisfying $\varphi^{(r)}$.

The following is the main theorem of [12].

Theorem 3.1 (Gaifman)**.** *Every first-order sentence φ is equivalent over $\mathfrak{K}$ to a Boolean combination of basic local sentences, and this translation is effective.*

We now proceed to the proof that order-invariant CFO contains FO on arbitrary classes of finite colored graphs. First, we show that local formulae with one free variable are expressible

in CFO. Second, we show that basic local sentences are equivalent to asserting the existence of independent sets in locally defined graphs, and that the existence of independent sets can be expressed in order-invariant CFO.

To show that local formulae with one free variable are expressible in CFO, we effectively show that we can quantify over paths of length r. The proof is very similar to that of [15, Remark 2.1], so we only very briefly describe it.

Lemma 3.2. *For every r-local $\varphi^{(r)}(x)$ with one free variable, any valid S, and any $(w,i) \in S$, there exists $\varphi^{(r)}_{(w,i)}$ in* CFO_S *such that for any interpretation $\nu : X_S \to A$, $(A,\nu) \models \varphi^{(r)}_{(w,i)}$ if and only if $A \models \varphi^{(r)}(\nu(x^i_w))$.*

Proof. Fix a valuation $\nu : X_S \to A$, and let $a = \nu(x^i_w)$. Each quantifier in $\varphi^{(r)}(x)$ is bounded to $B^A_r(x)$. Following [15, Remark 2.1], each such bounded quantifier can be simulated in cluster w by quantifying over paths in A starting at x^i_w of length $s \leq r$, with the end of the path being the quantified variable (and the set of possible endpoints is exactly $B^A_r(a)$). Because all of the variables introduced in this translation are in the cluster w, arbitrary edge and equality comparisons are allowed between them. Also, note that this translation is effective.

□

The following lemma is the heart of the proof of our main containment theorem. Because we can express every r-local $\varphi^{(r)}(x)$, the only remaining thing to prove to show that we can express basic local formulae is to show that we can express the existence of $2r$-independent sets of size n with each element satisfying $\varphi^{(r)}(x)$. The proof in general ends up being very similar to showing that there exists an independent set of size n, so the following lemma is the key step of the proof, and the lemma following it generalizes it suitably.

Lemma 3.3. *For every n, there exists a sentence ι_n in order-invariant* CFO *which says that there exists an independent set of size n in a finite graph A.*

Proof. First, consider a finite connected graph A. The existence of an independent set of size n in A is equivalent to the existence of some $a \in A$ such that the ball $B^A_{2n-2}(a)$ contains an independent set of size n. Clearly the latter condition implies the former.

Conversely, suppose that A contains an independent set of size n. If the diameter of A is at most $2n-2$, then $A = B^A_{2n-2}(a)$ for every $a \in A$, so any a and our assumed independent set of size n witness the latter condition. If the diameter of A is greater than $2n-2$, we can find a path $v_0, v_1, \ldots, v_{2n-2}$ which is the shortest possible between v_0 and v_{2n-2}. Therefore, there are no edges between vertices in this path other than between successive vertices, so the set $v_0, v_2, \ldots, v_{2n-2}$ is independent and contained in $B^A_{2n-2}(v_0)$. In particular, note that this second chain of reasoning does not rely on our assumed independent set of size n.

There is a $(2n-1)$-local first-order $\beta^{(2n-1)}_n(x)$ which says that $B^A_{2n-2}(x)$ contains an independent set of size n. Similarly, for $m < n$, we can write an FO formula $\mu^{(2n-1)}_{n,m}(x)$, which references a linear order $<$ on A but is $(2n-1)$-local with respect to the edge relation on A, which says that

(1) there is no y such that $\delta^A(x,y) = 2n-1$,

(2) x is the minimal element of $B_{2n-2}(x)$ with respect to $<$, and
(3) the largest independent subset of $B^A_{2n-2}(x)$ has m elements.

Thus, $\mu^{(2n-1)}_{n,m}(x)$ defines the minimal x with respect to $<$ of a connected component whose largest independent set has size $m < n$. The correspondence between the minimality condition and Item 2 is clear, the existence of a y as described in Item 1 would imply the existence of an independent set of size $n > m$, and the non-existence of such a y implies that $B^A_{2n-2}(x)$ is equal to the connected component C, which when combined with Item 3 completes our explanation.

Via the same logic as Lemma 3.2 adapted to allow for the order comparisons in Item 2, we can write for valid S and $(w, i) \in S$ a CFO_S formula $\varphi_{\mu,n,m}$ which says that x^i_w satisfies $\mu^{(2n-1)}_{n,m}$. In particular, the order comparisons are possible because all variables in the local formula are in the same cluster.

We now describe our construction of ι_n. First, what we have established so far implies that $\exists x \beta^{(2n-1)}_n(x)$ is equivalent to the existence of a connected component in A that contains an independent set of size n, and we can write a CFO sentence $\varphi_{\beta,n}$ which is equivalent to $\exists x \beta^{(2n-1)}_n(x)$. We construct $\varphi_{\beta,n}$ as

$$\varphi_{\beta_n} \equiv \exists x^0_\varepsilon \beta^{(2n-1)}_{(\varepsilon,0)}(x^0_\varepsilon),$$

where $\beta^{(2n-1)}_{(\varepsilon,0)}(x^0_\varepsilon)$ is constructed as in Lemma 3.2. If each connected component of A has no x satisfying $\beta^{(2n-1)}_n(x)$, then for each connected component C there is some $c \in C$ and some $m < n$ such that $\mu^{(2n-1)}_{n,m}(c)$ holds. Take $p \leq n$ and $\vec{m} = (m_1, \ldots, m_p)$ such that $m_1 + \cdots + m_p \geq n$. We can write an order-invariant CFO sentence $\varphi_{\mu,\vec{m},n}$ which says that there exist $x_1 < \cdots < x_p$ such that $\mu^{(2n-1)}_{n,m_q}(x_q)$ for every x_q. The construction is

$$\varphi_{\mu,\vec{m},n} \equiv \exists x^0_\varepsilon \exists x^0_\alpha \ldots \exists x^0_{\alpha^{p-1}} \left(\bigwedge_{1 \leq q \leq p-1} x^0_{\alpha^{q-1}} < x^0_{\alpha^q} \wedge \varphi_{\mu,n,m_q}(x^0_{\alpha^{q-1}}) \wedge \varphi_{\mu,n,m_p}(x^0_{\alpha^{p-1}}) \right).$$

Finally, we have

$$\iota_n \equiv \varphi_{\beta_n} \vee \bigvee_{p,\vec{m}} \varphi_{\mu,\vec{m},n},$$

where p ranges over $1 \leq p \leq n$ and $\vec{m}$ ranges over the tuples $\vec{m} = (m_1, \ldots, m_p)$ such that $1 \leq m_i < n$ for every i and $m_1 + \cdots + m_p \geq n$. □

Lemma 3.4. *Take some r_1-local $\varphi^{(r_1)}(x)$ and some r_2-local $\psi^{(r_2)}(x, y)$. Then there exists an order-invariant* CFO *sentence $\iota_{\varphi,\psi,r_3,n}$ which says for every finite colored graph A there is an independent set of size n in the graph G with domain $\{a \in A : A \models \varphi^{(r_1)}(a)\}$ and edge relation*

$$\{(a, a') \in A^2 : A \models (\psi^{(r_2)}(a, a') \vee \psi^{(r_2)}(a', a)) \wedge \delta(a, a') \leq r_3 \wedge a \neq a'\}.$$

Proof. If y lies in $B^G_{2n-1}(x)$, then $\delta^A(x, y) \leq r_3(2n-1)$. Also, whether $z \in A$ is a vertex of G is determined by $N^A_{r_1}(z)$, so whether (z, z') is an edge in G is determined by $N^A_{\max(r_1,r_2,r_3)}(z, z')$. Therefore, the value of any $(2n-1)$-local formula $\phi^{(2n-1)}(x)$ in G is equivalent to an $r_3(2n-$

$1) + \max(r_1, r_2, r_3)$-local formula in A. Therefore the lemma follows from similar reasoning to Lemma 3.3, in which the reasoning inside of each component is almost identical and the use of the ordering between the minima of components is more or less unchanged. □

We can now complete the proof that order-invariant CFO contains FO on colored graphs.

