McColm's Conjecture

Yuri Gurevich* Neil Immerman† Saharon Shelah‡

U of Michigan U of Massachusetts Hebrew U and Rutgers

LICS (1994), 10–19

Abstract

Gregory McColm conjectured that positive elementary inductions are bounded in a class K of finite structures if every (FO + LFP) *formula is equivalent to a first-order formula in K. Here* (FO + LFP) *is the extension of first-order logic with the least fixed point operator. We disprove the conjecture. Our main results are two model-theoretic constructions, one deterministic and the other randomized, each of which refutes McColm's conjecture.*

1 Introduction

Gregory McColm conjectured in [M] that, for every class K of finite structures, the following three claims are equivalent:

M1 Every positive elementary induction is bounded in K.

M2 Every (FO + LFP) formula is equivalent to a first-order formula in K.

M3 Every $L^{\omega}_{\infty\omega}$-formula is equivalent to a first-order formula in K.

The definitions of $L^{\omega}_{\infty,\omega}$ and (FO+LFP) are recalled in the next section.

Clearly, M1 implies M2. McColm observed that M3 implies M1. Phokion Kolaitis and Moshe Vardi proved that M1 implies M3 [KV]. A nice exposition of all of that is found in [D] The question whether M2 implies M1 has been open though McColm made the following important observation.

Let $\bar{n}$ be the set $\{0, 1, .., n-1\}$ with the standard order. It is easy to see that no infinite class of structures $\bar{n}$ satisfies M1. List all (FO + LFP) sentences in vocabulary $\{<\}$: $\varphi_0, \varphi_1, \ldots$. Let $K_i = \{\bar{n} \mid \bar{n} \models \varphi_i\}$ and construct an infinite K such that every intersection $K \cap K_i$ is either finite or co-finite. Each φ_i is equivalent to a first-order sentence in K. Thus M1 does not follow from the restriction of M2 to formulas without free variables.

The main results of this paper are two model theoretic constructions, one deterministic and the other randomized, each of which gives a counterexample to the implication M2→M1. Actually, each construction implies the stronger result that M2 fails to imply M1 even when (FO + LFP) is replaced in M2 by an arbitrary countable subset of $L^{\omega}_{\infty,\omega}$, see Corollary 3.10 and Theorem 4.1. We present the deterministic construction in full detail in Section 3. The randomized construction is presented in Section 4; but, some of the proofs are omitted due to lack of space.

Both constructions depend on the fact that the language $L^{\omega}_{\infty,\omega}$, and thus (FO + LFP) is unable to count the number of vertices in a large clique. The deterministic construction extends naturally to Theorem 3.13: an extension of our counterexample to the stronger language (FO + LFP + COUNT) in which counting is present.

Recall that (FO + ITER), is first-order logic plus an unbounded iteration operator (equivalent to the "while", and "partial fixed point" operators). It is known that the language (FO + ITER) captures PSPACE on ordered structures [I82, V]. Abiteboul and Vianu [AV] showed that P = PSPACE if and only if, (FO + LFP) = (FO + ITER) on all sets of finite structures.

*Partially supported by NSF, ONR and BSF. Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI 48109-2122, USA, gurevich@umich.edu

†Supported by NSF grant CCR-9207797. Computer Science Department, University of Massachusetts, Amherst, MA 01003, USA, immerman@cs.umass.edu

‡Publication 525. Partially supported by USA–Israel Binational Science Foundation. Mathematics Department, Hebrew University, Jerusalem 91904, Israel, shelah@cs.huji.ac.il, and Mathematics Department, Rutgers University, New Brunswick, NJ 08903, USA, shelah@math.rutgers.edu

In light of this, another interesting consequence of the deterministic construction is Corollary 3.14 which says that if P is not equal to PSPACE, then there is a set of finite structures on which FO = (FO + LFP), but on which FO ≠ (FO + ITER).

2 Background

We briefly recall some background material. More information on Descriptive Complexity and Finite Model Theory can be found for example in [I89] and [G].

Proviso Structures are finite. Vocabularies are finite and do not contain function symbols of positive arity. In particular, the vocabulary of any $L^{\omega}_{\infty,\omega}$-formula is finite. Classes of structures are closed under isomorphism. □

If M is a structure then $|M|$ is the universe of M. If X is a nonempty subset of M (that is, of $|M|$) then $M \mid X$ is the induced substructure with universe X.

An *r-ary global relation* ρ on a class K of structures of the same vocabulary is a function that, given a structure $M \in K$, produces an r-ary (local) relation ρ^M on $|M|$. By definition, $M \models \rho(\bar{a})$ if and only if $\bar{a} \in \rho^M$. It is supposed that, for every isomorphism η from M to a structure N and every r-tuple $x_1, \ldots, x_r$ of elements of M, $M \models \rho(x_1, \ldots, x_r) \iff N \models \rho(\eta(x_1, \ldots, x_r))$.

In this paper, an *infinitary formula* means an $L^{\omega}_{\infty,\omega}$ formula of finite vocabulary. Recall that $L^{\omega}_{\infty,\omega}$ is the generalization of first-order logic that allows arbitrary infinite conjunctions and disjunctions provided that the total number of individual variables, bound or free, in the resulting formula is finite [B]. In other words, infinitary formulas are built from atomic formulas by means of negation, existential quantification, universal quantification and the following rule:

- If $\{\varphi_i \mid i \in I\}$ is a collection of infinitary formulas that uses only a finite vocabulary and a finite number of individual variables then $\bigvee_i \varphi_i$ and $\bigwedge_i \varphi_i$ are infinitary formulas.

The semantics is obvious. $A \models \bigvee_i \varphi_i(\bar{a})$ if and only if $A \models \varphi_i(\bar{a})$ for some i, and $A \models \bigwedge_i \varphi_i(\bar{a})$ if and only if $A \models \varphi_i(\bar{a})$ for all i. Let $L^{k}_{\infty,\omega}$ be the subset of $L^{\omega}_{\infty,\omega}$ in which at most the k distinct variables $\{x_1, x_2, \ldots, x_k\}$ occur.

