Logical Methods in Computer Science
Volume 17, Issue 1, 2021, pp. 2:1–2:16
https://lmcs.episciences.org/

Submitted Apr. 02, 2018
Published Jan. 22, 2021

TAMENESS IN LEAST FIXED-POINT LOGIC AND MCCOLM'S CONJECTURE

SIDDHARTH BHASKAR AND ALEX KRUCKMAN

Department of Computer Science, University of Copenhagen, Denmark
e-mail address: sbhaskar@di.ku.dk

Department of Mathematics and Computer Science, Wesleyan University, USA
e-mail address: akruckman@wesleyan.edu

Abstract. We investigate four model-theoretic tameness properties in the context of least fixed-point logic over a family of finite structures. We find that each of these properties depends only on the elementary (i.e., first-order) limit theory, and we completely determine the valid entailments among them. In contrast to the context of first-order logic on arbitrary structures, the order property and independence property are equivalent in this setting.

McColm conjectured that least fixed-point definability collapses to first-order definability exactly when proficiency fails. McColm's conjecture is known to be false in general. However, we show that McColm's conjecture is true for any family of finite structures whose limit theory is model-theoretically tame.

1. Introduction

Least fixed-point (LFP) logic is obtained by extending first-order (FO) logic by a quantifier denoting the least fixed-point of a relational operator. The difference between FO and LFP definability over classes of finite structures is a central question in finite model theory. McColm conjectured that the existence of arbitrarily long elementary inductions would suffice to separate LFP from FO [McC90]. This conjecture was refuted by two separate constructions due to Gurevich, Immerman, and Shelah [GIS94]. However, instances of this conjecture remain interesting: in particular, any resolution of the *ordered conjecture*, which states that LFP is more expressive than FO over every class of totally ordered structures, would resolve a longstanding open problem in computational complexity [KV92].

Recent work, e.g., Adler and Adler [AA14], has shown that certain model-theoretic tameness properties introduced by Shelah [She90] are relevant to finite model theory, generalizing assumptions like *bounded cliquewidth* and *bounded treewidth* which permit, e.g., fast algorithms for formula evaluation [Cou90]. Here we study four such properties, NOP, NIP, NSOP, and NTP_2. We show that any counterexample to McColm's conjecture must fail all of them, thus establishing in a precise sense that any such counterexample must be complicated (Theorem 2.18).

Key words and phrases: Least fixed-point logic, inductive definability, finite model theory, model-theoretic dividing lines.

DOI:10.23638/LMCS-17(1:2)2021

This result was foreshadowed by McColm himself, who observed in [McC90] that any counterexample to his conjecture must fail another model-theoretic tameness property, viz., the negation of the finite cover property (NFCP). Our work complements Lindell and Weinstein [LW00], who show that any counterexample cannot be *recursion-theoretically* tame.

In the course of this investigation, we formulate LFP versions of these four properties. In FO logic, NOP $\Rightarrow$ NSOP and NOP $\Rightarrow$ NIP $\Rightarrow$ NTP_2, but no other entailments hold in general between these four properties. By contrast, in LFP logic over finite structures, we have NOP $\Leftrightarrow$ NIP $\Rightarrow$ NTP_2 $\Rightarrow$ NSOP (Corollary 3.2). Moreover, both of the implications in NIP $\Rightarrow$ NTP_2 $\Rightarrow$ NSOP are strict (Theorem 3.5). We find the equivalence NOP $\Leftrightarrow$ NIP especially remarkable: intuitively, it says that "order implies randomness" in this context.

Finally, we find that each of the properties of families of finite structures that we study, viz., FO = LFP, proficiency, and LFP-(NOP, NIP, NSOP, NTP_2), depend only on the elementary limit theory of the class of structures (Lemma 2.2, Corollary 2.7, and Corollary 2.17). This relatively innocuous observation seems not to have been explicitly mentioned before, but the resulting shift in perspective, from classifying structures to classifying theories, brings these questions closer to the spirit of classical model theory.

1.1. **Least fixed-point logic.** We assume familiarity with FO and LFP definability, and we very briefly review the latter. (See, e.g., Libkin [Lib04] for a reference.) Ordinary (first-order) variables are denoted by lowercase Latin letters, e.g., x, y, z. Relational (second-order) variables are denoted by uppercase Latin letters, e.g., P, Q, R, S. Every relational variable comes with an arity, but this is not made explicit in the notation.

Following the typical model-theoretic convention, we will also use, e.g., x, to denote a tuple of (first-order) variables, not just a single variable. We write $|x|$ to indicate the length of the tuple x. Below, when we write, e.g., $\varphi(x, S)$, we mean that φ is a formula, x is a tuple of first-order variables (which includes all free first-order variables in φ), and S is a single second-order variable which is free in φ (but φ may have other free second-order variables).

We use boldface letters, especially $\mathbf{A}$ and $\mathbf{M}$, to denote structures. Their respective domains are denoted in lightface, e.g., A and M. We denote a class of structures by $\mathcal{C}$; its elements are always assumed to share a common signature.

Definition 1.1. A formula $\varphi(x, S)$ is *positive elementary in* S if each occurrence of S is in the scope of an even number of negations. In addition, $\varphi(x, S)$ is *operative* if it is positive elementary in S and the arity of S is $|x|$.

An operative formula $\varphi(x, S)$ is so called because for any structure $\mathbf{A}$, given together with interpretations of all free second-order variables in φ other than S, φ defines a monotone operator $A^{|x|} \to A^{|x|}$ by

$$R \mapsto \{a \in A^{|x|} \mid \mathbf{A} \models \varphi(a, R)\}.$$

For every ordinal α, we define

$$\varphi^\alpha = \begin{cases} \emptyset & \text{if } \alpha = 0 \\ \{a \in A^{|x|} \mid \mathbf{A} \models \varphi(a, \varphi^\beta)\} & \text{if } \alpha = \beta + 1 \\ \bigcup_{\beta<\alpha} \varphi^\beta & \text{if } \alpha \text{ is a limit} \end{cases}$$

The relations φ^α are called the *stages* of φ on $\mathbf{A}$. The *closure ordinal* $\|\varphi\|_{\mathbf{A}}$ is the least ordinal Γ such that $\varphi^\Gamma = \varphi^{\Gamma+1}$. (Note that if A is finite, $\|\varphi\|_{\mathbf{A}}$ must be finite.) The relation φ^Γ is the least fixed-point of the monotone operator defined by φ, and is written φ^∞.

Definition 1.2. The class of LFP *formulas* is obtained from the class of atomic formulas (in first- and second-order variables) by closing under boolean connectives, first-order quantifiers, and the least fixed-point quantifier: If $\varphi(x, S)$ is an operative LFP formula and t is a tuple of terms of length $|x|$, then $[\mathbf{lfp}\, Sx.\varphi](t)$ is an LFP formula, in which the free first-order variables are those in t and the free second-order variables are those in φ, except for S, which is bound by the quantifier.

The standard semantics of first-order logic are extended to the least fixed-point quantifier as follows: A structure $\mathbf{A}$ (given together with an interpretation of the free first- and second-order variables) satisfies $[\mathbf{lfp}\, Sx.\varphi](t)$ if and only if the interpretation of the tuple of terms t is in the relation φ^∞.