Theorem 3.5. *Every* FO *sentence φ over σ is equivalent over the colored graphs over σ to an order-invariant* CFO *sentence ψ.*

Proof. Consider a basic local sentence

$$\exists x_1, \ldots, x_n \left(\bigwedge_i \alpha^{(r)}(x_i) \wedge \bigwedge_{i<j} \delta(x_i, x_j) > 2r \right).$$

Applying Lemma 3.4 with $\varphi^{(r_1)} = \alpha^{(r)}$, $\psi^{(r_2)}$ a tautology, and $r_3 = 2r$, we can write the above basic local sentence as an equivalent order-invariant CFO sentence. Thus, we can write any Boolean combination of basic local sentences as an equivalent order-invariant CFO sentence. The theorem then follows from Theorem 3.1. □

The generalization to local-order-invariant FO is straightforward, which completes the proof of our main containment.

Corollary 3.6. *Let φ be a local-order-invariant sentence over a class $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$ of finite colored graphs in the vocabulary $\sigma \cup \{\prec\}$. Then φ is equivalent over $\mathfrak{K}$ to an order-invariant* CFO *sentence ψ. In other words, on finite colored graphs, local-order-invariant* FO *is contained in order-invariant* CFO*.*

Proof. Applying Theorem 3.1 to φ, we have that φ is a Boolean combination of basic local sentences of the form

$$\exists x_1, \ldots, x_n \left(\bigwedge_i \alpha^{(r)}(x_i) \wedge \bigwedge_{i<j} \delta(x_i, x_j) > 2r \right),$$

where $\delta(x_i, x_j) > 2r$ and the localization of $\alpha^{(r)}(x)$ are with respect to the Gaifman graph of $(A, \prec)$. Because $B_1^{(A,\prec)}(x) = B_2^A(x)$, then $\delta^{(A,\prec)}(x_i, x_j) > 2r$ is equivalent to $\delta^A(x_i, x_j) > 4r$ and $\alpha^{(r)}(x)$ is equivalent to a formula in the vocabulary $\sigma \cup \{\prec\}$ such that the domain of the quantified variables is bounded to $B_{2r}^A(x)$. Given a linear order $<$ on A, then for each $\alpha^{(r)}(x)$, we can replace each atom $\prec(x, y, z)$ with a formula that says that x is adjacent to y, x is adjacent to z, and $y < z$, via similar logic to Lemma 3.2. Thus, via similar logic to Lemma 3.4, we have the corollary. □

We now conclude with several corollaries of our main containment theorem. As noted in the introduction, our first corollary generalizes Theorem 2.5.

Corollary 3.7. *Let φ be a local-order-invariant sentence over a class $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$ of colored graphs in the vocabulary $\sigma \cup \{\prec\}$. Then φ is equivalent over $\mathfrak{K}$ to an* FO *sentence ψ.*

Definition 3.8. Let $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$ and let $\mathcal{L}$ be a logic. The model checking problem for $\mathcal{L}$ on $\mathfrak{K}$ is the problem of determining, with input $A \in \mathfrak{K}$ and a sentence φ in $\mathcal{L}$, whether $A \models \varphi$. We say that this problem is fixed-parameter tractable with parameter φ if there is some computable function $f(\varphi)$ and some constant c such that the model-checking problem can be solved in time $f(\varphi) \cdot O(|A|^c)$.

Theorem 3.9 (Ghasemi–Grange)**.** *For every class $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$ of finite colored graphs, model checking for order-invariant* CFO *on $\mathfrak{K}$ is fixed-parameter tractable.*

Corollary 3.10. *For every class $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$ of finite colored graphs, model checking for local-order-invariant* FO *on $\mathfrak{K}$ is fixed-parameter tractable.*

Proof. Our construction of a CFO sentence expressing a basic local sentence in the vocabulary $\sigma \cup \{\prec\}$ is effective. Also, as is noted in Theorem 3.1, the translation from an arbitrary local-order-invariant FO sentence to Gaifman normal form is effective. Combining these, we have an effective translation from local-order-invariant FO sentences to order-invariant CFO sentences, so Theorem 3.9 extends to local-order-invariant FO. □

4. Generalizations via Interpretations

Our main containment result was proved for colored graphs, because this is the setting in which CFO was originally defined and in which Theorem 2.9 applies. We now generalize Theorem 2.9, Corollary 3.6, and Corollary 3.10 to general vocabularies, affirming Conjecture 2.10.

We begin by defining interpretations themselves.

Definition 4.1. A d-dimensional interpretation I between $\mathfrak{K} \subseteq \mathfrak{K}_\sigma$ and $\mathfrak{K}_\tau$ is a domain formula $\delta(\vec{x})$ with $\vec{x}$ of length d, an equivalence formula $\varepsilon(\vec{x}, \vec{y})$ with $\vec{x}, \vec{y}$ of length d, and a relation formula $\rho_R(\vec{x}_1, \dots, \vec{x}_{a_R})$ for each $R \in \tau$ with each $\vec{x}_i$ of length d. Each of these formulae is in the vocabulary σ. For a given $A \in \mathfrak{K}$, we write $D_I(A)$ for the set $\{\vec{a} \in A^d : \delta(\vec{a})\}$, and we require ε to define an equivalence relation on $D_I(A)$ for every $A \in \mathfrak{K}$, and further we require that if we have $\vec{a}_1, \dots, \vec{a}_{a_R} \in D_I(A)$ and $\vec{a}'_1, \dots, \vec{a}'_{a_R} \in D_I(A)$ such that $A \models \varepsilon(\vec{a}_i, \vec{a}'_i)$ for every i, then $A \models \rho_R(\vec{a}_1, \dots, \vec{a}_{a_R})$ if and only if $A \models \rho_R(\vec{a}'_1, \dots, \vec{a}'_{a_R})$. We define $I(A)$ as the structure over τ with underlying set $D_I(A)/\varepsilon$, with relations $R^{I(A)}([\vec{a}_1], \dots, [\vec{a}_{a_R}])$ such that $R^{I(A)}([\vec{a}_1], \dots, [\vec{a}_{a_R}])$ in $I(A)$ if and only if $A \models \rho_R(\vec{a}_1, \dots, \vec{a}_{a_R})$.

Theorem 4.2. (1) *Let σ be an arbitrary vocabulary, and let $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$. Every sentence φ which is local-order-invariant over $\mathfrak{K}$ is equivalent on $\mathfrak{K}$ to an order-invariant* CFO *sentence ψ.*

(2) *Conversely, let $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$. Every* CFO *sentence φ which is order-invariant over $\mathfrak{K}$ is equivalent over $\mathfrak{K}$ to a first-order sentence ψ.*

In particular, order-invariant CFO*, local-order-invariant* FO*, and plain* FO *are equivalent over any class of finite structures of bounded degree.*

Proof. The containment of local-order-invariant FO in order-invariant CFO on classes of finite structures over arbitrary vocabularies is an almost verbatim generalization of the proof of Corollary 3.6.