We next recall the definition of (FO + LFP). Consider a first-order formula $\varphi(P, v_1, \ldots, v_r, v_{r+1}, \ldots, v_s)$ with free individual variables $v_1, \ldots, v_s$ where an r-ary predicate P has only positive occurrences; let $\tau = \text{Vocabulary}(\varphi) - \{P\}$. Given a τ-structure M and elements $a_{r+1}, \ldots, a_s$ of M, we have the following r-ary relations on the universe $|M|$ of M:

$$P_0 = \emptyset, \qquad P_{i+1} =$$

$$\{(v_1, \ldots, v_r) \mid M \models \varphi(P_i, v_1, \ldots, v_r, a_{r+1}, \ldots, a_s)\}$$

Since P is positive in φ, $P_0 \subseteq P_1 \subseteq P_2 \subseteq \ldots$. M1 asserts that, for every such φ, there exists a positive integer j such that, for every $M \in K$ and any $a_{r+1}, \ldots, a_s \in M$, $P_j = \bigcup_i P_i$.

The least fixed point operator LFP can be applied to the formula φ. The result is a new formula

$$\text{LFP}_{P;v_1,\ldots,v_r}\varphi(v_1, \ldots, v_s)$$

of vocabulary τ. If M is a τ-structure, $a_1, \ldots, a_s$ are elements of M and relations P_i are as above then

$$M \models \text{LFP}_{P;v_1,\ldots,v_r}\varphi(a_1, \ldots, a_s) \Leftrightarrow (a_1, \ldots, a_r) \in \bigcup_i P_i.$$

(FO + LFP) is the extension of first-order logic with this new formula-constructor. Applications of LFP can be nested and interleaved with other formula-constructors. It is obvious that (FO+LFP) is a subset of $L^{\omega}_{\infty,\omega}$.

Pebble games are a convenient tool to deal with infinitary formulas. A k-pebble game $\Gamma^k_\tau(A, B)$ is played by Spoiler and Duplicator on structures A, B of vocabulary τ. For each $i \in \{1, \ldots, k\}$, there are two pebbles numbered i; there are $2k$ pebbles altogether. Starting with Spoiler, the players alternate making moves. A move consists of placing a free pebble at an element of one of the two structures or removing one of the pebbles from some element. If Spoiler puts a pebble of number i at an element x of A (resp., an element y of B), Duplicator must answer by placing the other pebble number i at some element y of B (resp., some element x of A). If Spoiler removes a pebble number i, Duplicator must remove the other pebble number i. Initially, all pebbles are free. At each even-numbered state S, the pebbles define a partial map η_S from A to B. Dom(η_S) consists of the elements of A covered by pebbles. If $x \in A$ is covered by a pebble i then $\eta_S(x)$ is the element of B covered by the other pebble i. Initially, all $2k$ pebbles are free. The goal of Duplicator is to ensure that every such η_S is a partial isomorphism. If the game reaches an even state S such that η_S is not a partial isomorphism, Spoiler wins; otherwise the game continues forever and Duplicator wins.

Fact 2.1 ([B, I82]) *Let $l \leq k$ and consider the version of Γ^k_τ where the initial state is as follows: pebbles $1, \ldots, l$ are placed at elements $x_1, \ldots, x_l$ of A and at elements $y_1, \ldots, y_l$ of B. If Duplicator has a winning strategy in that game then, for every τ-formula $\varphi(v_1, \ldots, v_l) \in L^k_{\infty,\omega}$,*

$$A \models \varphi(x_1, \ldots, x_l) \iff B \models \varphi(y_1, \ldots, y_l).$$

3 The Deterministic Construction

We are now ready to state our main theorem:

Theorem 3.1 *There exists a set of finite directed graphs, $\mathcal{G} = \{G_1, G_2, \ldots\}$, such that $\mathcal{G}$ admits fixed points of unbounded depth and yet on $\mathcal{G}$,* FO = (FO + LFP)*, i.e. every formula expressible with a least fixed point operator is already first-order expressible.*

The proof of Theorem 3.1 has two main ideas. The first is the idea of a standard oracle construction from Structural Complexity Theory. The second is Lemma 3.5: a formula in (FO + LFP) with only k distinct variables cannot distinguish a k-clique from any larger clique. We divide the proof up into several parts, that of the oracle construction (Section 3.1), that with one free variable (Section 3.2), and finally the general case (Section 3.3).

3.1 With Lots of Relation Symbols

In this subsection we concentrate on the oracle construction by temporarily introducing infinitely many new relation symbols of each arity: R^j_i, $i, j \geq 1$. For convenience in the proofs we will use the notation $\mathrm{var}(\varphi)$ to denote the number of distinct variables free or bound occurring in φ. Let $\mathrm{free}(\varphi)$ denote the number of free variables occurring in φ.

Lemma 3.2 *There exists a set of finite directed graphs, $\mathcal{D} = \{D_1, D_2, \ldots\}$, which also interpret the new relations: R^j_i, $i, j \geq 1$, such that $\mathcal{D}$ admits fixed points of unbounded depth; and yet on $\mathcal{D}$,* FO = (FO+LFP)*, i.e., every formula expressible with a least fixed point operator is already first-order expressible.*

proof Let $\Delta_1, \Delta_2, \ldots$ be a listing of all formulas in (FO + LFP) in this expanded language. Let $u_i = \mathrm{free}(\Delta_i)$, the number of free variables occurring in Δ_i. Let S_i be one of the new relation symbols of arity u_i such that,

$$S_i \text{ does not occur in } \Delta_r \text{ for any } r \leq i. \tag{3.3}$$

We will let the graph $D^0_j = \langle V_j, E_j \rangle$ be a directed segment of length $j - 1$:

$$V_j = \{d_1, d_2, \ldots, d_j\}$$
$$E_j = \{\langle d_k, d_{k+1} \rangle \mid 0 \leq k < j\}$$

We next show how to interpret the new relation symbols in the D_j's such that: For all i, for all $j \geq i$, and for all $a_1, a_2, \ldots, a_{u_i} \in |D_j|$,

$$D_j \models (\Delta_i(a_1, a_2, \ldots, a_{u_i}) \leftrightarrow S_i(a_1, a_2, \ldots, a_{u_i})) \tag{3.4}$$

From Equation 3.4, it follows that each Δ_i is equivalent to a first-order formula – in fact, to an atomic formula – for all but finitely many structures. Of course, on any fixed finite structure, the formula Δ_i is equivalent to a first-order formula. Lemma 3.2 follows immediately.

Now we construct the D_j's so that Equation 3.4 holds. D^0_j defined above is just a graph, which may be thought of as interpreting all of the new relations as false. Assuming D^{i-1}_j has been defined, let D^i_j be the same as D^{i-1}_j except that for all $a_1, a_2, \ldots, a_{u_i} \in |D_j|$, we interpret S_i so that

$$D_j \models (\Delta_i(a_1, a_2, \ldots, a_{u_i}) \leftrightarrow S_i(a_1, a_2, \ldots, a_{u_i}))$$

Note that by Equation 3.3, this doesn't affect any of the previous steps.