Definition 1.3. A *query* $R(x)$ of arity n over a class of structures $\mathcal{C}$ is an isomorphism-invariant family of n-ary relations $R^A \subseteq A^n$ for each $\mathbf{A} \in \mathcal{C}$. It is LFP-*definable* in case it is defined by some LFP formula with no free second-order variables, uniformly over all structures in $\mathcal{C}$. An important special case is a *boolean* query, whose arity is zero. LFP-definable boolean queries are defined by LFP sentences.

We say LFP = FO over a class of structures $\mathcal{C}$ if every query which is LFP-definable over $\mathcal{C}$ is defined by a first-order formula. Otherwise, we say LFP > FO over $\mathcal{C}$.

Definition 1.4. An operative formula $\varphi(x, S)$ is *basic* if it is first-order (i.e., it does not contain any instances of the least fixed-point quantifier) and it has no free second-order variables other than S.

Remark 1.5. We are primarily concerned with definability over families of finite structures. Immerman [Imm86], building on the work of Moschovakis [Mos74], proves the following normal form for LFP formulas over finite structures:

$$(Qy)\,([\mathbf{lfp}\, Sx.\varphi](t))$$

for a basic operative formula φ, a tuple of terms t, and a string of first-order quantifiers Qy binding some of the free variables in t.

When working with a class of finite structures, this normal form allows us to restrict attention to LFP formulas containing a single least fixed-point quantifier. In particular, we only need to consider least fixed-point quantification of basic operative formulas. On the other hand, when we want to show that a particular relation is LFP-definable, we will freely make use of the full syntax in Definition 1.2.

Definition 1.6. Operative formulas $\varphi(x, S)$ and $\psi(y, T)$ are *complementary on finite structures* in case $|x| = |y|$ and for every finite structure $\mathbf{A}$,

$$\mathbf{A} \models \forall z\ ([\mathbf{lfp}\, Sx.\varphi](z) \leftrightarrow \neg[\mathbf{lfp}\, Ty.\psi](z))\,.$$

Fact 1.7 [Imm86]**.** *For every basic operative formula $\varphi(x, S)$, there exists a basic operative formula $\psi(y, T)$, such that $\varphi(x, S)$ and $\psi(y, T)$ are complimentary on finite structures.*

1.2. **Proficiency and McColm's conjecture.**

Definition 1.8. We say that a class of finite structures $\mathcal{C}$ is *proficient* if there exists a basic operative formula $\varphi(x, S)$ such that

$$\sup\{\|\varphi\|_{\mathbf{A}} : \mathbf{A} \in \mathcal{C}\} = \omega.$$

Remark 1.9. For any basic operative formula $\varphi(x, S)$ and any finite n, the stage φ^n is definable by a first-order formula, uniformly over all structures. It follows immediately that if a class of finite structures $\mathcal{C}$ is not proficient, then LFP = FO over $\mathcal{C}$. This observation was made originally in [McC90].

In 1990, McColm [McC90] conjectured that the following three properties are equivalent, for any family of finite structures $\mathcal{C}$:

(1) $\mathcal{C}$ is proficient.
(2) FO < LFP over $\mathcal{C}$.
(3) FO < $\mathcal{L}^{\omega}_{\infty\omega}$ over $\mathcal{C}$.

The implication from 2 to 1 is Remark 1.9, and the implication from 3 to 1 is also easy. In 1992, Kolaitis and Vardi [KV92] proved the equivalence of 1 and 3. In 1994, Gurevich, Immerman, and Shelah [GIS94] constructed two examples of a proficient family of structures for which FO = LFP, thus establishing that 1 does not imply 2.

Historically, the equivalence of 1 and 2 has been called *McColm's first conjecture* and the equivalence of 1 and 3 *McColm's second conjecture* (see, e.g., [KV92]). In the interests of brevity, we will simply say *McColm's conjecture* to mean the equivalence of 1 and 2, following the usage in [GIS94].

1.3. **Elementary limit theories.**

In model theory, the most important invariant of a structure is its theory. In the present paper, the most important invariant of a family of finite structures is its limit theory.

Definition 1.10. Let $\mathcal{C}$ be a class of finite structures. For a first-order sentence φ, we write $\mathcal{C} \models \varphi$ in case all but finitely many structures in $\mathcal{C}$ satisfy φ. The *(elementary) limit theory* of a family $\mathcal{C}$ of finite structures is

$$\mathrm{Th}(\mathcal{C}) = \{\varphi \mid \varphi \text{ is a first-order sentence, and } \mathcal{C} \models \varphi\}.$$

Unlike the theory $\mathrm{Th}(\mathbf{A})$ of a structure $\mathbf{A}$, limit theories are not always complete — nor consistent! But it is easy to see that $\mathrm{Th}(\mathcal{C})$ is consistent if and only if $\mathcal{C}$ is infinite. Henceforth, we will only consider infinite families of finite structures.

Lemma 1.11. *For any class of finite structures $\mathcal{C}$, $\mathrm{Th}(\mathcal{C})$ is closed under logical consequence.*

Proof. Suppose $\mathrm{Th}(\mathcal{C}) \models \varphi$. By compactness, $\Delta \models \varphi$ for some finite subset $\Delta \subseteq \mathrm{Th}(\mathcal{C})$. Each sentence in Δ is true in all but finitely many structures in $\mathcal{C}$, and Δ is finite, so all but finitely many structures in $\mathcal{C}$ satisfy Δ, and hence satisfy φ. Thus $\varphi \in \mathrm{Th}(\mathcal{C})$. □

Definition 1.12. A first-order theory T has the *finite model property* if for every sentence φ, if $T \models \varphi$, then φ has a finite model.

Lemma 1.13. *Let T be a countable first-order theory. Then T has the finite model property if and only if $T \subseteq \mathrm{Th}(\mathcal{C})$ for some infinite class of finite structures $\mathcal{C}$.*

Proof. Suppose T has the finite model property. Enumerate T as $\{\varphi_n \mid n \in \mathbb{N}\}$, and let $\psi_n = \bigwedge_{i=0}^{n} \varphi_i$ for each $n \in \mathbb{N}$. Then $T \models \psi_n$, so ψ_n has a finite model $\mathbf{A}_n$ (which we may assume to be distinct from $\mathbf{A}_m$ for all $m < n$). Letting $\mathcal{C} = \{A_n \mid n \in \mathbb{N}\}$, each sentence $\varphi_n \in T$ is satisfied by $\mathbf{A}_m$ for all $m \geq n$, so $\mathcal{C} \models \varphi_n$, and hence $T \subseteq \mathrm{Th}(\mathcal{C})$.

Conversely, suppose $T \subseteq \mathrm{Th}(\mathcal{C})$ for some infinite class of finite structures $\mathcal{C}$. If $T \models \varphi$, then $\varphi \in \mathrm{Th}(\mathcal{C})$, by Lemma 1.11. Since $\mathcal{C}$ is infinite, there exists some $\mathbf{A} \in \mathcal{C}$ such that $\mathbf{A} \models \varphi$, so T has the finite model property. □

1.4. **Model-theoretic dividing lines.** Stability theory originated in the work of Shelah in his program to classify the models of certain complete first-order theories. He discovered a robust division of theories into "stable" and "unstable." The former are "tame" in the sense that their definable sets are highly structured, in a way that makes it possible to classify their models (under certain additional hypotheses); the latter are "wild" in the sense that they interpret combinatorial objects such as infinite linear orders and random graphs, and they necessarily have too many models to admit a nice classification.