We now describe the generalization of the collapse result on classes of bounded degree. For structures with one element the translation is trivial, so we assume from here that our structures have at least two elements. Take a structure A over σ. Define the colored simple graph $\mathrm{Inc}(A)$ as follows.

(1) Use the color U for the domain of A, the color T_R for each $R \in \sigma$, and the color $P_{R,i}$ for each $1 \leq i \leq a_R$ where a_R is the arity of R.
(2) For each element a of A, there is a vertex u_a of $\mathrm{Inc}(A)$ colored U.
(3) For each relation R and tuple $\vec{a}$ of elements of A such that $A \models R(\vec{a})$, there is a vertex $t_{R,\vec{a}}$ of $\mathrm{Inc}(A)$ colored T_R.
(4) For each vertex $t_{R,\vec{a}}$, and each $1 \leq i \leq a_R$, there is a vertex $p_{R,\vec{a},i}$ of $\mathrm{Inc}(A)$ colored $P_{R,i}$.
(5) For every vertex $t_{R,\vec{a}}$ and $1 \leq i \leq a_R$, we have $E(t_{R,\vec{a}}, p_{R,\vec{a},i})$ and $E(p_{R,\vec{a},i}, u_{a_i})$ in $\mathrm{Inc}(A)$.

There is a uniform first-order interpretation I such that $I(\mathrm{Inc}(A)) \cong A$. This construction Inc preserves boundedness of degree, in the sense that for all $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$, the class $\mathrm{Inc}(\mathfrak{K}) = \{\mathrm{Inc}(A) : A \in \mathfrak{K}\}$ is of bounded degree. To see why, take $a_\sigma = \max\{a_R : R \in \sigma\}$. Then for all $a \in A$ and any $R \in \sigma$, there are at most $(d+1)^{a_\sigma}$ tuples $\vec{a}$ of elements of A such that $R(\vec{a})$ for some $R \in \sigma$. Therefore each u_a has degree at most

$$\sum_{R \in \sigma} a_R (d+1)^{a_R} \leq |\sigma| a_\sigma (d+1)^{a_\sigma},$$

each $t_{R,\vec{a}}$ has degree exactly a_R, and each $p_{R,\vec{a},i}$ has degree exactly 2, so $\mathrm{Inc}(A)$ has degree at most

$$\max(|\sigma| a_\sigma (d+1)^{a_\sigma}, 2).$$

Every φ_1 which is an order-invariant CFO sentence on $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$ can be translated to an order-invariant CFO sentence φ_2 on $\mathrm{Inc}(\mathfrak{K})$. Such a translation can proceed inductively, by encoding each atom $R(\vec{x})$ via similar logic to Lemma 3.2, translating guarded quantification to quantification over element vertices of distance exactly 4 again by similar logic to Lemma 3.2, and preserving order comparisons and unguarded quantification. By Theorem 2.9, this translated sentence is equivalent on $\mathrm{Inc}(\mathfrak{K})$ to a first-order sentence φ_3. Finally, we claim that Inc is a uniform first-order interpretation, so φ_3 can be translated to a first-order sentence φ_4 which is equivalent on $\mathfrak{K}$ to φ_1. We now describe this interpretation explicitly.

Let

$$\mathcal{S} = \{U\} \cup \{T_R : R \in \sigma\} \cup \{P_{R,i} : R \in \sigma, i \leq a_R\}.$$

The equality type of a tuple $\vec{a}$ of elements of A is the set $\{(i,j) : a_i = a_j\}$. Choose h to be large enough so that there are at least $|\mathcal{S}|$ equality types for elements of $\{0,1\}^h$. For each $s \in \mathcal{S}$, choose the equality type τ_s of some element of $\{0,1\}^h$, and construct a formula $\eta_{\tau_s}(\vec{y})$ which says that $\vec{y}$ has equality type τ_s. We define an $(a_\sigma + h)$-dimensional interpretation J between $\mathfrak{K}$ and $\mathfrak{K}_{\{E\} \cup \mathcal{S}}$ such that $J(A) \cong \mathrm{Inc}(A)$ for every $A \in \mathfrak{K}$ of size at least two. Informally, the first a_σ

coordinates will represent the elements themselves, and the latter h coordinates will represent the sort from $\mathcal{S}$. More precisely, we have the following.

(1) If $\eta_{\tau_s}(\vec{y})$ holds with $s = U$, then $\delta(\vec{x}, \vec{y})$ is true, and $\varepsilon(\vec{x}, \vec{y}, \vec{x}', \vec{y}')$ holds if and only if $x_1 = x'_1$ and $\eta_{\tau_s}(\vec{y}')$ holds.
(2) If $\eta_{\tau_s}(\vec{y})$ holds with $s = T_R$, then $\delta(\vec{x}, \vec{y})$ holds if and only if $R(x_1, \dots, x_{a_R})$ holds, and $\varepsilon(\vec{x}, \vec{y}, \vec{x}', \vec{y}')$ holds if and only if $x_i = x'_i$ for $i \leq a_R$ and $\eta_{\tau_s}(\vec{y}')$ holds.
(3) If $\eta_{\tau_s}(\vec{y})$ holds with $s = P_{R,i}$, then $\delta(\vec{x}, \vec{y})$ holds if and only if $R(x_1, \dots, x_{a_R})$ holds, and $\varepsilon(\vec{x}, \vec{y}, \vec{x}', \vec{y}')$ holds if and only if $x_i = x'_i$ for $i \leq a_R$ and $\eta_{\tau_s}(\vec{y}')$ holds.
(4) If $\eta_{\tau_s}(\vec{y})$ does not hold for any $s \in \mathcal{S}$, then $\delta(\vec{x}, \vec{y})$ does not hold.
(5) For $s \in \mathcal{S}$, the formula $\rho_s(\vec{x}, \vec{y})$ holds if and only if $\eta_{\tau_s}(\vec{y})$ holds.
(6) The formula $\rho_E(\vec{x}, \vec{y}, \vec{x}', \vec{y}')$ holds if and only if either
   (a) $x_1 = x'_i$ (or $x_i = x'_1$), $\eta_{\tau_s}(\vec{y})$ and $\eta_{\tau_{s'}}(\vec{y}')$ hold with $s = U$ and $s' = P_{R,i}$ (or vice versa); or
   (b) $x_i = x'_i$ for $i \leq a_R$, and $\eta_{\tau_s}(\vec{y})$ and $\eta_{\tau_{s'}}(\vec{y}')$ hold with $s = T_R$ and $s' = P_{R,i}$ (or vice versa).

That $J(A) \cong \mathrm{Inc}(A)$ is clear, so we are done. □

Corollary 4.3. *For any vocabulary σ, model checking for local-order-invariant* FO *on a class $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$ is fixed-parameter tractable.*

Proof. The translation from a local-order-invariant FO sentence to an order-invariant CFO sentence is effective, and the translation from φ_1 to φ_2 in Theorem 4.2 is effective, so the corollary follows from Theorem 3.9. □

A question of interest is whether local-order-invariant FO is as expressive as order-invariant FO, which is nontrivial because they have the same expressiveness on classes with Gaifman graphs of bounded diameter by a result of [4]. In particular, an answer to this question would answer the important question of whether order-invariant FO collapses to FO on classes of bounded degree. We give an equivalent formulation for this question in terms of interpretations.