Let $D_j = D^j_j$. This completes the construction, guaranteeing that Equation 3.4 holds. This completes the proof of Lemma 3.2. □

3.2 One Free Variable Case: Relations Replaced by Cliques

Now, we get rid of the new relation symbols, replacing them by cliques attached to the vertices in the D_j's. The main result we will need is that formulas from $L^k_{\infty,\omega}$, i.e. infinitary formulas with at most k variables, cannot distinguish k-cliques from r-cliques for any $r > k$.

Lemma 3.5 *Let F be a finite, directed graph and let v be a vertex in F. For $i \geq 1$, let F_i be the result of replacing v by a clique of i new vertices: $v_1, \ldots, v_i$. Each edge $\langle v, w \rangle$ or $\langle z, v \rangle$ from F is replaced with i new edges: $\langle v_j, w \rangle$ or $\langle z, v_j \rangle$, $j = 1, 2, \ldots, i$. Let $1 \leq k < r$ be natural numbers. Then F_k and F_r agree on all formulas with at most k variables from $L^\omega_{\infty,\omega}$.*

proof This is proved by using the game Γ^k_τ from Fact 2.1. We have to show that the Duplicator has a winning strategy for the k-pebble game on F_k and

F_r. Her strategy is to answer any move outside of the cliques with the same vertex in the other graph. A move on one of the new cliques is likewise matched by a move on the new clique in the other graph. Since there are only k pebbles, there is always an unpebbled vertex in either of the cliques to match with. Thus the Duplicator has a winning strategy. It follows that F_k and F_r agree on all formulas from $L^k_{\infty,\omega}$. □

To make the deterministic construction easier to understand we begin by doing it just for formulas with only one free variable:

Lemma 3.6 *There exists a set of finite directed graphs, $\mathcal{H} = \{H_1, H_2, \ldots\}$, such that $\mathcal{H}$ admits fixed points of unbounded depth, and yet on $\mathcal{H}$, every formula with at most one free variable that is expressible with a least fixed point operator is already first-order expressible.*

proof Let $\Theta_1, \Theta_2, \ldots$ be the set of all formulas in $(\mathrm{FO} + \mathrm{LFP})$ that have at most one free variable. The construction of the H_j's is similar to that of the D_j's of Lemma 3.2. The difference is that instead of making the relation $S_i(d)$ hold, we will modify the size of a certain clique that is connected to d.

We next define the sequence of natural numbers: $v_0 < v_1 < v_2 < \cdots$ that will be the sizes of the initial cliques. Let $v_0 = 0$, and inductively, let $v_i = \max(\mathrm{var}(\Theta_i), v_{i-1} + 2^{i+1})$. In the construction of H_j we will modify the sizes of cliques that are initially of size v_i. The modification will add a number of vertices to these cliques while keeping them smaller than v_{i+1}.

Define the graph H^0_j as follows: First, H^0_j contains D^0_j, the directed segment of length $j - 1$. For each $d \in |D^0_j|$ and for each $i \leq j$, H^0_j also contains the size v_i clique $C_{d,i}$ which has edges from each of its elements to the vertex d.

Assuming H^{i-1}_j has been defined, let H^i_j be the same as H^{i-1}_j except that for each $d \in |D^0_j|$ we add $n(d,i)$ vertices to the v_i-vertex clique $C_{d,i}$. The number $n(d,i)$ is an $i+1$ bit binary number such that:

$$(\text{“Bit 0 of } n(d,i) \text{ is one.”}) \quad \Leftrightarrow \quad (H^{i-1}_j \models \Theta_i(d))$$

And, for $1 \leq s \leq i$, let a_s be a vertex in $C_{d,s}$. Then,

$$(\text{“Bit s of } n(d,i) \text{ is one.”}) \quad \Leftrightarrow \quad (H^{i-1}_j \models \Theta_i(a_s))$$

Finally, let $H_j = H^j_j$. Define the notation $S \preceq_k T$ to mean that S is a k-variable elementary substructure of T. That is, S is a substructure of T and for all first-order formulas φ with $\mathrm{var}(\varphi) \leq k$, and for all $a_1, a_2, \ldots, a_k \in |S|$,

$$S \models \varphi(a_1, a_2, \ldots, a_k) \quad \Leftrightarrow \quad T \models \varphi(a_1, a_2, \ldots, a_k)$$

We have constructed the H_j's so that,

$$H^{i-1}_j \preceq_{v_i} H_j \tag{3.7}$$

Equation 3.7 follows from Lemma 3.5 and the fact the the construction of H^r_j for $r \geq i$ proceeds by increasing the size of cliques whose size is at least v_i.

Let $a \in |H_j|$. If $a = d \in |D^0_j|$ then,

$$\begin{aligned}(H_j \models \Theta_i(a)) \quad &\Leftrightarrow \quad (H^{i-1}_j \models \Theta_i(d)) \\ &\Leftrightarrow \quad (\text{“Bit 0 of } n(d,i) \text{ is one.”})\end{aligned}$$

If a is a member of a clique $C_{d,r}$, let $s = \min(i,r)$. Then,

$$\begin{aligned}(H_j \models \Theta_i(a)) \quad &\Leftrightarrow \quad (H^{i-1}_j \models \Theta_i(a)) \\ &\Leftrightarrow \quad (\text{“Bit s of } n(d,i) \text{ is one.”})\end{aligned}$$

Remember that v_{i+1} is a fixed constant. Furthermore, there are at most 2^{i+1} possible values for $n(d,i)$. It follows that there is a first-order formula $\varphi_i(a)$ that finds the appropriate d and s, and determines $n(d,i)$ which is the size of largest maximal clique connected to d that has fewer than v_{i+1} vertices. Next, compute bit s of $n(d,i)$ by table look up, and let $\varphi_i(a)$ be true iff this bit is one.