A critical observation about the stable/unstable dichotomy is that stability can be defined by the absence of a simple combinatorial configuration in definable sets, namely the *order property* (described below). Subsequent work in "neo-stability" has generalized stability theory to increasingly more inclusive notions of tameness, each of which is defined by the absence of a particular configuration in definable sets. The goal is to find robust dividing lines, such that it is possible to prove structure theorems on the tame side and non-structure theorems on the wild side. For an interactive guide to the these dividing lines, see [Con]; see Hodges [Hod87] for a discussion of structure theorems.

Any formula $\varphi(x;y)$, whose free variables are partitioned into a tuple x and a tuple y, defines a bipartite graph relation R_φ between $M^{|x|}$ and $M^{|y|}$ for any structure $\mathbf{M}$, by $(a,b) \in R_\varphi$ if and only if $\mathbf{M} \models \varphi(a;b)$. From another point of view, such a formula defines a family of S_φ of subsets of $M^{|x|}$: writing $\varphi(M;b)$ for $\{a \in M^{|x|} \mid \mathbf{M} \models \varphi(a;b)\}$, we let $S_\varphi = \{\varphi(M;b) \mid b \in M^{|y|}\}$. By a combinatorial configuration, we usually mean some concrete property of the graph R_φ or the family of sets S_φ.

We will now give the precise definitions of the combinatorial properties we will consider in this paper, all of which are originally due to Shelah [She90]. See below for further discussion.

Definition 1.14. Let $\varphi(x;y)$ be any formula (FO or LFP), whose free variables are partitioned into a tuple x and a tuple y. Let n be a natural number, and let $\mathbf{M}$ be a structure.

- φ has an *n-instance of the order property* (OP(n)) in $\mathbf{M}$ if there exist tuples $a_1,\dots,a_n \in M^{|x|}$ and $b_1,\dots,b_n \in M^{|y|}$ such that $\mathbf{M} \models \varphi(a_i;b_j)$ if and only if $i \le j$.
- φ has an *n-instance of the independence property* (IP(n)) in $\mathbf{M}$ if there exist tuples $a_i \in M^{|x|}$ for all $i \in \{1,\dots,n\}$ and $b_X \in M^{|y|}$ for all $X \subseteq \{1,\dots,n\}$ such that $\mathbf{M} \models \varphi(a_i;b_X)$ if and only if $i \in X$.
- φ has an *n-instance of the strict order property* (SOP(n)) in $\mathbf{M}$ if there exist tuples $b_1,\dots,b_n \in M^{|y|}$ such that $\varphi(M;b_i) \subseteq \varphi(M;b_j)$ if and only if $i \le j$.
- We say $\varphi(x;y)$ has an *n-instance of the tree property of the second kind* ($\mathrm{TP}_2(n)$) in $\mathbf{M}$ if there are tuples $b_{i,j} \in M^{|y|}$ for $1 \le i,j \le n$ such that for any i and any $j \ne k$, $\varphi(M;b_{i,j}) \cap \varphi(M;b_{i,k}) = \emptyset$, but for any function $f\colon \{1,\dots,n\} \to \{1,\dots,n\}$,

$$\bigcap_{i=1}^{n} \varphi(M;b_{i,f(i)}) \ne \emptyset.$$

Definition 1.15. Let $\mathcal{C}$ be a class of structures. For each property P in $\{\mathrm{OP},\mathrm{IP},\mathrm{SOP},\mathrm{TP}_2\}$, we say $\mathcal{C}$ *has* LFP-P (resp. FO-P) if there exists an LFP formula (resp. an FO formula) $\varphi(x;y)$ such that for each n, there exists a structure $\mathbf{M} \in \mathcal{C}$ such that φ has $P(n)$ in $\mathbf{M}$.

Let T be a theory. We say that T has FO-P if the class of models of T has FO-P.

If a class of structures $\mathcal{C}$ or a theory T does not have (FO/LFP)-OP (resp. IP, SOP, TP_2), we say it has (FO/LFP)-NOP (resp. NIP, NSOP, NTP_2).

Remark 1.16. Our definitions of these properties differ from the standard definitions in model theory, which consist of a single infinite configuration, rather than a sequence of finite configurations. For example, according to the standard definition, a theory T has the order property if there exists an FO formula $\varphi(x;y)$, a model $\mathbf{M} \models T$, and tuples $(a_n)_{n\in\mathbb{N}}$ in $M^{|x|}$ and $(b_n)_{n\in\mathbb{N}}$ in $M^{|y|}$ such that $\mathbf{M} \models \varphi(a_i, b_j)$ if and only if $i \leq j$.

In the context of a first-order theory T, our definitions are equivalent to the standard ones, by an application of the compactness theorem. But compactness is not available in the context of LFP definability, and the standard infinitary definitions are not meaningful over classes of finite structures.

Another advantage of using finite configurations is that the presence or absence of an n-instance of one of our properties in a structure $\mathbf{M}$ is expressible by a single sentence. For example, $\varphi(x;y)$ has $\mathrm{OP}(n)$ in $\mathbf{M}$ if and only if

$$\mathbf{M} \models \exists x_1 \ldots \exists x_n \exists y_1 \ldots \exists y_n \left(\bigwedge_{i\leq j} \varphi(x_i;y_j) \wedge \bigwedge_{i>j} \neg\varphi(x_i;y_j) \right).$$

For $P \in \{\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP}_2\}$, we denote by $P_\varphi(n)$ the sentence expressing that $\varphi(x;y)$ has $P(n)$. This leads immediately to the following lemma.

Lemma 1.17. *Let $\mathcal{C}$ be a class of finite structures. For any P in $\{\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP}_2\}$, $\mathcal{C}$ has* FO-P *if and only if* $\mathrm{Th}(\mathcal{C})$ *has* FO-P.

Proof. Suppose $\mathcal{C}$ does not have FO-P. Then for every FO formula $\varphi(x;y)$, there is some $n \in \mathbb{N}$ such that for all $\mathbf{M} \in \mathcal{C}$, $\varphi(x;y)$ does not have $P(n)$ in $\mathbf{M}$. That is, $\neg P_\varphi(n)$ is true in every structure in $\mathcal{C}$, so $\neg P_\varphi(n) \in \mathrm{Th}(\mathcal{C})$. Thus $\varphi(x;y)$ does not have $P(n)$ in any model of T, so T does not have FO-P.

Conversely, suppose $\mathrm{Th}(\mathcal{C})$ does not have FO-P. Then for every FO formula $\varphi(x;y)$, there is some $n \in \mathbb{N}$ such that for all $\mathbf{M} \models \mathrm{Th}(\mathcal{C})$, $\varphi(x;y)$ does not have $P(n)$ in $\mathbf{M}$. By Lemma 1.11, $\neg P_\varphi(n) \in \mathrm{Th}(\mathcal{C})$, so $\varphi(x;y)$ does not have $P(n)$ in any structures in $\mathcal{C}$, except for finitely many exceptions. Each of these exceptional structures are finite, so there is some maximum N such that $\varphi(x;y)$ has $P(N)$ in any structure in $\mathcal{C}$. Thus $\mathcal{C}$ does not have FO-P. ☐

Intuitively, a formula $\varphi(x;y)$ has the order property if arbitrarily long linear orders are represented in the bipartite graph R_φ, in the sense that the "half-graphs" appear as induced subgraphs: $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ with $a_i R_\varphi b_j$ if and only if $i \leq j$. The independence property and the strict order property are two natural strengthenings of this condition: IP is equivalent to the condition that *arbitrary* bipartite graphs appear as induced subgraphs of R_φ, and SOP says that arbitrarily long linear orders appear as chains in the family of sets S_φ.

