Using Non-Standard Techniques to Analyze First-Order Definability over Finite Structures

by

Steven Lindell
Haverford College
April 18th, 1999

Scott Weinstein
University of Pennsylvania
Institute for Research in Cognitive Science

Abstract

We present a positive solution to McColm's conjecture for the special case of *decidable theories*. This gives a uniform explanation of the separation of fixed-point from first-order definability over a wide range of commonly studied classes of finite structures (including ordered ones). We demonstrate why this result generalizes the standard constructive proofs, and show why it is unlikely that these techniques will ever be able to perform any separation of complexity-theoretic significance!

Techniques

The rest of the talk explores examples of the use of infinitary techniques to effectively produce known separations; all without the use of any tedious combinatorics. We hope these will provide inspiration for the application of non-constructive methods to achieve lower-bound arguments.

First-order Definability (FO)

$$\mathfrak{A} = \langle \overbrace{\underset{\text{finite domain}}{A}, \underbrace{R_1, \dots, R_k}_{\text{relations}}}^{\text{input}}, \overset{\text{query}}{Q} \rangle \qquad R_i^{\mathfrak{A}} \subseteq A^{a_i}; A = |\mathfrak{A}| \qquad Q^{\mathfrak{A}} \subseteq A^{a} \text{ (output)}$$

Global augmentation of a class via languages:

first-order formulas

variables over A	$x = y$	$\wedge$ $\vee$	$\exists x$
e.g. x, y, z	$R(\bar{x})$	$\neg$	$\forall x$

$\varphi(\bar{x}) \in FO \quad \varphi^{\mathfrak{A}} = \{\bar{a} \in A^{|\bar{x}|} : \mathfrak{A} \vDash \varphi[\bar{a}]\}$

Fixed-point Definability (FP)

inductions:

new S ; arity$(S) = \bar{x}$; $\varphi(\bar{x}, S)$ is S-pos.

$$\emptyset = \varphi^0 \subseteq \cdots \subseteq \varphi^i \subseteq \varphi(\varphi^i) = \varphi^{i+1} \subseteq \cdots \subseteq \varphi^\infty$$

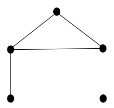
$S_0 \subseteq |\mathfrak{A}|^{|\bar{x}|} \quad \varphi^{\mathfrak{A}}(S_0) = \{\bar{a} : (\mathfrak{A}, S_0) \vDash \varphi[\bar{a}]\}$

monotone: $S_1 \subseteq S_2 \Rightarrow \varphi(S_1) \subseteq \varphi(S_2)$

Simple Graphs

$G = \langle V, E \rangle$

$E \subseteq V^2$



Simplicity is FO (elementary)

$\neg xEx$

$xEy \to yEx$

Connectivity is in FP

Start by defining *paths* in graphs

Transitive Closure:

$$\varphi(x, y, E^{+}) \equiv xEy \wedge (\exists z) xEzE^{+}y$$

= diameter

$$\varphi^{\infty}(x, y) \Longleftrightarrow \text{there is a path from } x \text{ to } y$$

G is connected $\Longleftrightarrow \forall x \forall y \varphi^{\infty}(x, y)$

Binary Strings

$\langle B, <, U \rangle$ | ○ | ○

$< \subseteq B^2; U \subseteq B$ ● < ○ < ● < ○

1 2 3 4

Orderedness is elementarily definable

$$x \nless x$$
$$x \neq y \rightarrow x < y \vee y < x$$
$$x < y \wedge y < z \rightarrow x < z$$

Define Parity in FP

$$\left.\begin{matrix} O(x) \\ E(x) \end{matrix}\right\} \Longleftrightarrow |\{y \leq x : U(y)\}| \text{ is } \begin{cases} \text{odd} \\ \text{even} \end{cases}$$

$$\varphi_O(x, O, E) \equiv x = \min \wedge U(x) \vee E(\text{pred}(x)) \wedge U(x) \vee$$