Thus, we have that for all $j \geq i$ and for all $a \in |H_j|$,

$$H_j \models (\Theta_i(a) \leftrightarrow \varphi_i(a))$$ □

3.3 General Case: Arbitrary Arity

The reason that the general case is more complicated than the arity one case is that we must include gadgets that identify tuples of nodes. We then must contend with having arguments from these gadgets and so the arities seem to multiply. We must therefore be careful so that the arities remain bounded.

proof of Theorem 3.1: Let $\Gamma_1, \Gamma_2, \ldots$ be a listing of all formulas in $(\mathrm{FO} + \mathrm{LFP})$. As we have mentioned, arities might multiply. The base arity of the formula Γ_i is $f_i = \mathrm{free}(\Gamma_i)$. We will use increased arities $A_0 < A_1 < \ldots < A_j$ defined by $A_0 = 1$, and inductively,

$$A_i = 1 + (A_{i-1})(2f_i) \tag{3.8}$$

Next define the sequence of natural numbers: $w_0 < w_1 < w_2 < \cdots$ that will be the sizes of the initial cliques. Let $w_0 = 0$, and inductively, let $w_i = \max(\mathrm{var}(\Gamma_i), 1 + w_{i-1} + A_{i-1})$.

To define the graph G_j, we begin as usual by including the directed segment D_j^0. For each i, we include enough gadgets: T_i^r, $r = 1, 2, \ldots, n_i$, to encode all possible sequences of length at most A_i of elements of $|D_j^0|$. (Here, n_i is equal to $(j+1)^{A_i}$.)

Each gadget T_i^r consists of $j \cdot A_i$ cliques of size w_i. For each $d \in |D_j^0|$ there are A_i of these cliques, $C_{d,i}^r$, with edges to d. T_i^r also contains one vertex t_i^r with edges to all the $C_{d,i}^r$'s, $d = 1, \ldots, j$. When we want T_i^r to encode the sequence $d_1, d_2, \ldots, d_{A_i}$ we will choose A_i cliques, $C_{d_1,i}^r, C_{d_2,i}^r, \ldots, C_{d_{A_i},i}^r$ and increase their sizes by $1, 2, \ldots, A_i$ vertices respectively. Note that we have enough copies of each $C_{d,i}^r$ to tolerate any number of repetitions of the same d. To skip one of the members of the sequence, say d_t, we increase no clique by exactly t vertices. In this case we write $d_t = 0$. Thus, we have shown how to modify the gadget T_i^r so that it codes any sequence of length A_i from the alphabet $\{0, 1, \ldots, j\}$. Note that no formula Γ_t with $t \le i$ can detect this modification!

Define G_j^0 to include D_j^0 plus all of the T_i^r's, $1 \le i \le j$, $1 \le r \le n_i$.

Inductively, assume that G_j^{i-1} has been constructed. Now, for each tuple $a_1, a_2, \ldots, a_{f_i} \in |G_j^{i-1}|$, **if** $G_j^{i-1} \models \Gamma_i(a_1, a_2, \ldots, a_{f_i})$, **then** we will modify one of the gadgets T_i^r to encode the tuple, $a_1, a_2, \ldots, a_{f_i}$.

Let's first consider the case that a_1 is a vertex from some $T_{i-1}^{r_1}$. In this case, $T_{i-1}^{r_1}$ codes a sequence,

$$b_{11}, b_{12}, \ldots, b_{1,A_{i-1}}, \quad \text{each } b_{1t} \in \{0, 1, \ldots, j\} \qquad (3.9)$$

To reencode this sequence, we first just copy it. Next, we have to indicate which vertex in $T_{i-1}^{r_1}$, a_1 is. (It could be the vertex $t_{i-1}^{r_1}$, or a vertex in one of the unused cliques, $C_{d,i-1}^{r_1}$, or in one of the cliques $C_{b_{1q},i-1}^{r_1}$ that codes the q^{th} element of the sequence of Equation 3.9. In each case, we use the A_{i-1} extra slots to encode which of these cases apply[1]. This is the reason for the factor of 2 in Equation 3.8 and while this is slightly wasteful, it is simple and we are just trying to prove that something is finite.

We have just explained how to encode a_1 in the first $2A_{i-1}$ slots of T_i^r. Similarly, code $a_2, \ldots, a_{f_i}$ into the next $2A_{i-1}(f_i - 1)$ slots. (If one of the a_s's comes from a shorter sequence, then leave the rest of its positions 0.) Finally, in the one remaining slot, put a 1.

[1] For those who want to know, the coding is done as follows: If a_1 is the vertex $t_{i-1}^{r_1}$, then the extra A_{i-1} slots are all 0's. If a_1 is in an unused $C_{d,i-1}^{r_1}$, then the first two extra slots contain d's and the rest are 0's. Finally, if a_1 is in $C_{b_{1q},i-1}^{r_1}$ then put b_{1q} into the q^{th} extra slot and leave the rest 0.

Let $G_j = G_j^j$. It follows just as in Equation 3.7 that, $G_j^{i-1} \preceq_{w_i} G_j$.

Again recall that each A_i and w_{i+1} is a fixed constant. Thus, given a tuple, $a_1, \ldots, a_{f_i}$ from $|G_j|$, a first-order formula, $\psi_i(a_1, \ldots, a_{f_i})$, can express the existence of the gadget T_i^r that codes this tuple. Thus, for all $j \ge i$,

$$G_j \models (\Gamma_i(a_1, \ldots, a_{f_i}) \leftrightarrow \psi_i(a_1, \ldots, a_{f_i}))$$

This complete the proof of Theorem 3.1. □

We should note that Theorem 3.1 did not use any properties of (FO + LFP) except that the language is countable and each formula had a constant number of variables. We thus have the following extension:

Corollary 3.10 *Let $\mathcal{L}$ be any countable subset of formulas about graphs from $L^{\omega}_{\infty\omega}$. Then there exists a set of finite graphs, $\mathcal{F}$, that admits unbounded fixed points and such that over $\mathcal{F}$ every formula from $\mathcal{L}$ is equivalent to a first-order formula.*

3.4 Two Extensions and an Open Problem

The deterministic construction relied heavily on Lemma 3.5. This in turn depends on the fact that $L^{\omega}_{\infty,\omega}$ on unordered structures is not expressive enough to count.

In [CFI] a lower bound was proved on the language (FO + COUNT + LFP). This is a language over two-sorted structures: one sort is the numbers: $\{0, 1, \ldots, n-1\}$ equipped with the usual ordering. The other sort is the vertices: $\{v_0, v_1, \ldots, v_{n-1}\}$ with the edge predicate. The interaction between the two sorts is via counting quantifiers. For example, the formula,

$$(\exists i\, x)\varphi(x)$$

means that there exist at least i vertices x such that $\varphi(x)$. Here i ranges over numbers and x over vertices. The least fixed point operator may be applied to relations over a combination of number and vertex variables. Define the language $(L + \text{COUNT})^{\omega}_{\infty,\omega}$ to be the superset of (FO + COUNT + LFP) obtained by adding counting quantifiers to $L^{\omega}_{\infty,\omega}$.