Definition 4.4. We say that a logic $\mathcal{L}$ is closed under first-order interpretations if, for every interpretation I from $\mathfrak{K} \subseteq \mathfrak{K}_\sigma$ to $\mathfrak{K}_\tau$ and every $\mathcal{L}$-definable $\mathfrak{J} \subseteq I(\mathfrak{K})$, the preimage $I^{-1}(\mathfrak{J})$ is $\mathcal{L}$-definable in $\mathfrak{K}$.

The proof of the following appears in Section A.

Proposition 4.5. *Over $\mathfrak{F}_\sigma$, local-order-invariant* FO *is closed under first-order interpretations on all classes of finite structures if and only if local-order-invariant* FO *is as expressive as order-invariant* FO*.*

The following corollary follows from similar techniques. We defer the definition of tree-width and path-width to Section A, as well as the proof of the corollary.

Corollary 4.6. *If local-order-invariant* FO *collapses to* FO *on every class of bounded tree-width (resp. bounded path-width), then order-invariant* FO *collapses to* FO *on such classes as well.*

This interpretation perspective also gives a short proof of a collapse of order-invariant FO to FO when restricted to bounded shrub-depth classes. The proof is simply a combination of results of [10], [7], and [24].

Definition 4.7. For a rooted forest F, let $\overline{F}$ be the graph obtained by adding an edge between every vertex and its ancestors. The tree depth of G is the minimum height of a forest F such that $G \subseteq \overline{F}$. A class $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$ has bounded tree-depth if there is some d such that $\mathcal{G}(A)$ has tree-depth at most d for every $A \in \mathfrak{K}$.

Definition 4.8. An m-color tree-model of depth d for G consists of a rooted tree T whose leaves all have depth d, a bijection between the leaves and G, a function $c : G \to [m]$, and a symmetric map $\eta : [m]^2 \times [d] \to \{0, 1\}$. When two leaves u, v have distance $2l$ in F, we have $\eta(c(u), c(v), l) = 1$ if (u, v) is an edge in G. A class of graphs is of bounded shrub-depth if every graph in the class has a tree-model of m colors and depth d for some fixed m, d.

We use the following results from [24, 10, 7], though Theorem 4.12 had previously been claimed in [13].

Definition 4.9. A simple interpretation is a 1-dimensional interpretation with $\varepsilon(x, y) \equiv (x = y)$. A transduction T consists of an integer k, a binary relation S, unary relations $U_1, \ldots, U_n$, and a simple interpretation I_T from $\mathfrak{K} \subseteq \mathfrak{K}_{\sigma \cup \{S, U_1, \ldots, U_n\}}$ to $\mathfrak{K}_\tau$. Given $A \in \mathfrak{K}$, write A_k for the set $\{1, \ldots, k\} \times A$, and write (A_k, S) for the structure over $\sigma \cup \{S\}$ such that as structures over σ each $\{i\} \times A$ is isomorphic to A, and S is the equivalence relation on A_k such that $S((i, a), (j, a'))$ if and only if $a = a'$. We define the output $T(A)$ to be the set of structures of the form $I_T(A_k, S, U_1, \ldots, U_n)$ where each U_i is some unary relation on (A_k, S).

Theorem 4.10 (Ohlmann–Pilipczuk–Przybyszewski–Toruńczyk)**.** *Let $\mathfrak{K}$ be a bounded shrub-depth class of finite graphs. There exist*

1. *a finite relational vocabulary σ of arity at most 2,*
2. *an integer d,*
3. *a transduction T from $\mathfrak{K}$ to $\mathfrak{F}_\sigma$,*
4. *a unary P in σ, and*
5. *a first-order interpretation I from $\mathfrak{F}_\sigma$ to the class of finite graphs,*

such that, for every $G \in \mathfrak{K}$, some output $B \in T(G)$ has tree-depth at most d, the elements in P are the original vertices of G, and $D_I(B)$ is $\{b \in B : B \models P(b)\}$.

Theorem 4.11 (Eickmeyer–Elberfeld–Harwath)**.** *Let φ be an order-invariant sentence on $\mathfrak{F}_\sigma$, and let $\mathfrak{F}_\sigma^{\mathrm{td},d}$ be the class of finite σ structures of tree-depth at most d. Then there exists a first-order sentence ψ which is equivalent to φ on $\mathfrak{F}_\sigma^{\mathrm{td},d}$.*

Theorem 4.12 (Chen–Flum)**.** *On every class $\mathfrak{K}$ of finite graphs of bounded shrub-depth, every* MSO *sentence φ is equivalent on $\mathfrak{K}$ to some first-order sentence ψ.*

Theorem 4.13. *Let $\mathfrak{K}$ be a class of finite graphs of bounded shrub-depth, and let φ be order-invariant over the class of finite graphs. Then φ is equivalent on $\mathfrak{K}$ to a first-order sentence ψ.*

Proof. The proof is a relatively straightforward combination of Theorem 4.11, Theorem 4.10, Theorem 4.12, the closure of MSO under interpretations, and the definability of tree-depth in MSO. The full proof can be found in Section A. □

To attempt to generalize Theorem 4.13 past graphs, one might attempt to apply similar methods to Theorem 4.2. A first step might be to generalize Theorem 4.13 to colored graphs.

Indeed, let $\mathfrak{K}$ be a class of finite colored graphs of bounded shrub-depth, and let φ be order-invariant over the class of finite colored graphs. Then φ is equivalent on $\mathfrak{K}$ to a first-order sentence ψ. The proof of the above result generalizes to colored graphs straightforwardly. In slightly more detail, we can add finitely many unary color relations as unary predicates, track them through the transduction and interpretation from Theorem 4.10, apply Theorem 4.11 identically while taking these unary relations into account, and apply Theorem 4.12 identically while taking these unary relations into account.

Remark 4.14. However, the proof of Theorem 4.2 cannot generalize to classes of bounded shrub-depth. That is, there is no uniform encoding of finite relational structures into finitely colored graphs such that

(1) A is first-order interpretable in its encoding, and
(2) the encoding takes a class $\mathfrak{K}$ of bounded shrub-depth to a class of colored graphs of bounded shrub-depth.

In particular, there is an obstruction even for finite linear orders.

5. Deterministic Frozen-Witness Logic

We conclude with an application of our collapse of local-order-invariant FO to FO on classes of bounded degree. This recovers, on classes of bounded degree, the collapse of deterministic frozen-witness FO to FO conjectured by Abiteboul, Simon, and Vianu [2] (which was disproved on arbitrary finite structures by Otto [25]). The proof takes inspiration from the erroneous proof of [4] that epsilon-invariant FO is contained in local-order-invariant FO on classes of bounded degree. However, we apply the theorem to a weaker logic (deterministic frozen-witness logic), make substantial repairs to the proof, and remove the bounded degree hypothesis. Further, the erroneously claimed collapse of epsilon-invariant FO to FO on classes of bounded degree in [4] inherits the necessity of invariance on an ambient substructure-closed class from Theorem 2.5, whereas because we have removed that hypothesis in Theorem 4.2, our collapse inherits the lack of the hypothesis.

We begin by introducing deterministic frozen-witness logic. Our presentation is slightly non-standard, but it is clearly equivalent to the standard definition, is close to the presentation of [25], and is much more convenient for the purposes of the following proofs.