φ_E similarly

$$O(\text{pred}(x)) \wedge \neg U(x)$$

Define in FO: $\text{pred}(x); x \neq \min$

compute simultaneous fixed-points:

$$\left.\begin{matrix} \varphi_O^{\infty}(\max) \\ \varphi_E^{\infty}(\max) \end{matrix}\right\} \Longleftrightarrow \text{parity of entire string is } \begin{cases} \text{Odd} \\ \text{Even} \end{cases}$$

Descriptive Complexity

Fixed-Point characterization of polynomial-time over classes of ordered structures:

Ptime = FP(<, …) [Immerman] [Vardi]

A First-Order characterization of $O(1)$ parallel-time (uniform – AC^0) requires arithmetic:

CRAM($O(1)$) = FO(<, bit,…) [Immerman]

= FO(+, ∗, …) = FO(bit, …) [DDLW]

Separation of FO and FP

Observe $\left\langle \underbrace{2^{\cdot^{\cdot^{\cdot^{2}}}}}_{n}, \text{bit} \right\rangle \cong \langle V_n, \in \rangle$

Fact: FO ≠ FP over $\{\langle V_n, \in \rangle : n \in w\}$
iff HTIME($O(n)$) ≠ DTIME($2^{O(n)}$).

Goal: Explain why separation is so difficult to prove (apparently).

McColm's Conjecture

Definition: A class $\mathcal{C}$ of finite structures is *proficient* if there is an unbounded induction over $\mathcal{C}$, i.e.:

$$\sup\{\|\varphi\|_{\mathfrak{A}} : \mathfrak{A} \in \mathcal{C}\} = \infty,$$

where $\|\varphi\|$ is the closure ordinal of φ

Observation: $\mathcal{C}$ not proficient ⇒ FP = FO

Conjecture: $\mathcal{C}$ proficient ⇒ FP ≠ FO

FALSE [GIS] counterexample

Important Open Subcases

Conjecture: FP ≠ FO on every class of ordered structures [KV]
positive solution implies complexity separations

Instead, consider the subcase

$$\mathrm{Th}(\mathcal{C}) = \{\theta \in \mathrm{FO} : \mathcal{C} \vDash \theta\} \text{ is decidable}$$

Theorem: If $\mathcal{C}$ is also proficient, then FP ≠ FO on $\mathcal{C}$.

Decidable Theories

Definition: $\mathrm{Th}(\mathcal{C}) = \{\theta : \forall \mathfrak{A} \in \mathcal{C}, \mathfrak{U} \vDash \theta\}$
"Theory of $\mathcal{C}$", all *valid* sentences over $\mathcal{C}$

$\mathrm{Th}^{\infty}(\mathcal{C}) = \{\theta : |\{\mathfrak{A} \in \mathcal{C} : \mathfrak{U} \nvDash \theta\}| < \infty\}$
"Limit Theory", sentences *eventually* true on $\mathcal{C}$

Observation: $\mathrm{Th}^{\infty}(\mathcal{C})$ is complete iff every FO sentence is eventually true or eventually false

Results

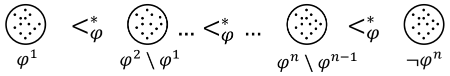
Theorem [McColm]: If $\mathcal{C}$ is r.e. and proficient then $\mathrm{Th}^{\infty}(\mathcal{C})$ complete $\Rightarrow$ FP ≠ FO on $\mathcal{C}$.

Lemma: If $\mathcal{C}$ is proficient and $\mathrm{Th}(\mathcal{C})$ is decidable, then there exists r.e. proficient $\mathcal{D} \subseteq \mathcal{C}$ such that $\mathrm{Th}^{\infty}(\mathcal{D})$ is complete.

Theorem: If $\mathcal{C}$ is proficient and $\mathrm{Th}(\mathcal{C})$ decidable then FP ≠ FO on $\mathcal{C}$.

Proof of Theorem

Suppose $n = \|\varphi\|_{\mathfrak{A}} \longrightarrow \infty$ for $A \in \mathcal{C}$
Consider equivalence classes determined by the stages of $\varphi(\bar{x}, S)$:



Suppose this (stage comparison) along with inductive definitions for + and × are all elementary over $\mathcal{C}$.

