Tameness in least fixed-point logic and McColm's conjecture

Siddharth Bhaskar [1] Alex Kruckman [2]

[1]Datalogisk Institut, Københavns Universitet

[2]Department of Mathematics and Computer Science, Wesleyan University

Motivating questions

- What does classification theory say about finite model theory/descriptive complexity theory, and vice versa?
- What is the interaction of model-theoretic tameness with classical open problems in FM/DCT?
- To what extent are there robust dividing lines among classes of finite models, and not just for first-order definability?
- Main motivating question: difference between FO and LFP logic on classes of finite structures.

Least fixed-point logic

- Consider first-order L-formulas $\varphi(\vec{x}, S)$ with second-order (relation) variables S.
- If the length of $\vec{x}$ and the arity of S agree, then φ defines a functional $\mathcal{P}(A^n) \to \mathcal{P}(A^n)$ for any L-structure A
- The *stages* of an operative positive elementary L-formula φ are defined by
  - $\varphi^0 = \emptyset$
  - $\varphi^{n+1} = \{x : \varphi(x, \varphi^n)\}$
  - $\varphi^\alpha = \bigcup_{\beta<\alpha} \varphi^\beta$
- If S occurs *positively*, then this functional is monotone, so has a least fixed-point, over every L-structure

Least fixed-point logic

- For an L-structure A, the *closure ordinal* $\|\varphi\|_A$ is the least Γ such that $A \models \varphi^\Gamma = \varphi^{\Gamma+1}$.
- LFP formulas are obtained by extending FO logic by the *LFP quantifier*,

$$(\mathbf{lfp}_{\vec{x},S}\,\varphi)(\vec{t}) := \varphi^\Gamma$$

  for any operative, positive elementary $\varphi(\vec{x}, S)$
- For $x_1, x_2 \in A$, we say $x_1 \prec x_2$ in case x_1 "appears in" φ^α before x_2. (This is the stage comparison preorder over φ)
- Moschovakis: the stage comparison preorder of any LFP formula $\varphi(x, S)$ is itself LFP-definable, uniformly over all structures.

Proficiency

- A family of finite structures $\mathcal{C}$ is proficient in case there exists some φ such that $\|\varphi\|_A$ is unbounded in ω as A ranges over $\mathcal{C}$.[1]
- Observation: if $\mathcal{C}$ is not proficient, then $\mathrm{FO} = \mathrm{LFP}$ over $\mathcal{C}$.
- McColm, 1986: conjecture that if $\mathcal{C}$ is proficient, then $\mathrm{FO} \neq \mathrm{LFP}$ over $\mathcal{C}$.
- Slogan: non-proficiency is a finite-variable version of *countable categoricity!*
- Roughly speaking, $\mathcal{C}$ is non-proficient if for each $m \leq n$, it realizes at most finitely many types consisting of of FO^n formulas of arity m.

[1] For any finite structure A, $\|\varphi\|_A < \omega$.

A brief history

- **McColm, 1986:** If $\mathcal{C}$ is proficient, then $\mathrm{FO} \neq \mathrm{LFP}$ over $\mathcal{C}$.
- **Gurevich, Immerman, Shelah, 1994:** Not so. There are proficient families of structures over which $\mathrm{FO} = \mathrm{LFP}$.
- **Kolaitis and Vardi, 1992:** If $\mathcal{C}$ is ordered, then $\mathrm{FO} \neq \mathrm{LFP}$ over $\mathcal{C}$.
- A resolution either way of the ordered conjecture would resolve a major open problem in computational complexity!
  - Positive resolution: LH $\lneq$ ETIME (**Dawar, Hella, 1995**)
  - Negative resolution: PTIME $\lneq$ PSPACE (**Dawar, Lindell, Weinstein, 1995**)

Our contribution

Theorem (BK, 2017)
McColm's conjecture holds for tame classes of finite structures $\mathcal{C}$.

- The proof involves investigating LFP analogues of FO dividing lines like $\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP2}$.
- We also completely classify the implications among the LFP versions of these properties, namely

$$\mathrm{LFP}-\mathrm{SOP} \implies \mathrm{LFP}-\mathrm{TP2} \implies \mathrm{LFP}-\mathrm{IP} \iff \mathrm{LFP}-\mathrm{OP}$$

- Note the suprising equivalence between the order and independence property!