In [CFI] it is shown that the language (FO + COUNT+LFP) – and in fact even $(L+\text{COUNT})^{\omega}_{\infty,\omega}$ – does not express all polynomial-time properties, even over structures of color class size four. Such structures are "almost ordered": they consist of an ordered set of $n/4$ color classes, each of size four. Only the vertices inside these color classes are not ordered. We glean the following fact from [CFI].

Fact 3.11 ([CFI]) *For each $n > 0$ there exist nonisomorphic graphs T_n and $\widetilde{T_n}$ each with $O(n)$ vertices, such that T_n and $\widetilde{T_n}$ are indistinguishable by all formulas with at most n variables from* (FO + LFP + COUNT), *or even from* $(L + \text{COUNT})^{\omega}_{\infty,\omega}$.

Useful in the proof of Fact 3.11 as well as in the next theorem is the following modification of the game Γ^k_τ of Fact 2.1. Given a pair of τ-structures G and H define the $\mathcal{C}^k_\tau$ game on G and H as follows: Just as in the Γ^k_τ game, we have two players and k pairs of pebbles. The difference is that each move now has two parts.

1. Spoiler picks up the pair of pebbles numbered i for some i. He then chooses a set A of vertices from one of the graphs. Now Duplicator answers with a set B of vertices from the other graph. B must have the same cardinality as A.
2. Spoiler places one of the pebbles numbered i on some vertex $b \in B$. Duplicator answers by placing the other pebble numbered i on some $a \in A$.

The definition for winning is as before. What is going on in the two part move is Spoiler asserts that there exist $|A|$ vertices in G with a certain property. Duplicator answers with the same number of such vertices in H. Spoiler challenges one of the vertices in B and Duplicator replies with an equivalent vertex from A. This game captures expressibility in $(L + \text{COUNT})^{\omega}_{\infty,\omega}$:

Fact 3.12 ([IL]) *The Duplicator has a winning strategy for the $\mathcal{C}^k_\tau$ game on G, H if and only if G and H agree on all formulas with at most k variables from* $(L + \text{COUNT})^{\omega}_{\infty,\omega}$.

Using the above facts, we now prove a counterexample to a weaker version of McColm's Conjecture:

Theorem 3.13 *There exists a set of finite directed graphs, $\mathcal{J} = \{J_1, J_2, \ldots\}$, such that $\mathcal{J}$ admits fixed points of unbounded depth and yet on $\mathcal{J}$,* FO = (FO + COUNT+LFP), *i.e., every formula expressible with a least fixed point operator and counting is already first-order expressible. In fact, this statement remains true when* (FO+COUNT+LFP) *is replaced by an arbitrary countable subset of* $(L + \text{COUNT})^{\omega}_{\infty,\omega}$.

proof The idea of this construction is that everywhere we started with a clique of size n in the previous proof, we will start with a chain of copies of the graph T_n from Fact 3.11. Then where previously we increased the size of the clique to code some number b of bits, we will instead flip some copies of T_n to $\widetilde{T_n}$, in a particular length b chain of T_n's.

The main differences are that unlike the cliques, there is not an automorphism mapping every point in T_n to every other point in T_n. Furthermore, T_n is distinguishable from T_{n+1} using a small number of variables.

Let $f(j)$ be the number of formulas that are handled by the structure G_j, and let $v(j)$ be $v_{f(j)}$, the number of variables to be handled as in the proof of Theorem 3.1. Observe that $f(j)$ and thus $v(j)$ may be chosen to grow very slowly. In particular, we will make sure that $f(j)$, and in fact the number of vertices in each $T_{v(j)}$ is less than j. Recall also that the graphs T_n from Fact 3.11 are ordered up to sets of size four. We introduce two new binary relations: Red edges from each vertex in each $T_{v(i)}$ to the vertex $i \in D^0_j$, and Blue edges from each of the four vertices numbered k in any of the $T_{v(i)}$'s to the vertex $k \in D^0_j$. Thus, any vertex chosen from G_j will have a "name" that consists of a pair of vertices from D^0_j, together with a bounded number of bits.

The construction and proof now follow as in the proof of Theorem 3.1. □

We also show,

Corollary 3.14 *If* P ≠ PSPACE, *then there exists a set $\mathcal{C}$ of finite structures such that* FO = (FO + LFP) *on $\mathcal{C}$; but,* FO ≠ (FO + ITER) *on $\mathcal{C}$.*

proof Let $\mathcal{G}$ be the set of all finite, ordered graphs. If P ≠ PSPACE, then there is a property $S \subset \mathcal{G}$ such that $S \in$ PSPACE − P. Now, do the construction of Theorem 3.1, starting with $\mathcal{G}$. This construction assures that FO = (FO + LFP) on the resulting set $\mathcal{C}$. However, any first-order formula φ has a fixed number, k, of variables. Thus, to φ, the noticeable changes during the construction involve at most k PTIME properties. Therefore, S is still not recognizable in FO over $\mathcal{C}$. □

One special case of McColm's conjecture remains open. This is a fascinating question in complexity theory and logic related to uniformity of circuits and logical descriptions, cf. [BIS]. Consider the structures $\mathcal{B} = \{B_1, B_2, \ldots\}$ where $B_i = \langle\{0, 1, \ldots, i-1\}, \leq, \text{BIT}\rangle$. Here $\leq$ is the usual ordering on the natural numbers and BIT(x, y) holds iff the x^{th} bit in the binary representation of the number y is a one.

Question 3.15 *Is* FO = (FO + LFP) *over $\mathcal{B}$?*

The answer to Question 3.15 is "Yes," iff every polynomial-time computable numeric predicate is already computable in (FO + BIT). Equivalently, the answer to Question 3.15 is "Yes," iff deterministic log-time uniform AC^0 is equal to polynomial-time uniform AC^0, cf. [BIS]. A resolution of this question would thus answer an important question in complexity theory.