Polynomials

Express any Diophantine equation:

$$p(\bar{x}) = c_d\bar{x}^d + \cdots + c_0\bar{x}^0 = 0 \quad c_i \in \mathbb{N}$$

and see that it has a solution iff

$$\theta \equiv (\exists \bar{x})p(\bar{x}) = 0 \in \mathrm{Sat}(\mathcal{C})$$

$$\Leftrightarrow \neg\theta \notin \mathrm{Val}(\mathcal{C}) = \mathrm{Th}(\mathcal{C}) \text{ which is decidable}$$

A contradiction to the unsolvability of Hilbert's 10th.

Corollary

Corollary: A proficient subclass $\mathcal{C}' \subseteq \mathcal{C}$ with $\mathrm{Th}(\mathcal{C}')$ decidable, $\Rightarrow$ FP ≠ FO on $\mathcal{C}$.

Proof: Separations always extend "upward".

Applications

Gives uniform explanation of separation from FO on a wide variety of examples:

- Strings (colored orderings)
- Trees
- Graphs of bounded tree-width

In fact, their mSO theory is decidable,
⟹ FP ≠ mSO on above classes.

Extensions

Presburger ⟹ $\mathrm{Th}(\{\langle n,+\rangle : n\in\mathbb{N}\})$ decidable.

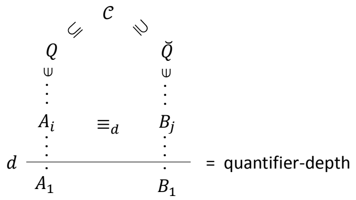
Moreover, consider FO with monadic counting on ordered structures: $\exists_x^{\geq i}\varphi(x)$

Fact: mFOC(<) = FO(+)

Therefore, FP ≠ mFOC (in fact, DTC(<) ≠ mFOC)

Separation Arguments

Theorem: $Q \notin \mathrm{FO}$ on $\mathcal{C}$ iff



Proof: Perform refinement on complete *d*-types that *Q* splits into ∞ pieces on $\mathcal{C}$.

Observations

Let $\mathcal{C}' = \{A_i, B_i : i = 1, 2, \ldots\}$

1. The subclass $\mathcal{C}'$ trivializes FO, i.e. every elementary query is finite/co-finite.

Typically (E-F game), we have that $\mathcal{C}'$ is constructive, i.e. $d \longmapsto A_d, B_d$ effective.

2. In this case, $\mathrm{Th}(\mathcal{C}')$ is decidable!

This implies "effective trivialization" of FO

Logtime ≠ Trivial on r.e. classes

Theorem: Let $\mathcal{C}$ be an infinite r.e. class. Then there is a non-trivial Logtime Boolean query on $\mathcal{C}$.

Proof: Let *M* be a machine enumerating $\mathcal{C} = \{A_1, A_2, \ldots\}$. For A ∈ $\mathcal{C}$, define $q(A)$ to be the computation:

Computation

1. Calculate $|A|$ in binary.
2. Run *M* for log $|A|$ steps to obtain *B* = last structure generated (∅ if none)
3. Set $q = \begin{cases} \text{false if } B = \emptyset \\ \neg q(B) \text{ otherwise} \end{cases}$

Observe that $q \in O(\log|A|)$ time (using a random-access Turing machine). Notice that construction guarantees an alternating subsequence of truth values for q over $\mathcal{C}_\square$

Implications for Complexity Theory

Using probabilistic techniques [FSS]

$$\text{Parity} \in \text{Ptime} \setminus \text{CRAM}(O(1))$$

Why not use descriptive complexity as a logical approach to separations??

$$Q \in \text{FP(bit)} \setminus \text{FO(bit)}$$

Using ‘monotone’ pebble game induces:

$$Q \upharpoonright_{\mathcal{C}'} \in \text{FP} \setminus \text{FO} = \text{trivial on } \mathcal{C}'$$

Pebble Games

In the presence of arithmetic:

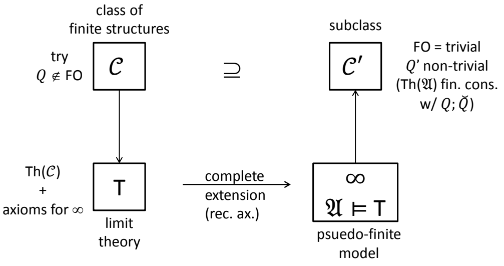
FO = Logtime – H $\supseteq$ DLOGTIME [BIS].