It is not hard to see that any formula with $\mathrm{IP}(n)$ or $\mathrm{SOP}(n+1)$ in a structure has $\mathrm{OP}(n)$ in that structure. At the level of of complete first-order theories, the converse is true (but the same formula need not serve as the witness). This important dichotomy is due to Shelah: an unstable (OP) theory must exhibit order (SOP) or randomness (IP).

Fact 1.18 [She90, Theorem II.4.7]**.** *A first-order theory T has* FO-OP *if and only if it has* FO-IP *or* FO-SOP.

The tree property of the second kind is admittedly somewhat less intuitive. Roughly speaking, a formula $\varphi(x;y)$ has the tree property of the second kind if the family of sets S_φ

includes arbitrarily many arbitrarily large families of disjoint sets, such that these families interact independently. It is not hard to see that any formula with $\mathrm{TP}_2(n)$ in a structure has $\mathrm{IP}(n)$ in that structure.

The name TP_2 comes from another important dichotomy identified by Shelah: a theory is called *simple* if it does not have the tree property (TP), and a theory has the tree property if and only if it has the tree property of the first kind (TP_1) or the tree property of the second kind (TP_2). Unlike TP_2, the configurations defining the properties TP and TP_1 are visibly related to trees. We will not consider the properties TP and TP_1 in this paper.

1.5. **Examples via Fraïssé limits.** Fraïssé theory is a fruitful source of examples in model theory and provides an important connection between classes of finite structures and the model theory of complete first-order theories. If a class $\mathcal{C}$ of finite structures is isomorphism-closed and countable up to isomorphism, and has the hereditary property, the joint embedding property, and the amalgamation property, then it admits a unique countable *Fraïssé limit* $\mathbf{M}_\mathcal{C}$. This structure is universal and homogeneous for $\mathcal{C}$, in the sense that a finite structure is in $\mathcal{C}$ if and only if it embeds in $\mathbf{M}_\mathcal{C}$, and any two such embeddings are conjugate by an automorphism of $\mathbf{M}_\mathcal{C}$. The theory $\mathrm{Th}(\mathbf{M}_\mathcal{C})$ is called the *generic theory* of $\mathcal{C}$. For more on Fraïssé theory, see [Hod93, Section 7.1].

Here are some examples of generic theories, and which properties from Definition 1.14 they do and do not satisfy:

- T_∞, the theory of an infinite set with no additional structure. This is the generic theory of the class of finite sets. It has FO-NOP (and hence FO-NIP, FO-NSOP, and FO-NTP_2).
- DLO, the theory of dense linear orders without endpoints. This is the generic theory of the class of finite linear orders. It has FO-SOP (and hence FO-OP), but FO-NIP (and hence FO-NTP_2).
- T_{rg}, the theory of the random graph. This is (by definition) the generic theory of the class of finite graphs. It has FO-IP (and hence FO-OP), but FO-NSOP and FO-NTP_2.
- T_{org}, the generic theory of the class of finite graphs equipped with an ordering of their vertices. It has FO-IP and FO-SOP (and hence FO-OP), but FO-NTP_2.
- T_{feq}, the generic theory of the class of finite parameterized equivalence relations. The language consists of two unary predicates O and P (for "objects" and "parameters"), and one ternary relation symbol E. A parameterized equivalence relation is a structure such that O and P partition the domain, and for every element a satisfying P, the binary relation $E(a, x, y)$ is an equivalence relation on the elements satisfying O. This theory has FO-TP_2 (and hence FO-IP and FO-OP) but FO-NSOP.
- T_{aba}, the theory of atomless Boolean algebras. This is the generic theory of the class of finite Boolean algebras. It has FO-SOP and FO-TP_2 (and hence FO-IP and FO-OP).

We will return to several of these examples in Section 3 below.

2. FO characterizations of LFP properties

Many important properties of a class of finite structures $\mathcal{C}$ depend only on the elementary limit theory $\mathrm{Th}(\mathcal{C})$, in the sense that any two classes of structures with the same limit theory agree on the property in question. Lemma 1.17 shows this holds of the FO properties from Definition 1.14. In this section, we show that it additionally holds for proficiency, "FO = LFP over $\mathcal{C}$," and all the LFP properties from Definition 1.14.

2.1. **Proficiency and** FO = LFP**.** By Remark 1.9, each finite stage φ^n of a basic operative formula $\varphi(x, S)$ is definable by a first order formula, uniformly over all structures. We also use $\varphi^n(x)$ to denote any such formula. The context that the symbol φ^n appears in will distinguish whether we mean the formula or the relation it defines.

Definition 2.1. A theory T is *not proficient* in case for every basic operative formula $\varphi(x, S)$, there is a natural number n such that

$$T \models (\forall x)(\varphi^n(x) \leftrightarrow \varphi^{n+1}(x)).$$

Otherwise, we say that T is *proficient*.

Lemma 2.2. *Let $\mathcal{C}$ be a class of finite structures. Then $\mathcal{C}$ is proficient if and only if* $\mathrm{Th}(\mathcal{C})$ *is proficient.*

Proof. Suppose that $\mathcal{C}$ is not proficient. Then for each basic operative formula $\varphi(x, S)$, there is a bound $n \in \mathbb{N}$ such that $\|\varphi\|_{\mathbf{A}} \leq n$ for all $\mathbf{A} \in \mathcal{C}$. Therefore, for all $\mathbf{A} \in \mathcal{C}$, $\mathbf{A} \models (\forall x)(\varphi^n(x) \leftrightarrow \varphi^{n+1}(x))$, and hence $\mathrm{Th}(\mathcal{C}) \models (\forall x)(\varphi^n(x) \leftrightarrow \varphi^{n+1}(x))$.

Conversely, suppose that $\mathrm{Th}(\mathcal{C})$ is not proficient, and let φ be any basic operative formula. Then, there is some $n \in \mathbb{N}$ such that $\mathrm{Th}(\mathcal{C}) \models (\forall x)(\varphi^n(x) \leftrightarrow \varphi^{n+1}(x))$. Hence, for all but finitely many $\mathbf{A} \in \mathcal{C}$, $\mathbf{A} \models (\forall x)(\varphi^n(x) \leftrightarrow \varphi^{n+1}(x))$, and therefore $\|\varphi\|_{\mathbf{A}} \leq n$ for all but finitely many $\mathbf{A} \in \mathcal{C}$. Since there are only finitely many exceptional structures, $\sup\{\|\varphi\|_{\mathbf{A}} : \mathbf{A} \in \mathcal{C}\}$ must still be finite. □

Corollary 2.3. *Let $\mathcal{C}$ and $\mathcal{D}$ be classes of finite structures. If* $\mathrm{Th}(\mathcal{C}) = \mathrm{Th}(\mathcal{D})$*, then $\mathcal{C}$ is proficient if and only if $\mathcal{D}$ is proficient.*

The proof of Lemma 2.2 actually shows that if $\mathrm{Th}(\mathcal{C}) \subseteq \mathrm{Th}(\mathcal{D})$, and $\mathcal{D}$ is proficient, then $\mathcal{C}$ is proficient. Similar refinements, replacing equality of limit theories with containment, can be observed for many of the results in this section.

Definition 2.4. Let $\varphi(x, S)$ be a basic operative formula, and let $\mathcal{C}$ be a class of finite structures. We say that φ^∞ *is elementary over* $\mathcal{C}$ if there is a first-order formula $\gamma(x)$ which defines the query φ^∞ over $\mathcal{C}$.