Elementary limit theories

Definition
The elementary limit theory $\mathrm{Th}^\infty(\mathcal{C})$ of $\mathcal{C}$ is the set of sentences which hold in cofinitely many structures of $\mathcal{C}$.

Lemma (Lindell)
The proficiency of $\mathcal{C}$ and whether or not $\mathrm{FO} = \mathrm{LFP}$ over $\mathcal{C}$ are both properties of $\mathrm{Th}^\infty(\mathcal{C})$.

Corollary
McColm's conjecture is a property of $\mathrm{Th}^\infty(\mathcal{C})$.

Elementary limit theories

Lemma
$(\mathbf{lfp}\,\varphi)(t)$ is first-order definable over $\mathcal{C}$ iff there is a first-order formula $\theta(t)$ such that the following sentences are in $\mathrm{Th}^\infty(\mathcal{C})$:

$$\varphi(x,\theta) \leftrightarrow \theta(x) \text{ (1)}$$

$$\psi(x,\neg\theta) \leftrightarrow \neg\theta(x) \text{ (2)}$$

where ψ is complementary to ϕ.

Proof.
($\Longleftarrow$) (1) says that θ is a fixed point of $S \mapsto \varphi(x,S)$, hence $(\mathbf{lfp}\,\varphi)(t) \to \theta(t)$. Similarly, (2) says that $(\mathbf{lfp}\,\psi)(t) \to \theta(t)$.

Since they are complementary, $(\mathbf{lfp}\,\varphi)(t) \leftrightarrow \theta(t)$. □

Tame classes of theories

Let $\varphi(x; y)$ be any formula (FO or LFP), $n \in \mathbb{N}$, and M be a structure.

- φ has an *n-instance of the order property* ($\mathrm{OP}(n)$) in M if there exist tuples $a_1, \ldots, a_n \in M^{|x|}$ and $b_1, \ldots, b_n \in M^{|y|}$ such that $M \models \varphi(a_i; b_j)$ if and only if $i \leq j$.
- φ has an *n-instance of the independence property* ($\mathrm{IP}(n)$) in M if there exist tuples $a_i \in M^{|x|}$ for all $i \in \{1, \ldots, n\}$ and $b_X \in M^{|y|}$ for all $X \subseteq \{1, \ldots, n\}$ such that $M \models \varphi(a_i; b_X)$ if and only if $i \in X$.
- φ has an *n-instance of the strict order property* ($\mathrm{SOP}(n)$) in M if there exist tuples $b_1, \ldots, b_n \in M^{|y|}$ such that $\varphi(M; b_i) \subseteq \varphi(M; b_j)$ if and only if $i \leq j$.

Tame classes of theories

Let $\varphi(x; y)$ be any formula $(\mathrm{FO}$ or $\mathrm{LFP})$, $n \in \mathbb{N}$, and M be a structure.

- $\varphi(x; y)$ has an *n-instance of the tree property of the second kind* $(\mathrm{TP}_2(n))$ in M if there are tuples $b_{i,j} \in M^{|y|}$ for $1 \leq i, j \leq n$ such that for any i and any $j \neq k$, $\varphi(M; b_{i,j}) \cap \varphi(M; b_{i,k}) = \emptyset$, but for any function $f\colon \{1, \dots, n\} \to \{1, \dots, n\}$,

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

Tame classes of theories

- A family $\mathcal{C}$ of finite structures has property (FO- or LFP-) P, for any $P \in \{\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP2}\}$, in case there is a (FO- or LFP-) formula φ with arbitrarily large instances of $P(n)$ in M, as M varies over structures in $\mathcal{C}$.
- Therefore, $\mathcal{C}$ does **not** have P, in case the n-instances of P, for any formula φ, are uniformly bounded as we vary over all $M \in \mathcal{C}$.
- $\mathcal{C}$ having FO-P is equivalent to $\mathrm{Th}^{\infty}(\mathcal{C})$ having P.
- We study these four, because:
  - Most commonly studied combinatorial dividing lines imply either NSOP or $\mathrm{NTP2}$, and
  - IP and OP are otherwise the "most important."