So, $\mathcal{C}'$ trivializes DLOGTIME

Hence, $\mathcal{C}'$ is not recursively enumerable

Therefore, monotone E-F pebble game technique is not **effective**!

Infinitary Techniques



Non-constructive?

$\mathcal{C}'$: try $Q \notin$ FO = trivial

1. Can avoid tedious combinatorial arguments
2. Allows potential for non-constructive techniques

T = Con(Th($\mathcal{C}$), axioms for ∞)

Background

Recall: The *limit theory* of $\mathcal{C}$ is $\text{Th}^{\infty}(\mathcal{C})$ = Cons(Th($\mathcal{C}$), axioms of infinity)

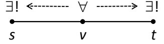
Call $A \models \text{Th}^{\infty}(\mathcal{C})$ pseudo-finite and note that this is equivalent to A having the FMP over $\mathcal{C}$

Definition: An infinite model A has the *finite model property* if whenever $A \models \theta$, there is a (finite) model which satisfies θ. (in $\mathcal{C}$)

Connectivity $\notin$ FO

Consider (fin. ax) theory of simple 2-regular graphs:

$\tau \equiv$

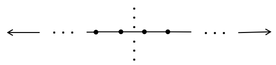


whose finite models are collections of cycles.

Let σ_n say there are no cycles of length *n*,

Claim: T = $\{\tau\} \cup \{\sigma_1, \sigma_2, \ldots\}$ is complete.

Proof: T is uncountably categorical, since every model is a collection of chains:

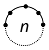


apply Łoś-Vaught test □

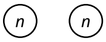
Compactness!

Observe: T is finitely consistent with

{connected graphs}
(finite)



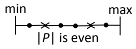
{disconnected graphs}
(finite)



Proposition: Connectivity is not FO-definable.

Proof: Apply compactness to both sides. QED

Sparse PARITY $\notin$ FO(<)



$\mathcal{C}$ = finite $\langle A, <, P\rangle$
$\Sigma = \mathrm{Th}(\mathcal{C}) \cup$
$\{P(\min), P(\max), |P| = \infty, P \text{ "sparse"}\}$
(adjacent elements of P are infinitely far apart)

Obvious: Σ is finitely consistent with both $\mathcal{C} \upharpoonright_{\text{even}}$ and $\mathcal{C} \upharpoonright_{\text{odd}}$.

Lemma: Σ is complete (this implies the result).

Proof of Lemma

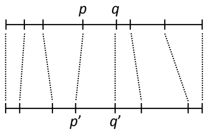
Let $\langle A, <, P\rangle \vDash \Sigma$ be $\aleph_1$- saturated.
Let Δ = theory of infinite discrete < (w/ end pts).

$\langle P, <\rangle \vDash \Delta$ is $\aleph_1$- saturated (relativize types).
For $p, q \in P$ adjacent, $\langle [p,q], <\rangle \vDash \Delta$ is also.

Assume CH: We aim to show that any two saturated models of power $2^{\aleph_0}$ are isomorphic.

Isomorphism

$\mathfrak{A} = \langle A, <, P\rangle$



$\mathfrak{A}' = \langle A', <, P'\rangle$

$\langle P, <\rangle \vDash \Delta$
$\cong$ complete
$\langle P', <\rangle \vDash \Delta'$

To extend $\cong$ to all of A and A', notice
$[\forall x \notin P \;\; \exists \text{ adj. } y, z \in P \;\; x \in (y,z)] \in \mathrm{Th}(\mathcal{C})$

Now combine with $\langle [p,q], <\rangle \cong \langle [p',q'], <\rangle$.