Definition 5.1. Let $\varphi(\vec{x}, y)$ be a first-order formula in the vocabulary σ, and let $A \in \mathfrak{K}_\sigma$. We say that a relation $R(\vec{x}, y)$ on A is a witness predicate for φ if it satisfies the following conditions.

(1) For every tuple $\vec{a}$ of elements of A and $a' \in A$, then $R(\vec{a}, a')$ implies $A \models \varphi(\vec{a}, a')$.
(2) If there exists $a' \in A$ such that $A \models \varphi(\vec{a}, a')$, then there is a unique $a'' \in A$ such that $R(\vec{a}, a'')$.

(3) If there does not exist $a' \in A$ such that $A \models \varphi(\vec{a}, a')$, then there is no $a'' \in A$ such that $R(\vec{a}, a'')$.

Definition 5.2. A frozen witness program over a vocabulary σ is a finite sequence of relations $R_1, \ldots, R_n$ with each R_i not in σ, a sequence of first-order formulae $\varphi_1, \ldots, \varphi_n$ with φ_i a formula in the vocabulary $\sigma \cup \{R_1, \ldots, R_{i-1}\}$, and a first-order sentence ψ in the vocabulary $\sigma \cup \{R_1, \ldots, R_n\}$. A run of the program on a structure $A \in \mathfrak{K}_\sigma$ consists of interpreting $R_1, \ldots, R_n$ in order, such that the interpretation of R_i is a witness predicate for φ_i on $(A, R_1, \ldots, R_{i-1})$. Finally, the program evaluates ψ once it has fixed all of these interpretations.

We say that a frozen witness program is deterministic on $\mathfrak{K} \subseteq \mathfrak{K}_\sigma$ if, for every $A \in \mathfrak{K}$, all runs of the program give the same evaluation of ψ. If every run evaluates ψ to be true, then we write $A \models \psi$.

The following lemma is from [4].

Lemma 5.3 (Aoki)**.** *For every r, there exists a formula $\lambda_r(x, y, z)$ in the vocabulary $\sigma \cup \{\prec\}$ which linearly orders $B_r^A(x)$ when x is fixed and when $\prec$ is a local order.*

Definition 5.4. We write $\mathrm{tp}_k(A)$ for the set of first-order sentences φ of quantifier rank at most k such that $A \models \varphi$, and we write $\mathrm{tp}_k^r(A, \vec{a})$ for the set of r-local first-order formulae $\varphi^{(r)}(\vec{x})$ of quantifier rank at most k such that $A \models \varphi^{(r)}(\vec{a})$.

The following is a well-known folklore result [22, Theorem 3.15].

Theorem 5.5. *For every $r, k, n \in \omega$, the sets $\{\mathrm{tp}_k(A) : A \in \mathfrak{K}_\sigma\}$ and*

$$\{\mathrm{tp}_k^r(A, \vec{a}) : A \in \mathfrak{K}_\sigma \text{ and } \vec{a} \in A^n\}$$

are both finite.

The following two lemmas form the core of the proof of our containment of deterministic frozen-witness FO in local-order-invariant FO. The basic idea is that, given a formula $\varphi(\vec{x}, y)$ for which we wish to define a witness predicate, we can choose a collection D of witnesses of bounded size, where the size depends on only the formula, such that for any $\vec{x}$ there is a witness y which is distance at most r from some point in either D or $\vec{x}$. Then, we use a local order to linearly order the elements distance at most r from some point in D or $\vec{x}$, and choose the minimum valid witness using this linear order. We establish the existence of D in the following lemma, and we establish this choice via local order in the lemma after.

Lemma 5.6. *Let $\theta(\vec{x}, y, \vec{c})$ be a first-order formula in some vocabulary τ. Then there are integers r, n, and a first-order formula $\gamma_\theta(\vec{c}, \vec{d})$ with $\vec{d}$ of length n such that the following hold for every $A \in \mathfrak{K}_\tau$.*

(1) *For every tuple $\vec{a}$ of elements of A of the same length as $\vec{c}$, there is some tuple $\vec{a}'$ such that $A \models \gamma_\theta(\vec{a}, \vec{a}')$.*
(2) *If $A \models \gamma_\theta(\vec{a}, \vec{a}')$, $\vec{b}$ is a tuple of elements of A, and there exists $b' \in A$ such that $\theta(\vec{b}, b', \vec{a})$, then $A \models \theta(\vec{b}, b'', \vec{a})$ for some $b'' \in B_r^A(\vec{a}, \vec{b})$ or for some $b'' \in B_{2r}^A(\vec{a}')$.*

Proof. Put $\theta(\vec{x}, y, \vec{c})$ in Gaifman normal form, let the radius be s, let $r = 2s$, and let the maximal quantifier rank of the s-local formulae in the Gaifman normal form be k. By Theorem 5.5, we can enumerate the quantifier rank k s-local types $\tau_1, \dots, \tau_m$. Let $\ell_{\vec{x}}$ be the length of $\vec{x}$, let $\ell_{\vec{c}}$ be the length of $\vec{c}$. Let $n = m(\ell_{\vec{x}} + \ell_{\vec{c}} + 1)$.

Denote by D the set $\{d_1, \dots, d_n\}$. Then $\gamma_\theta(\vec{c}, \vec{d})$ says that for each τ_i, either

(1) D contains a $2r$-independent subset of size $\ell_{\vec{x}} + \ell_{\vec{c}} + 1$ such that $\mathrm{tp}^s_k(A, d) = \tau_i$ for every d in this subset, or
(2) D contains a $2r$-independent set of size at most $\ell_{\vec{x}} + \ell_{\vec{c}}$, and this subset is maximal among the $2r$-independent subsets of $\{a \in A : \mathrm{tp}^s_k(A, a) = \tau_i\}$.

Note that this is first-order expressible because D is of bounded size.

Now suppose that $A \models \theta(\vec{b}, b', \vec{a})$ for tuples $\vec{b}, \vec{a}$ of elements of A and $b' \in A$, and suppose that we have constructed $\vec{a}'$ such that $A \models \gamma_\theta(\vec{a}, \vec{a}')$. If b' is in $B^A_r(\vec{a}, \vec{a}', \vec{b})$, then we are done. If b' is not in $B^A_r(\vec{a}, \vec{a}', \vec{b})$, then if Item 1 from our earlier dichotomy for D (where D is $\{a'_1, \dots, a'_n\}$) holds for $\tau_i = \mathrm{tp}^s_k(A, b')$, there is some $d \in D$ which is distance greater than r from every element of $\vec{b}$ and $\vec{a}$. Therefore, $\mathrm{tp}^s_k(A, \vec{b}, \vec{a}, b') = \mathrm{tp}^s_k(A, \vec{b}, \vec{a}, d)$, so $A \models \theta(\vec{b}, b', \vec{a})$ if and only if $A \models \theta(\vec{b}, d, \vec{a})$, and thus $A \models \theta(\vec{b}, d, \vec{a})$ with $d \in B_{2r}(\vec{a}, \vec{a}', \vec{b})$ because it is an element of $\vec{a}'$. If instead Item 2 from the dichotomy holds, then every $b'' \in A$ such that $\mathrm{tp}^s_k(A, b') = \mathrm{tp}^s_k(A, b'')$ is in $B^A_{2r}(\vec{a}')$ by maximality of the relevant $2r$-independent subset of D. □

Lemma 5.7. *Let r, n, and γ_θ be as in Lemma 5.6. Then there is a formula $\chi_\theta(\vec{x}, y, \vec{c}, \vec{d})$ such that for any $A \in \mathfrak{F}_\sigma$, any local order $\prec$ on A, and any $\vec{a}, \vec{a}'$ such that $(A, \prec) \models \gamma_\theta(\vec{a}, \vec{a}')$, the relation $R_{\vec{a},\vec{a}'}(\vec{x}, y) \equiv \chi_\theta(\vec{x}, y, \vec{a}, \vec{a}')$ is a witness predicate for $\theta(\vec{x}, y, \vec{a})$.*

Proof. Construct χ_θ as follows.