Lemma 2.5. *φ^∞ is elementary over $\mathcal{C}$ if and only if there is a first-order formula $\gamma(x)$ which defines the query φ^∞ over all but finitely many structures in $\mathcal{C}$.*

Proof. One direction is trivial. For the other direction, suppose $\gamma(x)$ defines φ^∞ over all but finitely many structures in $\mathcal{C}$. Since every finite structure is determined up to isomorphism by a first-order sentence, and every automorphism-invariant relation on a finite structure is definable by a first-order formula, we can modify $\gamma(x)$ so that it defines φ^∞ in each of the finitely many exceptional cases. □

Note that the sentence

$$\forall x\, ([\mathbf{lfp}\, Sx.\varphi](x) \leftrightarrow \gamma(x)),$$

which expresses that γ defines φ^∞, is not first-order. So it does not follow directly from Lemma 2.5 that elementarity of φ^∞ over $\mathcal{C}$ is a property of the limit theory $\mathrm{Th}(\mathcal{C})$. Nevertheless, this turns out to be true, as we will now show.

Lemma 2.6. *Let $\varphi(x, S)$ be a basic operative formula. Then φ^∞ is elementary over $\mathcal{C}$ if and only if there exists a first-order formula $\theta(x)$ such that*

$$\forall x\, (\varphi(x, \theta) \leftrightarrow \theta) \in \mathrm{Th}(\mathcal{C}), \textit{ and}$$

$$\forall x\, (\psi(x, \neg\theta) \leftrightarrow \neg\theta) \in \mathrm{Th}(\mathcal{C}),$$

where $\psi(x,S)$ is a basic operative formula which is complementary to $\varphi(x,S)$ on finite structures.

Proof. Suppose φ^∞ is elementary relative to $\mathcal{C}$, witnessed by θ. In an arbitrary structure $\mathbf{A} \in \mathcal{C}$, θ defines φ^∞, which is a fixed-point for φ. Since $\psi(x,S)$ and $\varphi(x,S)$ are complementary on finite structures, $\neg\theta$ defines ψ^∞, which is a fixed-point for ψ. So $\mathbf{A}$ satisfies the sentences in the statement of the lemma.

Conversely, suppose these two sentences are in $\mathrm{Th}(\mathcal{C})$. Then for all but finitely many structures $\mathbf{A} \in \mathcal{C}$, the relation $\theta^{\mathbf{A}}$ defined by θ over $\mathbf{A}$ is a fixed-point of φ, and its complement $(\neg\theta)^{\mathbf{A}}$, which is defined by $\neg\theta$, is a fixed point of ψ. Since φ^∞ and ψ^∞ are the least fixed-points of φ and ψ, $\varphi^\infty \subseteq \theta^{\mathbf{A}}$ and $\psi^\infty \subseteq (\neg\theta)^{\mathbf{A}}$. Since φ^∞ and ψ^∞ are complements, $\varphi^\infty = \theta^{\mathbf{A}}$. This is true for all but finitely many structures in $\mathcal{C}$, so φ^∞ is elementary over $\mathcal{C}$ by Lemma 2.5. □

Corollary 2.7. *If* $\mathrm{Th}(\mathcal{C}) = \mathrm{Th}(\mathcal{D})$*, then* LFP = FO *over* $\mathcal{C}$ *if and only if* LFP = FO *over* $\mathcal{D}$*.*

Proof. Suppose that every LFP-definable query over $\mathcal{C}$ is FO-definable. Consider an arbitrary LFP-definable query R over $\mathcal{D}$. By the normal form for LFP formulas over finite structures (Remark 1.5), we may assume that R is defined by the LFP formula $(Qy)\,([\mathbf{lfp}\, Sx.\varphi](t))$. Then it suffices to show that φ^∞ is elementary over $\mathcal{D}$.

By assumption, φ^∞ is elementary over $\mathcal{C}$. But $\mathrm{Th}(\mathcal{C}) = \mathrm{Th}(\mathcal{D})$, so by Lemma 2.6, φ^∞ is elementary over $\mathcal{D}$. The converse follows in the same way. □

Remark 2.8. Even though FO = LFP is a property of limit theories, we have *not* proven (and in fact it is not true) that FO = LFP over $\mathcal{C}$ if and only if FO = LFP over the class of models of $\mathrm{Th}(\mathcal{C})$. This stands in contrast to the properties FO-P and proficiency, which do pass (Lemmata 1.17 and 2.2) between $\mathcal{C}$ and models of $\mathrm{Th}(\mathcal{C})$.

We learned Lemmata 2.2 and 2.6 from Steven Lindell through personal communication. As an immediate consequence of Lemma 2.2 and Corollary 2.7, we deduce the following.

Corollary 2.9. *If* $\mathrm{Th}(\mathcal{C}) = \mathrm{Th}(\mathcal{D})$*, then* $\mathcal{C}$ *satisfies McColm's conjecture if and only if* $\mathcal{D}$ *does.*

Remark 2.10. If we were to define the LFP *limit theory* to be the set of all LFP sentences that hold of all but finitely many structures in $\mathcal{C}$, then both FO = LFP and proficiency could easily seen to depend only on the LFP limit theory. One might naturally wonder whether the LFP limit theory itself depends only on the (elementary) limit theory; this would contain Lemma 2.2 and Corollary 2.7 as special cases.

However, this is not the case. For example, the family of even-sized linear orders and the family of odd-sized linear orders are two families with the same (elementary) limit theory: the complete theory of infinite discrete linear orders with endpoints. But they are distinguished by their LFP limit theories, since parity of the domain is an LFP-definable boolean query over ordered structures.

2.2. **Model-theoretic dividing lines.** We will now show that each LFP property defined in Definition 1.14 depends only on the elementary limit theory of a family of finite structures. We start by identifying proficiency with LFP-SOP. Key to this argument is the LFP-definability of the *stage comparison relation* [Mos74].

Definition 2.11. For any basic operative formula $\varphi(x, S)$ with $|x| = k$, any structure $\mathbf{A}$, and any $a \in A^k$, the *stage of* a, $\|a\|_\varphi$, is the least ordinal α such that $a \notin \varphi^\alpha$, or ∞ if $\alpha \notin \varphi^\infty$. The *stage comparison relation* $\preceq_\varphi$ is defined by $a \preceq_\varphi b$ if and only if $\|a\|_\varphi \leq \|b\|_\varphi$.

Fact 2.12 [Mos74, Theorem 2A.2]**.** *For any basic operative formula* $\varphi(x, S)$*, the stage comparison relation* $\preceq_\varphi$ *is* LFP*-definable over the class of all structures.*

By a *partial preorder*, we mean a reflexive transitive relation $\preceq$ on some set X. We get a partial order if we take the quotient by the equivalence relation $x \preceq y \wedge y \preceq x$. A partial preorder is *linear* if the associated partial order is linear. By a chain in a partial preorder $\preceq$, we mean a subset X which is linearly ordered by $\preceq$. In particular, for $x, y \in X$, $x \preceq y$ and $y \preceq x$ implies $x = y$.

Note that for any basic operative formula $\varphi(x, S)$ with $|x| = k$ and any structure $\mathbf{A}$, the stage comparison relation $\preceq_\varphi$ is always a linear preorder on A^k, whose associated linear order is a well-order.