4 The Randomized Construction

We now sketch a quite different construction that also disproves McColm's conjecture. Throughout this construction, P is a binary predicate. We will prove:

Theorem 4.1 *Suppose that K_1 is a class of structures of some vocabulary τ_1, and $\mathcal{L}$ is an arbitrary countable subset of $L^{\omega}_{\infty,\omega}$. Let τ_2 be the extension of τ_1 with an additional binary predicate P. There exist a class K_2 of τ_2-structures such that:*

1. *K_1 is precisely the class of τ_1-reducts of substructures $M_2 \mid \{x \mid P(x,x)\}$ where M_2 ranges over K_2.*

2. *Every $\mathcal{L}$-formula is equivalent to a first-order formula in K_2.*

The idea of the proof is relatively simple. Let $\rho_1, \rho_2, \ldots$ be a list of all $\mathcal{L}$-definable global relations on K_1. We attach a graph G to every $M \in K_1$ and define a projection function from elements of the new sort to elements of the old sort. Relations ρ_i^M on the old sort are coded by cliques of G; a tuple $\bar{a}$ belongs to ρ_i^M if and only if there is clique of cardinality i projected in a certain way onto $\bar{a}$. The necessity to have appropriate cliques is the only constraint on G; otherwise the graph is random. We check that every $\mathcal{L}$-definable global relation reduces by first-order means to $\mathcal{L}$-definable global relations on the old sort and thus is first-order expressible. In fact, we beef $\mathcal{L}$ up before executing the idea.

Let H be a hypergraph of cardinality ≥ 2.

Definition 4.2 An *envelope* for H is a $\{P\}$-structure E satisfying the following conditions:

- $|H| \subseteq |E|$, and P is the identity relation on $|H|$.
- P is irreflexive and symmetric on $|E| - |H|$.
- For every $x \in |E| - |H|$, there is a unique $a \in H$ with $E \models P(x,a)$.
- For every $a \in |H|$ and every $x \in |E| - |H|$, $E \models \neg P(a,x)$.

□

Let E range over envelopes for H such that $|E| - |H| \neq \emptyset$.

Definition 4.3 Elements of H are *nodes* of E and elements of $|E|-|H|$ are *vertices* of E. G_E is the graph formed by P on the vertices. If $E \models P(x,a)$ and $a \in H$ then a is called the *projection* of x and denoted $F(x)$ (or Fx). If X is a set of elements of E then $F(X)$ is the multiset $\{\!\{Fx \mid x \in X\}\!\}$. If $\bar{x}$ is a sequence $(x_1, \ldots, x_l)$ of elements of E then $F(\bar{x}) = (F(x_1), \ldots, F(x_l))$. □

Let k be a positive integer ≥ 3.

Definition 4.4 A clique X of G_E is a *k-clique* if $F(X) \in \mathrm{HE}(H)$ and $||X|| < k$. A vertex that does not belong to any k-clique is *k-plebeian.* The *k-closure* $C_k(X)$ of a subset X of E is the union of X and all k-cliques intersected by X. □

Definition 4.5 E is *k-good* for H if it satisfies the following conditions.

$G_0(k)$ All k-cliques are pairwise disjoint.

$G_1(k)$ For every $X \subseteq |E|$ of cardinality $< k$, there is a k-plebeian vertex $z \in |E| - X$ with a predefined projection Fz which is P-related to $C_k(X)$ in any predefined way that does not destroy any k-clique $C \subseteq C_k(X)$. In other words, if a is a node, $Y \subseteq C_k(X)$ and Y does not include any k-clique, then there is a k-plebeian vertex $z \in F^{-1}(a) - X$ adjacent to every vertex in Y and to no vertex in $C_k(X) - Y$.

$G_2(k)$ For every $X \subseteq |E|$ of cardinality $< k$, there is a k-clique $\{y_1, \ldots, y_l\} \subseteq |E| - X$ with any predefined projections Fy_m and any predefined pattern $R = \{(x,m) \mid E \models P(x, y_m)\}$ that does not destroy any k-clique $C \subseteq C_k(X)$. In other words, if $\bar{a} = (a_1, \ldots, a_l)$ is a tuple of nodes, $l < k$, $\mathrm{MS}(\bar{a})$ is a hyperedge, $R \subseteq C_k(X) \times \{1, \ldots, l\}$, no vertex is R-adjacent to all the numbers, and no number is R-adjacent to all vertices of any k-clique $C \subseteq C_k(X)$, then there is a tuple $\bar{y} = (y_1, \ldots, y_l)$ of distinct vertices such that $F(\bar{y}) = \bar{a}$, $\{y_1, \ldots, y_l\}$ is a clique disjoint from X, and $E \models P(x, y_m) \iff (x,m) \in R$ for all $x \in C_k(X)$ and all m.

□

Lemma 4.6 1. *If E is k-good, $X \subseteq E$ and $||E|| < k$ then $||C_k(X)|| \leq (k_1)^2$.*

2. *If E is k-good then every hyperedge of cardinality $< k$ is the projection of some k-clique.*

3. *In every k-good envelope, every clique C of cardinality $< k$ is a k-clique. Moreover, if a clique $C \subseteq C_k(X)$ for some X of cardinality $< k$ then C is a k-clique.*

4. *Let H' be the hypergraph obtained from H by discarding all hyperedges of cardinality $\geq k$. Then E is k-good for H if and only if it is k-good for H'.*

5. *If E is k'-good for H where $k' > k$ then E is k-good for H.*

proof Omitted due to lack of space. □

Theorem 4.7 *There exists a k-good envelope for H.*

proof Omitted due to lack of space. □

4.1 The Game

Let M be a structure of some vocabulary τ_0 such that every element of M interprets some individual constant. It is supposed that τ_0 does not contain the fixed binary predicate P. Let H be a hypergraph on $|M|$, so that $|H| = |M|$. An envelope E for H can be seen as a structure of vocabulary $\tau = \tau_0 \cup \{P\}$ where the τ_0-reduct of the substructure $E \mid |H|$ equals M and no τ_0 relation involves elements of $|E| - |H|$.

Fix a positive integer k and let E and E' range over k-good envelopes for H. We will prove that Duplicator has a winning strategy in $\Gamma^k_\tau(E, E')$.

Definition 4.8 A partial isomorphism η from E to E' is *k-correct* if it satisfies the following conditions where x ranges over $\mathrm{Dom}(\eta)$.

- If x is a node then $\eta(x) = x$.
- If x is a vertex then $\eta(x)$ is a vertex and $F(\eta(x)) = Fx$.
- x is k-plebeian if and only if $\eta(x)$ is k-plebeian.
- If x belongs to some k-clique X then $\eta(x)$ belongs to some k-clique X' such that $F(X') = F(X)$.