(1) If there are no y such that $\theta(\vec{x}, y, \vec{c})$, then set $\chi_\theta(\vec{x}, y\vec{c}, \vec{d})$ to be false for every y.
(2) If some d_i in $\vec{d}$ satisfies $\theta(\vec{x}, d_i, \vec{c})$, then let j be the least i such that $\theta(\vec{x}, d_i, \vec{c})$, and set $\chi(\vec{x}, y, \vec{c}, \vec{d})$ to be true if and only if $y = d_j$
(3) Finally, suppose that there is some y such that $\theta(\vec{x}, y, \vec{c})$, but no d_i is such a y. Then there exists some y in $B^{(A,\prec)}_r(\vec{x}, \vec{c}) \cup B^{(A,\prec)}_{2r}(\vec{d})$ such that $\theta(\vec{x}, y, \vec{c})$.

    We construct a linear order on $B^{(A,\prec)}_{2r}(\vec{x}, \vec{c}, \vec{d})$ as follows. Define the tuple $\vec{z}$ as the concatenation

$$z_1, \dots, z_k = x_1, \dots, x_m, c_1, \dots, c_l, d_1, \dots, d_n.$$

    For each $p \in B^{(A,\prec)}_{2r}(\vec{x}, \vec{c}, \vec{d})$, let $\kappa(p)$ be the least i such that $p \in B^{(A,\prec)}_{2r}(z_i)$. For $p, q \in B^{(A,\prec)}_{2r}(\vec{x}, \vec{c}, \vec{d})$, let p precede q if and only if $\kappa(p) < \kappa(q)$ or $\kappa(p) = \kappa(q)$ and $\lambda_{4r}(z_i, p, q)$ holds, with λ_{4r} as in Lemma 5.3. This works because $B^{(A,\prec)}_{2r}(\vec{a})$, for any $A \in \mathfrak{K}_\sigma$, any tuple $\vec{a}$ of elements of A, any local order $\prec$ on A, and any $r \in \omega$, is equal to $B^A_{4r}(\vec{a})$.

    Now, using this linear order, which is definable in the vocabulary $\sigma \cup \{\prec\}$, simply take the least element y of $B^{(A,\prec)}_r(\vec{x}, \vec{c}) \cup B^{(A,\prec)}_{2r}(\vec{d})$ such that $\theta(\vec{x}, y\vec{c})$.

These cases are exhaustive, and so we are done. □

We now complete the proof of our main containment theorem for this section and give several corollaries.

Theorem 5.8. *Take some deterministic frozen witness program on $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$ consisting of $R_1, \ldots, R_n$, $\varphi_1, \ldots, \varphi_n$, and ψ. Then there exists some local-order-invariant* FO *sentence ϕ on $\mathfrak{K}$ such that for every $A \in \mathfrak{K}$, $A \models \psi$ if and only if $A \models \phi$.*

Proof. We construct, by induction on i, formulae $\rho_i(\vec{x}_i, y_i, \vec{d}_{\leq i})$ in the vocabulary $\sigma \cup \{\prec\}$, where $\vec{d}_{\leq i} = (\vec{d}_1, \ldots, \vec{d}_i)$, such that for any suitable $\vec{d}_{\leq i}$ (in a sense which we will define), the relations $R_{\rho_i}(\vec{x}, y) \equiv \rho_i(\vec{x}, y, \vec{d}_{\leq i})$ are suitable interpretations for the R_i in the program.

For $i = 1$, we simply apply Lemma 5.7 to $\varphi_1(\vec{x}_1, y_1)$. This gives us a formula $\gamma_1(\vec{d}_1)$ and a formula $\rho_1(\vec{x}_1, y_1, \vec{d}_1)$ defining a witness predicate for φ_1.

Now assume that $\rho_1, \ldots, \rho_{i-1}$ have been constructed. In the formula φ_i, replace each $R_j(\vec{x}, y)$ with $\rho_j(\vec{x}, y, \vec{d}_{\leq j})$. Call this new formula $\varphi_i'(\vec{x}_i, y_i, \vec{d}_{\leq i-1})$. Then using Lemma 5.7 again, we can construct $\gamma_i(\vec{d}_{\leq i})$ and $\rho_i(\vec{x}_i, y, \vec{d}_{\leq i})$ which define a witness predicate for φ_i.

Now construct a formula $\psi'(\vec{d}_{\leq n})$ by replacing each $R_i(\vec{x}, y)$ in ψ with $\rho_i(\vec{x}, y, \vec{d}_{\leq i})$. Finally, construct the sentence

$$\psi'' \equiv \exists \vec{d}_1 \ldots \exists \vec{d}_n \left(\psi'(\vec{d}_{\leq n}) \wedge \bigwedge_{i \leq n} \gamma_i(\vec{d}_{\leq i}) \right).$$

For every $A \in \mathfrak{K}$ and every local order $\prec$ on A, there exists a tuple $\vec{d}_{\leq n}$ satisfying the conditions $\gamma_i(\vec{d}_{\leq i})$ for each i, and for any such tuple, the $\rho_i(\vec{x}_i, y_i, \vec{d}_{\leq i})$ define witness predicates, so for any local order $\prec$ on A, ψ'' says that $A \models \psi$ on a certain run of the program. Therefore, because ψ evaluates on $\mathfrak{K}$ in the same way regardless of the chosen witness predicates, $A \models \psi$ if and only if $A \models \psi''$ for any local order $\prec$, so ψ'' is equivalent to ψ and is local-order-invariant on $\mathfrak{K}$. □

Corollary 5.9. *Take some deterministic frozen witness program on $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$ consisting of $R_1, \ldots, R_n$, $\varphi_1, \ldots, \varphi_n$, and ψ. Then there exists some first-order sentence ϕ on $\mathfrak{K}$ such that for every $A \in \mathfrak{K}$, $A \models \psi$ if and only if $A \models \phi$.*

Proof. This is simply a combination of Theorem 4.2 and Theorem 5.8. □

Corollary 5.10. *For any $\mathfrak{K} \subseteq \mathfrak{F}_\sigma^d$, the model-checking problem for deterministic frozen-witness programs is fixed-parameter tractable on $\mathfrak{K}$.*

Proof. Our translation in Theorem 5.8 is effective, in particular because the translation to Gaifman normal form is effective by Theorem 3.1, so this is simply an application of Corollary 4.3. □

6. Conclusion

We have shown that order-invariant CFO contains local-order-invariant FO, affirming Conjecture 2.10. We have also shown that local-order-invariant FO contains deterministic frozen-witness FO, salvaging the proof strategy of a claimed result of [4]. Together, these prove the equality of FO, deterministic frozen-witness FO, local-order-invariant FO, and order-invariant CFO on classes of bounded degree. We also used our method of interpretations to extend the results beyond colored graphs and prove the collapse of order-invariant FO when restricted to classes of bounded shrub-depth. There are several natural remaining questions, such as whether local-order-invariant FO and order-invariant FO have the same expressiveness, or whether interpretations can be used in other ways to study invariant logics.