Theorem 2.13. *Let* $\mathcal{C}$ *be a class of finite structures. The following are equivalent:*

(1) $\mathcal{C}$ *is proficient.*
(2) *There is some* $n \in \mathbb{N}$ *and some* LFP *formula* ψ *such that* ψ *defines a linear preorder on* n*-tuples in every structure in* $\mathcal{C}$*, and this linear preorder has arbitrarily long finite chains in structures in* $\mathcal{C}$*.*
(3) *There is some* $n \in \mathbb{N}$ *and some* LFP *formula* ψ *such that* ψ *defines a partial preorder on* n*-tuples in every structure in* $\mathcal{C}$*, and this partial preorder has arbitrarily long finite chains in structures in* $\mathcal{C}$*.*

Proof. $(1 \Rightarrow 2)$ Suppose $\varphi(x, S)$ is a basic operative formula with $|x| = n$, which witnesses proficiency of $\mathcal{C}$. Its stage comparison relation $\preceq_\varphi$, which is LFP-definable by Fact 2.12, linearly preorders the n-tuples from each $\mathbf{M} \in \mathcal{C}$, and this linear preorder contains a chain of length $\|\varphi\|_{\mathbf{M}}$. By proficiency, there is no finite bound on the lengths of these chains.

$(2 \Rightarrow 3)$ Trivially.

$(3 \Rightarrow 1)$ Suppose $\lambda(y_1; y_2)$ defines a partial preorder $\preceq$ which has arbitrarily long finite chains in structures in $\mathcal{C}$. In any finite partial preorder, we define the *height* of an element to be one more than the maximum height among its (strict) predecessors, or 0 if it has none. Since λ has arbitrarily long chains, elements in the preorder defined by λ will have arbitrarily large heights.

Let $\varphi(y; T)$ say that all of y's strict predecessors are in T. In symbols,

$$\varphi(y; T) \equiv \forall y' \left((\lambda(y', y) \wedge \neg\lambda(y, y')) \rightarrow T(y') \right).$$

(Notice T occurs positively in φ.) Then it is easy to show by induction that the stages φ^n of φ are exactly those elements of height $< n$, and hence φ witnesses that $\mathcal{C}$ is proficient. □

Theorem 2.14. $\mathcal{C}$ *is proficient if and only if it has* LFP-SOP*.*

Proof. Suppose that the LFP formula φ witnesses that $\mathcal{C}$ has LFP-SOP. Define

$$\psi(y_1; y_2) \equiv \forall x \ (\varphi(x; y_1) \rightarrow \varphi(x; y_2)),$$

so that for each $\mathbf{M} \in \mathcal{C}$ and $b_1, b_2 \in M^{|y|}$,

$$\mathbf{M} \models \psi(b_1; b_2) \text{ if and only if } \varphi(M; b_1) \subseteq \varphi(M; b_2),$$

where (as in Section 1.4), $\varphi(M; b) = \{a \in M^{|x|} \mid \mathbf{M} \models \varphi(a; b)\}$. Then ψ defines a partial preorder on $M^{|y|}$. If φ has SOP(n) in M, then this partial preorder on $M^{|y|}$ contains a chain of length n. Since φ has the strict order property, we have arbitrarily long chains in structures in $\mathcal{C}$, so $\mathcal{C}$ is proficient by Theorem 2.13.

Conversely, suppose that $\mathcal{C}$ is proficient. By Theorem 2.13, there exists an LFP formula $\lambda(y_1; y_2)$ that defines a partial preorder with arbitrarily long chains in structures in $\mathcal{C}$. The formula λ itself witnesses the strict order property: given n, pick $\mathbf{M} \in \mathcal{C}$ which contains a chain $b_1, \ldots, b_n$. Then

$$\lambda(M^{|y_1|}; b_1) \subsetneq \cdots \subsetneq \lambda(M^{|y_1|}; b_n). \qquad \square$$

Among all the LFP properties from Definition 1.14, the strict order property turns out to be the strongest, in that in entails all the others. This contrasts with the first-order case where, in general, the strict order property does not imply the independence property or the tree property of the second kind.

Lemma 2.15. *If $\mathcal{C}$ has* LFP-SOP*, then it also has* LFP-OP*,* LFP-IP*, and* LFP-TP$_2$*.*

Proof. As noted in Section 1.4, LFP-SOP easily implies LFP-OP. For the other two properties, we consider the family $\mathcal{N}$ of all finite linear orders. Identify the unique linear order of size n with the set $n = \{0, 1, \ldots, n-1\}$ equipped with its natural ordering. It is well known that over $\mathcal{N}$, the graphs of addition and multiplication are LFP-definable; hence, so is the graph of exponentiation [Lin]. Therefore, since the relations $\mathrm{bit}(x; y)$:

"the x-th bit of y base 2 is 1"

and $\mathrm{factor}(x; y, z)$:

"y^z is the largest power of y dividing x"

are first-order definable over $\mathcal{N}$ with addition, multiplication, and exponentiation, they are LFP-definable over $\mathcal{N}$.

The relation $\mathrm{bit}(x; y)$ has $\mathrm{IP}(n)$ in m for sufficiently large m, witnessed by $a_i = i - 1$ for $i \in \{1, \ldots, n\}$ and $b_X = \sum_{j \in X} 2^{j-1}$ for $X \subseteq \{1, \ldots, n\}$. The relation $\mathrm{factor}(x; y, z)$ has $\mathrm{TP2}(n)$ in $(m, <)$ for sufficiently large m, witnessed by $b_{i,j} = (p_i, j)$, where $(p_i)_{i \in \omega}$ is an enumeration of the primes: for any function $f \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, we have $a_f = \prod_{i=1}^n p_i^{f(i)} \in \mathrm{factor}(M; p_i, f(i))$ for all $1 \le i \le n$. Hence, $\mathcal{N}$ has LFP-IP and LFP-TP$_2$.

Now suppose $\mathcal{C}$ has LFP-SOP. Since there is some LFP formula ψ which defines a linear preorder on n-tuples with arbitrarily long chains in structures in $\mathcal{C}$ (by Theorem 2.13 and Theorem 2.14), we can repeat the constructions of bit and factor above to get formulas witnessing that $\mathcal{C}$ has LFP-IP and LFP-TP$_2$.

To be a little more concrete, suppose $\varphi(x; y)$ witnesses LFP-IP or LFP-TP$_2$ over $\mathcal{N}$. Simply replace each variable v in φ by n variables $v_1, \ldots, v_n$, replace $v = w$ by $(v \le w \wedge w \le v)$, replace $v \le w$ by $\psi(v_1, \ldots, v_n; w_1, \ldots, w_n)$, and proceed by induction on the construction of φ in the obvious way. This gives us a new formula $\varphi^\star(x^\star; y^\star)$ witnessing LFP-IP or LFP-TP$_2$ over $\mathcal{C}$. $\square$

Theorem 2.16. *For any class $\mathcal{C}$ of finite structures and any P in $\{\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP}_2\}$, $\mathcal{C}$ has* LFP-P *if and only if $\mathcal{C}$ is proficient or $\mathcal{C}$ has* FO-P*.*

Proof. In the forwards direction, if $\mathcal{C}$ has LFP-P, but is not proficient, then FO = LFP over $\mathcal{C}$ (Remark 1.9), so $\mathcal{C}$ has FO-P.

Conversely, if $\mathcal{C}$ is proficient, then it has LFP-P by Theorems 2.14 and 2.15. Otherwise, if $\mathcal{C}$ has FO-P, then it trivially has LFP-P as well. $\square$

Theorem 2.16, when combined with Lemmata 1.17 and 2.2, has the immediate consequence that all the LFP properties depend only on the elementary limit theory of a family of structures.