□

Definition 4.9 A k-correct partial isomorphism η from E to E' is *k-nice* if there exists an extension of η to a k-correct partial isomorphism η^* with domain $C_k(\mathrm{Dom}(\eta))$. □

Lemma 4.10 *Suppose that η is a k-nice partial isomorphism from E to E'. Then η^* and η^{-1} are k-nice, $(\eta^*)^{-1} = (\eta^{-1})^*$, and $Range(\eta^*) = C_k(Range(\eta))$. η^* maps every k-clique onto k-clique of the same size, different k-cliques are mapped to different k-cliques.*

proof Obvious. □

Definition 4.11 An even-numbered state of $\Gamma^k_\tau(E, E')$ is *good* if the pebble-defined map is a k-nice partial isomorphism. A strategy of Duplicator in $\Gamma^k_\tau(E, E')$ is *good* if every move of Duplicator creates a good state. □

Theorem 4.12 *Every good strategy of Duplicator wins $\Gamma^k_\tau(E, E')$, and Duplicator has a good strategy.*

proof Omitted due to lack of space. □

Definition 4.13 A *0-table* is a conjunction $\alpha(v_1, \ldots, v_l)$ of atomic and negated atomic formulas in vocabulary $\{P\}$ which describes the isomorphism type of a $\{P\}$-structure of cardinality $\leq l$ which can be embedded into some envelope for some hypergraph. □

Definition 4.14 Let $j < k$ be a positive integer. A *(j,k)-table* is a first-order $\{P\}$-formula $\beta(v_1, \ldots, v_l)$ which says that there are distinct elements $u_1, \ldots, u_j$ such that $\{u_1, \ldots, u_j\}$ is a clique intersecting $\{v_1, \ldots, v_l\}$ and a particular 0-table $\beta_0(u_1, \ldots, u_s, v_1, \ldots, v_l)$ is satisfied. □

Definition 4.15 A *k-table* $\gamma(v_1, \ldots, v_l)$ is a conjunction such that:

- Some 0-table $\alpha(v_1, \ldots, v_l)$ is a conjunct of $\gamma(v_1, \ldots, v_l)$.
- If $j < k$ and $\beta(v_1, \ldots, v_l)$ is a (j,k)-table consistent with $\alpha(v_1, \ldots, v_l)$ then either $\beta(v_1, \ldots, v_l)$ or $\neg\beta(v_1, \ldots, v_l)$ is a conjunct of $\gamma(v_1, \ldots, v_l)$.
- There are no other conjuncts.

□

Fix a k-variable infinitary τ-formula $\varphi(u_1, \ldots, u_l, v_1, \ldots, v_m)$ and let $\Phi(\bar{u}, \bar{v})$ be the conjunction of $\varphi(\bar{u}, \bar{v})$ and some k-table $\gamma(\bar{v})$. Let $\bar{a}$ be an l-tuple of nodes of H and b be an m-tuple of nodes H. We introduce a relation $\Phi^-(\bar{u}, \bar{v})$ on H.

Definition 4.16

$$\Phi^-(\bar{a},\bar{b}) \iff E \models (\exists \bar{v})[(\Phi(\bar{a},\bar{v})) \wedge F(\bar{v}) = \bar{b}].$$

□

Lemma 4.17 *Φ^- does not depend on the choice of E: any other k-good envelope for H yields the same relation.*

proof It suffices to check that E' yields the same relation. Since Duplicator has a winning strategy in $\Gamma^k_\tau(E,E')$, no infinitary k-variable τ-sentence distinguishes between E and E'. In particular, no sentence

$$(\exists v_1,\ldots,v_m)[P(v_1,d_1) \wedge \ldots \wedge P(v_m,d_m) \wedge \Phi(c_1,\ldots,c_l,v_1,\ldots,v_m)],$$

where $c_1,\ldots,c_l,d_1,\ldots,d_m$ are individual constants, distinguishes between E and E'. □

Theorem 4.18 *Let $\bar{x}$ be an m-tuple of vertices in E. The following claims are equivalent:*

1. $E \models \Phi(\bar{a},\bar{x})$.
2. $H \models \Phi^-(\bar{a},F(\bar{x}))$ *and* $E \models \gamma(\bar{x})$.

proof Omitted due to lack of space. □

In the case $m = 0$, $\Phi = \Phi^- = \varphi$ and we have the following corollary.

Corollary 4.19

$$E \models \varphi(\bar{a}) \iff H \models \varphi(\bar{a}).$$

4.2 Proof of Theorem 4.1

We start with a couple of auxiliary definitions. Call an r-ary relation R *irreflexive* if every tuple in R consists of r distinct elements. Call a global relation ρ *irreflexive* if every local relation ρ^M is so.

Lemma 4.20 *Every global relation $\rho(v_1,\ldots,v_r)$ is a positive boolean combination of irreflexive global relations definable from ρ in a quantifier-free way.*

proof Omitted due to lack of space. □

Call a multiset A is *oriented* if the relation $\mathrm{MP}(a) < \mathrm{MP}(b)$ is a linear order on $\mathrm{Set}(A)$; let $\mathrm{OSet}(A)$ be the corresponding linearly ordered set.

Now we are ready to prove Theorem 4.1. Suppose that K_1 is a class of structures of some vocabulary τ_1, and τ_2 is the extension of τ_1 with binary predicate P. Let $\mathcal{L}$ be an arbitrary countable set of $L^\omega_{\infty,\omega}$-formulas.

A global relation ρ on a class K is *decidable* if there exists an algorithm that, given (the encodings of) a structure $M \in K$ and a tuple $\bar{a}$ of elements of M of appropriate length, decides whether $M \models \rho(\bar{a})$ or not. We are interested in a relativized version of this definition where K is the collection of all structures (that is, all finite structures) in the vocabulary of ρ. Let

$$\Omega = \{(\varphi, M, \bar{a}, 1) \mid \varphi \in \mathcal{L} \wedge M \models \varphi(\bar{a})\} \cup \{(\varphi, M, \bar{a}, 0) \mid \varphi \in \mathcal{L} \wedge M \not\models \varphi(\bar{a})\}$$

Definition 4.21 A global relation ρ of vocabulary τ is *$\mathcal{L}$-decidable* if there is an algorithm with oracle Ω that, given a τ-structure M and a tuple $\bar{a}$ of elements of M of appropriate length, decides whether $M \models \rho(\bar{a})$ or not. □

Every global relation defined by a formula in $\mathcal{L}$ is $\mathcal{L}$-decidable, and there there are only countably many $\mathcal{L}$-decidable relations. List all $\mathcal{L}$-decidable irreflexive global relations on K_1 of positive arities: $\rho_2, \rho_3\, \rho_4, \ldots$, and let r_i be the arity of ρ_i. We suppose that $r_i(r_i+1)/2 \leq i$. Let M range over K_1 and i range over positive integers ≥ 2.