Appendix A. Proofs for Section 4

Proof of Proposition 4.5. First, we note that order-invariant FO is closed under interpretations. Take a d-dimensional interpretation I between $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$ and $\mathfrak{F}_\tau$ defined by δ, ε, and ρ_R for each $R \in \tau$. Given $A \in \mathfrak{K}$ and a linear order $<$ on A, we can define a linear order on $I(A)$ by first taking the lexicographic order on A^d induced by $<$ and then ordering equivalence classes of $D_I(A)/\varepsilon$ by their minimal elements. That is, we will have $[\vec{a}_1] <_{I(A)} [\vec{a}_2]$ if $\vec{a}'_1$ precedes $\vec{a}'_2$ in the lexicographic order where $\vec{a}'_1$ is the minimal element of $[\vec{a}_1]$ and $\vec{a}'_2$ is the minimal element of $[\vec{a}_2]$. This order is definable from $<$, so we can translate any order-invariant FO sentence on $I(\mathfrak{K})$ back to an order-invariant sentence on $\mathfrak{K}$. For this backwards translation, the equality of interpreted elements is translated by ε, relation atoms are translated by ρ_R, and quantifiers in the translation are relativized to $D_I(A)$.

Now, if local-order-invariant FO is not closed under interpretations, then certainly it does not have the same expressive power as order-invariant FO by the above. Assume instead that it is closed under interpretations. Take any $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$. Define the 2-dimensional interpretation H between $\mathfrak{K}$ and $\mathfrak{F}_{\sigma \cup \{E,P\}}$ with P unary and E binary in which

(1) $\delta(x, y)$ is a tautology,
(2) $\varepsilon(\vec{x}, \vec{y})$ says

$$(x_1 = x_2 \wedge y_1 = y_2 \wedge x_1 = y_1) \vee (x_1 \neq x_2 \wedge y_1 \neq y_2),$$

(3) $R^{H(A)}([\vec{a}_1], \ldots, [\vec{a}_{a_R}])$ holds if and only if each $\vec{a}_i$ is of the form (a_i, a_i) and $A \models R(a_1, \ldots, a_{a_R})$.
(4) $P^{H(A)}([\vec{a}])$ holds if and only if $a_1 \neq a_2$, and
(5) $E^{H(A)}([\vec{a}], [\vec{a}'])$ holds if and only if $(a_1 = a_2 \wedge a'_1 \neq a'_2)$ holds or $(a_1 \neq a_2 \wedge a'_1 = a'_2)$ holds.

Effectively, we have added a hub node h to A which is colored P and is adjacent via the symmetric graph relation E to every element of A. Via this E-adjacency, any order-invariant FO sentence on $\mathfrak{K}$ can be translated to a local-order-invariant FO sentence on $H(\mathfrak{K})$ which uses the linear order on the substructure of $H(A)$ isomorphic to A given by the local order on $S_1^{H(A)}(h)$. Because we have assumed closure under interpretations, then this local-order-invariant FO sentence on $H(\mathfrak{K})$ can be translated to a local-order-invariant FO sentence on $\mathfrak{K}$. □

The above techniques also allow us to give a sufficient condition for the collapse of order-invariant FO to FO on classes of bounded tree-width and bounded path-width.

Definition A.1. A tree decomposition of G is a pair $(T, (X_t)_{t \in G})$, with T a tree and $X_t \subseteq G$, such that $G = \bigcup_{t \in T} X_t$, the set $\{t \in T : v \in X_t\}$ is connected for each $v \in G$, and for every edge (u, v) in G there is some t such that $u, v \in X_t$. The width is $\max_{t \in T} |X_t| - 1$, and the tree-width of G is the minimum width of the tree decompositions of G. We say that $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$ is of bounded tree-width if there is some k such that $\mathcal{G}(A)$ is of tree-width at most k for every $A \in \mathfrak{K}$. Replacing the tree T with a path P in the above definition yields the analogous notion of path-width.

Proof of Corollary 4.6. If A has tree-width at most d, then $H(A)$ has tree-width at most $d+1$, with H our hub interpretation. An order-invariant FO sentence on a class $\mathfrak{K} \subseteq \mathfrak{F}_\sigma$ of tree-width at most d can be translated to a local-order-invariant FO sentence on the class $H(\mathfrak{K})$ of tree-width at most $d+1$, which is equivalent on $H(\mathfrak{K})$ to some FO sentence via the hypothesis of the corollary, which can be pulled back to $\mathfrak{K}$ via H. For path-width, the same argument applies. □

Proof of Theorem 4.13. Let the interpretation I from Theorem 4.10 consist of formulae $\delta(x)$ and $\rho_E(x,y)$. We can translate φ to a sentence φ' in the vocabulary $\sigma \cup \{<\}$ such that $(B,<) \models \varphi'$ if and only if $(I(B), <\restriction D_I(B)) \models \varphi$. The order-invariance of φ on finite graphs transfers to the order-invariance of φ' on $\mathfrak{F}_\sigma$. By Theorem 4.11, there is some φ'' in the vocabulary σ which is equivalent to φ' on $\mathfrak{F}_\sigma^{\mathrm{td},d}$. Let $U_1, \dots, U_n$ be the parameters of the transduction T from Theorem 4.10, and let $T(G, \vec{U})$ be the image of $(G_k, S, \vec{U})$ under the interpretation I_T from the transduction T. We can define an MSO formula $\gamma(\vec{U})$ which says that

(1) $T(G, \vec{U})$ has tree-depth at most d,
(2) the interpreted domain $\delta(T(G, \vec{U}))$ is exactly the elements marked by P corresponding to G, and
(3) for all x, y in G, the interpreted edge formula $\rho_E(x,y)$ holds in $T(G, \vec{U})$ exactly when $E(x,y)$ holds in G.

The definitions of Item 2 and Item 3 are straightforward. For Item 1, we define the MSO formulae $\varphi_{\mathrm{td},d}(X)$ inductively for each d, which will say that the induced subgraph with domain X has tree-depth at most d. First, $\varphi_{\mathrm{td},0}(X)$ will simply say that X is empty. Next, there is a MSO formula $\varphi_C(X,Y)$ which says that Y is a connected component of the induced subgraph X, an MSO formula $\eta(Y, Y', x)$ which says that as sets $Y' = Y \setminus \{x\}$, and we define

$$\varphi_{\mathrm{td},d+1}(X) \equiv \forall Y(\varphi_C(X,Y) \to \exists x(Y(x) \wedge \exists Y'(\eta(Y,Y',x) \wedge \varphi_{\mathrm{td},d}(Y')))).$$

Translate φ'' through backwards through T to an MSO sentence $\varphi''_T(\vec{U})$, and define

$$\phi \equiv \exists U_1, \dots, U_n(\gamma(\vec{U}) \wedge \varphi''_T(\vec{U})).$$

If $G \models \varphi$, we can choose $\vec{U}$ such that $T(G, \vec{U}) \models \varphi''$. On the other hand, given $\vec{U}$ such that $\gamma(\vec{U}) \wedge \varphi''_T(\vec{U})$, then $G \models \varphi$. Thus ϕ is equivalent on $\mathfrak{K}$ to φ, and ϕ is equivalent on $\mathfrak{K}$ to some first-order ψ by Theorem 4.12. □

Proof of Remark 4.14. We say that a formula $\varphi(\vec{x}, \vec{y})$ is stable on a class $\mathfrak{K} \subseteq \mathfrak{K}_\sigma$ if there is an integer $k \in \omega$ such that for every $A \in \mathfrak{K}$, there do not exist tuples $\vec{a}_1, \dots, \vec{a}_k$ of elements of A of the same length as $\vec{x}$ and $\vec{a}'_1, \dots, \vec{a}'_k$ of elements of A of the same length as $\vec{y}$ such that for all $i, j \in \{1, \dots, k\}$, we have $A \models \varphi(\vec{a}_i, \vec{a}'_j)$ if and only if $i < j$. We say that $\mathfrak{K}$ is stable if every formula is stable on $\mathfrak{K}$. Roughly, this says that we cannot interpret linear orders of arbitrary length in elements of $\mathfrak{K}$.