LFP-SOP **and proficiency**

- Suppose a family of structures is proficient. Then the **stage comparison relation** witnesses LFP-SOP.
- Conversely, if $\varphi(\vec{x}, \vec{y})$ witnesses LFP-SOP, then $\psi(\vec{y}_1, \vec{y}_2) \equiv (\forall \vec{x})(\varphi(\vec{x}, \vec{y}_1) \to \varphi(\vec{x}, \vec{y}_2))$ defines a partial order with arbitrarily long chains.
- Given a partial order with arbitrarily long chains, we can define a linear order with arbitrarily long chains (roughly, by **comparing rank**)
- Given a linear order with arbitrarily long chains, we can define a proficient formula.

$\mathrm{LFP\text{-}SOP}$ implies all other properties

- $\mathrm{LFP\text{-}SOP} \implies \mathrm{LFP\text{-}OP}$, under general considerations.
- Consider the family $\mathcal{N}$ of finite initial segments of $(\mathbb{N}, <)$.
- Over $\mathcal{N}$, the BIT and FACTOR predicates are LFP-definable
  - $\mathsf{BIT}(x, y) \iff$ the x-th bit of y base 2 is 1.
  - $\mathsf{FACTOR}(x, y, z) \iff y^z$ is the largest power of y dividing x.
- BIT has IP, FACTOR has TP2
- Let $b_{i,j} = (p_i, j)$, where $(p_i)_{i\in\omega}$ is an enumeration of the primes: for any function $f : n \to n$, let $a_f = \prod_{i<n}^{n} p_i^{f(i)}$. Then,
  - $\mathsf{FACTOR}(\mathbb{N}; p_i, j)$ and $\mathsf{FACTOR}(\mathbb{N}; p_i, k)$ are disjoint, but
  - $a_f \in \bigcap_{i<n} \mathsf{FACTOR}(\mathbb{N}; p_i, f(i))$
- Hence, $\mathrm{LFP} - \mathrm{SOP} \implies \mathrm{LFP} - \mathrm{IP}$ and $\mathrm{LFP} - \mathrm{TP2}$ over *any* $\mathcal{C}$.

First-order characterizations of LFP-P

Lemma
For any $P \in \{\mathrm{OP}, \mathrm{IP}, \mathrm{SOP}, \mathrm{TP2}\}$, $\mathcal{C}$ has LFP-P iff $\mathcal{C}$ is proficient or $\mathcal{C}$ has FO-P.

Corollary
Whether or not $\mathcal{C}$ has LFP-P is a property of $\mathrm{Th}^{\infty}(\mathcal{C})$.

Corollary
For any property P, McColm's conjecture holds for any $\mathcal{C}$ that fails any FO-P.

Proof.
If $\mathcal{C}$ is proficient, then it satisfies LFP-P, but fails FO-P. $\square$

Implications among $\mathrm{LFP}\text{-}P$

- We know $\mathrm{LFP\text{-}SOP} \implies \mathrm{LFP\text{-}TP2} \implies \mathrm{LFP\text{-}IP} \implies \mathrm{LFP\text{-}OP}$. What about conversely?
- There are countably categorical theories with the finite model property that have:
  - IP but $\mathrm{NTP2}$ (e.g., theory of the random graph)
  - $\mathrm{TP2}$ but NSOP (e.g., generic theory of parameterized equivalence relations)

  which show the first two implications are strict.
- If $\mathcal{C}$ has $\mathrm{LFP\text{-}OP}$, but is not proficient, then $\mathrm{FO} = \mathrm{LFP}$, so by $\mathrm{OP} \iff \mathrm{IP} \vee \mathrm{SOP}$ (Shelah), it must have IP.
- Hence, $\mathrm{LFP\text{-}OP} \implies \mathrm{LFP\text{-}IP}$!

Future work

- Investigate combinatorial dividing lines for other fixed-point logics (e.g., transitive closure logic).
- Can we obtain some sort of asymptotic structure theory for tame, non-proficient classes of structures?
- To what extent can **FMT assumptions** of bounded cliquewidth and bounded treewidth be assimilated into model-theoretic tameness considerations?
- **Chen and Flum, 2012:** The ordered conjecture is **true** for families of finite structures of bounded cliquewidth and treewidth.
- **Open question:** Does the ordered conjecture hold for any tame family of finite structures?

Acknowledgements

- Thanks to Cameron Hill, Greg McColm, Steve Lindell, and Scott Weinstein for their help and encouragement.
- Thanks to Larry Moss and the logic group at Indiana University for their support.
- Thanks for listening! If you want to talk more about this, please contact us by email.