For each M and each i, let σ^M_i be the collection of oriented multisets A such that $\mathrm{OSet}(A) \in \rho^M_i$ and $||A|| = i$. Since $1 + 2 + \ldots + r_i = r_i(r_i+1)/2 \leq i$, σ^M_i is empty. Let $H(M)$ be the hypergraph

$$\left(|M|\,,\,\bigcup\{\sigma^M_i \mid 1 \leq i \leq ||M||\}\right).$$

Set $\tau_2 = \tau_1 \cup \{P\}$ and let $\mathcal{E}(M)$ be the collection of $||M||$-good envelopes for $H(M)$ of minimal possible cardinality. (The minimal cardinality is not important; we will use only the following two consequences: (i) $\mathcal{E}(M)$ is finite, and (ii) there is an algorithm that, given M constructs some $E \in \mathcal{E}(M)$.) View envelopes $E \in \mathcal{E}(M)$ as τ_2-structures where the τ_1-reduct of the substructure $E \mid |M|$ equals M and no τ_1-relation involves elements of $|E| - |M|$. For every $K \subseteq K_1$, let $\mathcal{E}(K) = \bigcup_{M \in K} \mathcal{E}(M)$. Finally, let $K_2 = \mathcal{E}(K_1)$. By the definition of envelopes (Definition 4.2), K_2 satisfies requirement 1 of Theorem 4.1. In order to prove requirement 2, it suffices to prove that every infinitary formula with $\mathcal{L}$-decidable global relation is first-order definable in K_2.

For any global relation $\rho(\bar{v})$ on K_1, let $\rho^+(\bar{v})$ be the global relation on K_2 such that

$$E \models \rho^+(\bar{x}) \iff M \models \rho(F(\bar{x}))$$

if $M \in K$, $E \in \mathcal{E}(M)$ and $\bar{x}$ is a tuple of elements of E of the appropriate length.

Lemma 4.22 *If ρ is $\mathcal{L}$-decidable then ρ^+ is first-order definable in K_2.*

proof Omitted due to lack of space. □

Now let φ be an arbitrary infinitary τ_2-formula whose global relation is $\mathcal{L}$-decidable. We prove that φ is equivalent to a first-order formula in K_2. Without loss of generality, $\varphi = \varphi(u_1, \dots, u_l, v_1, \dots, v_m)$ and φ implies

$$P(u_1, u_1), \dots, P(u_l, u_l), \neg P(v_1, \bar{v}_1), \dots, \neg P(v_m, v_m)$$

In other words, variables u_i are node variables, and variables v_j are vertex variables.

Let k be the total number of variables in φ, $K_1' = \{M \mid ||M|| \geq k\}$ and $K_2' = \mathcal{E}(K_1')$, so that every $E \in K_2'$ is k-good. Since $K_2 - K_2'$ is finite, it suffices to prove that $\varphi(\bar{u}, \bar{v})$ is equivalent to a first-order formula in K_2'. Let $\gamma(\bar{v})$ be an arbitrary k-table. Since there are only finite many k-tables, it suffices to prove that the formula $\Phi(\bar{v}) = \varphi(\bar{v}) \wedge \gamma(\bar{v})$ is equivalent to a first-order formula over K_2'.

Define a global relation Φ^- on K_1 as follows:

$$M \models \Phi^-(\bar{a}, \bar{b}) \iff (\exists \bar{x})[(E \models \Phi(\bar{x})) \wedge F(\bar{x}) = \bar{a}]$$

where $E \in \mathcal{E}(M)$. The choice of E does not matter. Indeed, extend τ_1 with individual constants for each element of M; call the resulting vocabulary τ_0. Now apply Lemma 4.17 with $H = H(M)$.

Lemma 4.23 *Φ^- is $\mathcal{L}$-decidable.*

proof Clear. □

It is not quite true that $(\Phi^-)^+$ is the global relation of the formula Φ on K_2' but this is close to truth. By virtue of Theorem 4.18,

$$\Phi(\bar{u}, \bar{v}) \iff [(\Phi^-)^+(\bar{u}, \bar{v}) \wedge \gamma(\bar{v})$$

on K_2'. Indeed, consider any $M \in K_1'$. Extend τ_1 with individual constants for each element of M; call the resulting vocabulary τ_0. Now apply Theorem 4.18 with $H = H(M)$. By Lemma 4.22, $(\Phi^-)^+$ is first-order definable in K_2. It follows that Φ is equivalent to a first-order formula on K_2'.

Acknowledgment

We are grateful to Anuj Dawar, Phokion Kolaitis, Steven Lindell, and Scott Weinstein for stimulating discussions.

References

[AV] S. Abiteboul and V. Vianu, "Generic Computation And Its Complexity," *32nd IEEE Symposium on FOCS* (1991), 209-219.

[B] J. Barwise, "On Moschovakis Closure Ordinals," J. Symb. Logic 42 (1977), 292-296.

[BIS] D. Barrington, N. Immerman, H. Straubing, "On Uniformity Within NC[1]," *JCSS 41, No. 3* (1990), 274 - 306.

[CFI] J. Cai, M. Fürer, N. Immerman, "An Optimal Lower Bound on the Number of Variables for Graph Identification," *Combinatorica* **12** (4) (1992) 389-410.

[D] A. Dawar, "Feasible Computation Through Model Theory," PhD Dissertation, University of Pennsylvania (1993).

[G] Y. Gurevich, "Logic and the Challenge of Computer Science," in *Current Trends in Theoretical Computer Science,* ed. E. Börger, Computer Science Press, 1988, 1-57.

[I82] N. Immerman, "Upper and Lower Bounds for First Order Expressibility," *JCSS* **25**, No. 1 (1982), 76-98.

[I89] N. Immerman, "Descriptive and Computational Complexity," *Computational Complexity Theory,* ed. J. Hartmanis, *Proc. Symp. in Applied Math.,* 38, American Mathematical Society (1989), 75-91.

[IL] N. Immerman and E. S. Lander, "Describing Graphs: A First-Order Approach to Graph Canonization," in *Complexity Theory Retrospective,* Alan Selman, ed., Springer-Verlag, 1990, 59-81.

[KV] Ph. Kolaitis and M. Vardi, "Fixpoint Logic vs. Infinitary Logic in Finite-Model Theory," LICS 1992, 46-57.

[M] G. McColm, "When is Arithmetic Possible?" *Annals of Pure and Applied Logic* 50 (1990), 29-51.

[V] M. Vardi, "Complexity of Relational Query Languages," *14th Symposium on Theory of Computation* (1982), 137-146.