The class of finite linear orders is not stable. Linear orders have complete Gaifman graphs, so they form a class of bounded shrub-depth. Stability is preserved under first-order interpretations, in the sense that if $\mathfrak{K}$ is stable then $I(\mathfrak{K})$ is stable. If an encoding as described in the remark existed, the class of the colored graphs encoding linear orders would be of bounded shrub-depth.

However, every graph class of bounded shrub-depth is stable (this is a consequence of the results in [14] that was first observed in [13]). □

References

[1] Serge Abiteboul, Richard Hull, and Victor Vianu. *Foundations of Databases*. Addison-Wesley, Reading, MA, 1995.

[2] Serge Abiteboul, Eric Simon, and Victor Vianu. Non-deterministic languages to express deterministic transformations. In *Proceedings of the Ninth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS 1990)*, pages 218–229. ACM, 1990. doi:10.1145/298514.298575.

[3] Serge Abiteboul and Victor Vianu. Non-determinism in logic-based languages. *Annals of Mathematics and Artificial Intelligence*, 3(2–4):151–186, 1991. doi:10.1007/BF01530924.

[4] Derek Aoki. Local-order-invariant logic on classes of bounded degree, 2025. arXiv:2512.02039.

[5] Michael Benedikt and Luc Segoufin. Towards a characterization of order-invariant queries over tame graphs. *Journal of Symbolic Logic*, 74(1):168–186, 2009. doi:10.2178/jsl/1231082307.

[6] Ashok K. Chandra and David Harel. Computable queries for relational data bases. *Journal of Computer and System Sciences*, 21(2):156–178, 1980. doi:10.1016/0022-0000(80)90032-X.

[7] Yijia Chen and Jörg Flum. FO-definability of shrub-depth. In *28th EACSL Annual Conference on Computer Science Logic (CSL 2020)*, volume 152 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 15:1–15:16. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CSL.2020.15.

[8] Bruno Courcelle. The monadic second-order logic of graphs X: Linear orders. *Theoretical Computer Science*, 160(1–2):87–143, 1996. doi:10.1016/0304-3975(95)00083-6.

[9] Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite Model Theory*. Springer Monographs in Mathematics. Springer, Berlin, second edition, 2005. doi:10.1007/3-540-28788-4.

[10] Kord Eickmeyer, Michael Elberfeld, and Frederik Harwath. Succinctness of order-invariant logics on depth-bounded structures. *ACM Transactions on Computational Logic*, 18(4):33:1–33:25, 2017. doi:10.1145/3152770.

[11] Ronald Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In Richard M. Karp, editor, *Complexity of Computation*, volume 7 of *SIAM-AMS Proceedings*, pages 43–73. American Mathematical Society, Providence, RI, 1974.

[12] Haim Gaifman. On local and non-local properties. In J. Stern, editor, *Proceedings of the Herbrand Symposium: Logic Colloquium '81*, volume 107 of *Studies in Logic and the Foundations of Mathematics*, pages 105–135. North-Holland, Amsterdam, 1982. doi:10.1016/S0049-237X(08)71879-2.

[13] Jakub Gajarský and Petr Hliněný. Kernelizing MSO properties of trees of fixed height, and some consequences. *Logical Methods in Computer Science*, 11(1:19):1–26, 2015. doi:10.2168/LMCS-11(1:19)2015.

[14] Robert Ganian, Petr Hliněný, Jaroslav Nešetřil, Jan Obdržálek, Patrice Ossona de Mendez, and Reshma Ramadurai. When trees grow low: Shrubs and fast MSO_1. In *Mathematical Foundations of Computer Science 2012*, volume 7464 of *Lecture Notes in Computer Science*, pages 419–430. Springer, 2012. doi:10.1007/978-3-642-32589-2_38.

[15] Fatemeh Ghasemi and Julien Grange. Order-invariant cluster first-order logic on graph classes of bounded degree, 2026. arXiv:2604.27693.

[16] Julien Grange. Successor-invariant first-order logic on classes of bounded degree. *Logical Methods in Computer Science*, 17(3):20:1–20:25, 2021. doi:10.46298/lmcs-17(3:20)2021.

[17] Julien Grange and Luc Segoufin. Order-invariant first-order logic over hollow trees. In *28th EACSL Annual Conference on Computer Science Logic (CSL 2020)*, volume 152 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 23:1–23:16. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CSL.2020.23.

[18] Neil Immerman. Relational queries computable in polynomial time (extended abstract). In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing (STOC 1982)*, pages 147–152. ACM, 1982. doi:10.1145/800070.802187.

[19] Neil Immerman. Languages that capture complexity classes. *SIAM Journal on Computing*, 16(4):760–778, 1987. doi:10.1137/0216051.

[20] Neil Immerman. Expressibility and parallel complexity. *SIAM Journal on Computing*, 18(3):625–638, 1989. doi:10.1137/0218043.

[21] Neil Immerman. *Descriptive Complexity*. Graduate Texts in Computer Science. Springer, New York, 1999. doi:10.1007/978-1-4612-0539-5.

[22] Leonid Libkin. *Elements of Finite Model Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, Berlin, 2004. doi:10.1007/978-3-662-07003-1.

[23] Steven Lindell, Henry Towsner, and Scott Weinstein. Hanf locality and invariant elementary definability, 2025. arXiv:2507.12450.

[24] Pierre Ohlmann, Michał Pilipczuk, Wojciech Przybyszewski, and Szymon Toruńczyk. Canonical decompositions in monadically stable and bounded shrubdepth graph classes. In *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*, volume 261 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 135:1–135:17. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ICALP.2023.135.

[25] Martin Otto. Epsilon-logic is more expressive than first-order logic over finite structures. *Journal of Symbolic Logic*, 65(4):1749–1757, 2000. doi:10.2307/2695073.

[26] Benjamin Rossman. Successor-invariant first-order logic on finite structures. *Journal of Symbolic Logic*, 72(2):601–618, 2007. doi:10.2178/jsl/1185803625.

[27] Moshe Y. Vardi. The complexity of relational query languages (extended abstract). In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing (STOC 1982)*, pages 137–146. ACM, 1982. doi:10.1145/800070.802186.

[28] Scott Weinstein. Some questions about invariant definability. In Anuj Dawar, Erich Grädel, Phokion G. Kolaitis, and Thomas Schwentick, editors, *Finite and Algorithmic Model Theory (Dagstuhl Seminar 17361)*, volume 7 of *Dagstuhl Reports*, pages 24–25. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2018. doi:10.4230/DagRep.7.9.1.

Proof School, 221 Main St., San Francisco, California, 94103, USA
Email address: derekaoki@acm.org