Corollary 2.17. *Suppose $\mathcal{C}$ and $\mathcal{D}$ are families of structures with the same limit theory, and let P be any property in* $\{\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP}_2\}$. *Then $\mathcal{C}$ has* LFP-P *if and only if $\mathcal{D}$ has* LFP-P.

It also has the nice consequence that any tame class of structures (in the sense of first-order model theory) satisfies McColm's conjecture.

Theorem 2.18. *For any family of finite structures $\mathcal{C}$, if $\mathcal{C}$ has* FO-NOP, FO-NIP, FO-NTP$_2$, *or* FO-NSOP, *then $\mathcal{C}$ satisfies McColm's conjecture.*

Proof. Let $P \in \{\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP}_2\}$ be a property such that $\mathcal{C}$ does not have FO-P. To show that $\mathcal{C}$ satisfies McColm's conjecture, it suffices to show that if $\mathcal{C}$ is proficient, then LFP $\neq$ FO over $\mathcal{C}$.

So we assume $\mathcal{C}$ is proficient. Then $\mathcal{C}$ has LFP-P, by Theorem 2.14 and Lemma 2.15. Since it does not have FO-P, LFP $\neq$ FO over $\mathcal{C}$, as desired. □

3. Entailments between LFP properties

We continue towards determining all valid entailments among the LFP properties. First, we show that an important dichotomy remains true in the LFP context.

Theorem 3.1. *For any family $\mathcal{C}$ of finite structures, $\mathcal{C}$ has* LFP-OP *if and only if $\mathcal{C}$ has* LFP-SOP *or* LFP-IP.

Proof. Since both LFP-SOP and LFP-IP entail LFP-OP, it suffices to show the forwards direction.

Suppose $\mathcal{C}$ has LFP-OP. If $\mathcal{C}$ is proficient, then it has LFP-SOP by Theorem 2.14. If $\mathcal{C}$ is not proficient, then $\mathcal{C}$ has FO-NSOP and FO-OP by Theorem 2.17, so Th($\mathcal{C}$) has FO-NSOP and FO-OP by Lemma 1.17. It follows that Th($\mathcal{C}$) has FO-IP by Fact 1.18. Hence $\mathcal{C}$ has FO-IP, and therefore LFP-IP. □

By Theorems 2.15, 3.1, and propositional reasoning, we obtain the following.

Corollary 3.2. *For any family $\mathcal{C}$ of finite structures,*

$$\mathcal{C} \models \text{LFP-OP} \iff \mathcal{C} \models \text{LFP-IP}.$$

Therefore,

$$\mathcal{C} \models \text{LFP-SOP} \implies \mathcal{C} \models \text{LFP-TP}_2 \implies \mathcal{C} \models \text{LFP-OP} \iff \mathcal{C} \models \text{LFP-IP}.$$

We would like to give examples showing that the first two implications above are strict. To do this, we will employ countably categorical first-order theories with the finite model property.

A theory is *countably categorical* if it has only one countable model up to isomorphism. Various equivalent formulations of countable categoricity were proven in the 50's and 60's by Ryll-Nardzewski, Svenonius, and Engler. These established countable categoricity as a robust and important property of first-order theories. Fraïssé theory is an important source of examples of countably categorical theories: every Fraïssé limit in a finite relational language has a countably category complete theory. For more information and history on countable categoricity, see Chapter 7 of Hodges [Hod93].

Though seemingly separate notions, proficiency and countable categoricity are intimately related: roughly speaking, non-proficiency is the "finite variable logic version" of countable categoricity. The Ryll-Nardzewski theorem [Hod93, Theorem 7.3.1] asserts that countable categoricity is equivalent to the finiteness of the set of complete n-types over T, for all n. Non-proficiency essentially weakens this condition to the finiteness of the set of all n-types *in m-variable logic*, for all n and $m \geq n$ [DLW96, Theorem 23]. We now prove more carefully that countable categoricity implies non-proficiency, for first-order theories.

Lemma 3.3. *Every countably categorical first-order theory is non-proficient.*

Proof. Suppose that T is countably categorical, fix a basic operative formula $\varphi(x, S)$, and consider the first-order formulas $\varphi^n(x)$ defining its stages. By the Ryll-Nardzewski theorem, there are only finitely many pairwise non-T-equivalent formulas with free variables from x. Thus there must be some $m \in \mathbb{N}$ and $n < m$ such that $T \models (\forall x)(\varphi^n(x) \leftrightarrow \varphi^m(x))$. Since, for all j, $T \models (\forall x)(\varphi^j(x) \rightarrow \varphi^{j+1}(x))$, it must be the case that $T \models (\forall x)(\varphi^m(x) \leftrightarrow \varphi^{m+1}(x))$, and $\|\varphi\|_{\mathbf{M}} \leq m$ for all models $\mathbf{M} \models T$.

Since φ was chosen arbitrarily, T is non-proficient. □

Remark 3.4. In particular, if T is a countable and countably categorical theory with the finite model property, then there is some family of finite structures $\mathcal{C}$ with limit theory T, by Lemma 1.13. Since T is countably categorical, it is non-proficient, and hence so is $\mathcal{C}$, by Lemma 2.2. Therefore, FO = LFP over $\mathcal{C}$. In addition, $\mathcal{C}$ inherits any property FO-P (or its negation) from T itself, by Lemma 1.17.

To complete the classification, we show that both of the one-way implications in Corollary 3.2 are strict.

Theorem 3.5. *There exists a class of finite structures with* LFP-IP *but without* LFP-TP_2 *and a class of finite structures with* LFP-TP_2 *but without* LFP-SOP*.*

Proof. To exhibit a class of structures with a certain combination of LFP-properties, it suffices to exhibit a class of structures with the same combination of FO-properties, over which FO = LFP. By Lemma 3.3 and Remark 3.4, it suffices to exhibit a countable, complete, and countably categorical theory with the finite model property, with the same combination of FO-properties. This is exactly what we do. See Section 1.4 for definitions of our example theories. In this proof, we drop the prefix FO-.

For IP but NTP_2, consider T_{RG}, the theory of the random graph. This is well-known to be countably categorical with the finite model property, to have IP, and to be simple; a first-order theory is simple if it does not have the tree property (TP), which implies that it does not have TP_2.

For TP_2 but NSOP, consider T_{feq}, the generic theory of parameterized equivalence relations. For discussions of this theory, see [CR16] and [Kru19]. In [CR16], Chernikov and Ramsey establish (Corollary 6.20) that T_{feq} does not have the property SOP_1, which implies that it does not have SOP, and (Corollary 6.18) that T_{feq} is not simple by witnessing TP_2 directly. A proof that T_{feq} has the finite model property is given in [Kru19]. □

We conclude with a list of some simple examples satisfying the various combinations of properties we have discussed in this paper (see Figure 1). Even though the LFP properties in each box in the table are not explicit, we can easily deduce them: in the column LFP = FO, the LFP properties agree with the FO properties, and in the column LFP $\neq$ FO, each family of structures is proficient, and hence (by Theorem 2.16) satisfies each of LFP-SOP, LFP-TP_2, LFP-IP, and LFP-OP. Since we have established (Lemma 2.15) that there are no classes satisfying (NIP and SOP and LFP = FO) or (IP and NTP_2 and SOP and LFP = FO), the table is complete.

Figure 1: Examples

FO properties	LFP = FO	LFP $\neq$ FO
NOP = (NIP and NSOP)	$\mathbb{N}_{\mathrm{fin}}$	$(\mathbb{N}_{\mathrm{fin}}, S)$
NIP and SOP	—	$(\mathbb{N}_{\mathrm{fin}}, <)$
IP and NTP_2 and NSOP	RG	$\mathrm{RG} + (\mathbb{N}_{\mathrm{fin}}, S)$
IP and NTP_2 and SOP	—	$\mathrm{RG} + (\mathbb{N}_{\mathrm{fin}}, <)$
TP_2 and NSOP	PEQ	$\mathrm{PEQ} + (\mathbb{N}_{\mathrm{fin}}, S)$
TP_2 and SOP	GIS	$\mathrm{PEQ} + (\mathbb{N}_{\mathrm{fin}}, <)$

Here are the definitions of the classes appearing in Figure 1:

- $\mathbb{N}_{\mathrm{fin}}$ is the class of initial segments of $\mathbb{N}$ with no extra structure. $(\mathbb{N}_{\mathrm{fin}}, S)$ and $(\mathbb{N}_{\mathrm{fin}}, <)$ are the classes of structures with the same domains, but equipped with the successor relation and the order relation, respectively.
- RG is any class of finite structures with limit theory T_{RG}, the theory of the random graph. The class of Paley graphs provides an explicit example (see [BEH81]).
- PEQ is any class of finite structures with limit theory T^*_{feq}, the generic theory of parameterized equivalence relations. Such a class exists by Lemma 1.13.
- GIS is any counterexample to McColm's conjecture. For example, one of the classes of finite structures devised by Gurevich, Immerman, and Shelah in [GIS94].
- Given classes of finite structures $\mathcal{C} = \{\mathbf{M}_i \mid i \in \omega\}$ and $\mathcal{C}' = \{\mathbf{M}'_i \mid i \in \omega\}$ in disjoint languages L and L', respectively, we denote by $\mathcal{C} + \mathcal{C}'$ the family $\{\mathbf{M}_i \sqcup \mathbf{M}'_i \mid i \in \omega\}$, where $\mathbf{M}_i \sqcup \mathbf{M}'_i$ is the disjoint union of $\mathbf{M}_i$ and $\mathbf{M}'_i$. We use the fact that for any property P in $\{\mathrm{SOP}, \mathrm{TP}_2, \mathrm{IP}, \mathrm{OP}\}$, $\mathcal{C} + \mathcal{C}'$ has FO-P if and only if $\mathcal{C}$ has FO-P or $\mathcal{C}'$ has FO-P.

4. Further work

Our results suggest that it may be fruitful to examine SOP, TP_2, IP, and OP beyond the first-order context. In particular, it would be interesting to examine the extend to which weaker fixed-point logics (like transitive closure logic) recover TP_2 and IP from SOP. Another direction is a program to recover complexity-theoretic tameness properties of families of finite structures (like fast formula evaluation) from model-theoretic tameness assumptions, generalizing assumptions like bounded treewidth and cliquewidth.

In the spirit of classification theory, we might hope to deduce some positive concrete information about, e.g., LFP-NOP classes of finite structures that distinguish them from the merely stable (FO-NOP) classes. One might hope to develop some kind of asymptotic structure theory (like Shelah's classification of models of certain stable theories) for finite classes which are stable and non-proficient.

Finally, we believe that the observation that FO = LFP is a property of the elementary limit theory of a class of finite structures strongly suggests a model-theoretic approach to difficult questions like the ordered conjecture. At the very least, it gives us a new set of powerful tools to test and clarify where, exactly, the difficulty lies.

Acknowledgements

We started this project in 2017 when we were both postdocs at Indiana University, Bloomington. We would like to thank Larry Moss and the logic group at IU for their support.

We would also like to acknowledge several people who have read earlier versions of this paper and discussed this material with us. In particular, we thank Cameron Hill for his comments on the first draft of this paper, Steve Lindell and Scott Weinstein for being invaluable sources of knowledge of finite model theory, and Greg McColm for inspiring the present line of inquiry. Finally, we are indebted to the anonymous referees, whose comments substantially improved the readability of this paper.

References

[AA14] Hans Adler and Isolde Adler. Interpreting nowhere dense graph classes as a classical notion of model theory. *European J. Combin.*, 36:322–330, 2014.

[BEH81] Andreas Blass, Geoffrey Exoo, and Frank Harary. Paley graphs satisfy all first-order adjacency axioms. *J. Graph Theory*, 5(4):435–439, 1981.

[Con] Gabriel Conant. Map of the universe. `https://forkinganddividing.com/`.

[Cou90] Bruno Courcelle. The monadic second-order logic of graphs. I. Recognizable sets of finite graphs. *Inform. and Comput.*, 85(1):12–75, 1990.

[CR16] Artem Chernikov and Nicholas Ramsey. On model-theoretic tree properties. *J. Math. Log.*, 16(2):1650009, 2016.

[DLW96] Anuj Dawar, Steven Lindell, and Scott Weinstein. First order logic, fixed point logic and linear order. In *Computer science logic (Paderborn, 1995)*, volume 1092 of *Lecture Notes in Comput. Sci.*, pages 161–177. Springer, Berlin, 1996.

[GIS94] Yuri Gurevich, Neil Immerman, and Saharon Shelah. Mccolm's conjecture [positive elementary inductions]. In Samson Abramsky, editor, *Proceedings of the Ninth Annual IEEE Symp. on Logic in Computer Science, LICS 1994*, pages 10–19. IEEE Computer Society Press, July 1994.

[Hod87] Wilfrid Hodges. What is a structure theory? *Bull. London Math. Soc.*, 19(3):209–237, 1987.

[Hod93] Wilfrid Hodges. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[Imm86] Neil Immerman. Relational queries computable in polynomial time. *Inform. and Control*, 68(1-3):86–104, 1986.

[Kru19] Alex Kruckman. Disjoint n-amalgamation and pseudofinite countably categorical theories. *Notre Dame J. Form. Log.*, 60(1):139–160, 2019.

[KV92] P.G. Kolaitis and M.Y. Vardi. Fixpoint logic vs. infinitary logic in finite-model theory. In Andre Scedrov, editor, *Proceedings of the Seventh Annual IEEE Symp. on Logic in Computer Science, LICS 1992*, pages 46–57. IEEE Computer Society Press, June 1992.

[Lib04] Leonid Libkin. *Elements of finite model theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2004.

[Lin] Steven Lindell. Exponentiation is elementarily definable from addition and multiplication on finite structures. `http://ww3.haverford.edu/cmsc/slindell/exponentiation.pdf`.

[LW00] Steven Lindell and Scott Weinstein. The role of decidability in first order separations over classes of finite structures. In Martin Abadi, editor, *Proceedings of the Fifteenth Annual IEEE Symp. on Logic in Computer Science, LICS 2000*, pages 45–50. IEEE Computer Society Press, June 2000.

[McC90] Gregory L. McColm. When is arithmetic possible? *Ann. Pure Appl. Logic*, 50(1):29–51, 1990.

[Mos74] Yiannis N. Moschovakis. *Elementary induction on abstract structures*. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1974. Studies in Logic and the Foundations of Mathematics, Vol. 77.

[She90] S. Shelah. *Classification theory and the number of nonisomorphic models*, volume 92 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, second edition, 1990.