A Finite Model–Theoretical Proof of a Property of Bounded Query Classes within PH

Leszek Aleksander Kołodziejczyk
Institute of Philosophy
Warsaw University
Krakowskie Przedmieście 3, 00-047 Warsaw, Poland
e-mail: l.kolodziejczyk@zodiac.mimuw.edu.pl

April 29th, 2004

Abstract

We use finite model theory (in particular, the method of FM–truth definitions, introduced in [MM01] and developed in [K04], and a normal form result akin to those of [Ste93] and [G97]) to prove:

Let $m \geq 2$. Then:

(A) If there exists k such that $NP \subseteq \Sigma_m TIME(n^k) \cap \Pi_m TIME(n^k)$, then for every r there exists k_r such that $P^{NP[n^r]} \subseteq \Sigma_m TIME(n^{k_r}) \cap \Pi_m TIME(n^{k_r})$;

(B) If there exists a superpolynomial time–constructible function f such that $NTIME(f) \subseteq \Sigma_m^p \cap \Pi_m^p$, then additionally $P^{NP[n^r]} \subsetneq \Sigma_m^p \cap \Pi_m^p$.

This strengthens a result by Mocas ([M96]) that for any r, $P^{NP[n^r]} \subsetneq NEXP$.

In addition, we use FM–truth definitions to give a simple sufficient condition for the Σ_1^1 arity hierarchy to be strict over finite models.

It is widely believed that the polynomial hierarchy is properly contained in the nondeterministic exponential time class $NEXP$. However, even a fairly restricted fragment of the polynomial hierarchy — P^{NP}, the class of problems solvable in polynomial time by a deterministic machine with access to an NP oracle — has not been proven to be smaller than $NEXP$. Moreover, the problem whether P^{NP} equals $NEXP$ seems difficult, as it is known to have contradictory relativizations (see [BT94] or [BFFT01] for an oracle under which $P^{NP} = NEXP$).

The largest parts of PH currently known to be separated from $NEXP$ are the bounded query fragments of P^{NP}. In [FLZ92], it was shown that $P^{NP[n^{o(1)}]} \subsetneq NEXP$. Later, Mocas ([M96]) improved this result by showing that for any fixed r, $P^{NP[n^r]}$ is also a proper subclass of $NEXP$.

In the present paper, we prove the following theorem:

Theorem 0.1. *Let $m \geq 2$. Then:*

(A) If there exists k such that $NP \subseteq \Sigma_m TIME(n^k) \cap \Pi_m TIME(n^k)$, then for every r there exists k_r such that $P^{NP[n^r]} \subseteq \Sigma_m TIME(n^{k_r}) \cap \Pi_m TIME(n^{k_r})$;

(B) If there exists a superpolynomial time–constructible function f such that $NTIME(f) \subseteq \Sigma_m^p \cap \Pi_m^p$, then additionally $P^{NP[n^r]} \subsetneq \Sigma_m^p \cap \Pi_m^p$.

Note that the theorem can be interpreted as a strengthening of Mocas' result, which immediately follows from either (A) or (B). Indeed, consider e.g. (B) and assume that $P^{NP[n^r]} = NEXP$. Then the hypothesis of (B) is satisfied for $m = 2$, so $P^{NP[n^r]} \subsetneq \Sigma_2^p \cap \Pi_2^p \subseteq NEXP = P^{NP[n^r]}$. Contradiction.

The proof uses finite model theory. We show that any $P^{NP[n^r]}$ property of finite models can be expressed by a second order sentence in a particular normal form, very similar to normal forms studied in [Ste93] and [G97]. We then show that if the assumption of (A) holds for a given m, then there is a Σ_m^1 FM–truth definition (in the sense of M. Mostowski; see [MM01]) for the class of sentences in the normal form, from which the conclusion of (A) follows via a result of [K04] and the fact that $P^{NP[n^r]}$ is closed under complementation. (B) is established in a similar way.

The paper is divided into four sections. After the preliminary section 1, we prove (A) of theorem 0.1 in section 2 and (B) in section 3. Section 4 offers some generalizations, remarks and related results, in particular a simple sufficient condition for the arity hierarchy of Σ_1^1 to be strict in finite models.

1 Preliminaries

All models are finite with built–in arithmetic. In other words, the universe of a model $\mathbf{M}$ is always an initial segment $M = \{0, \ldots, M-1\}$ of the natural numbers, and the vocabulary σ always contains a fixed arithmetical subvocabulary σ_0 (consisting, say, of the predicates $+, \times, \leq$ and the constants $0, MAX$) whose symbols are interpreted in the standard way (e.g. $+(x, y, z)$ holds in a model iff $x + y = z$ in $\mathbb{N}$).

Vocabularies are finite and relational, although individual constants are allowed. A model of vocabulary σ is referred to as a σ–model.

A logic is any function $\mathcal{L}$ which assigns to a vocabulary σ a decidable set of words over some alphabet (the set of $\mathcal{L}$–sentences over σ) and a decidable relation $\models^{\mathcal{L}}_{\sigma}$ between σ–models and $\mathcal{L}$–sentences over σ (the truth relation, normally denoted by just $\models$ if no confusion arises). If φ is an $\mathcal{L}$–sentence over σ, then $MOD(\varphi)$ denotes the set of σ–models $\mathbf{M}$ such that $\mathbf{M} \models \varphi$. We say that $MOD(\varphi)$ is defined by φ.

The notion of a logic closed e.g. under negation is defined in the natural way. We assume familiarity with commonly used logics such as first order logic (FO), second order logic (SO) and the prenex classes of second order logic (Σ^1_m). In second order formulae, we sometimes use numerical superscripts to indicate the arity of relational variables — thus, R^r is an r–ary variable. We also often use the same symbols to denote relations and relational variables.

If $\mathcal{L}$ and $\mathcal{L}'$ are logics, then $\mathcal{L} \leq \mathcal{L}'$ ("$\mathcal{L}'$ is at least as expressive as $\mathcal{L}$") means that for any $\mathcal{L}$–sentence φ there is an $\mathcal{L}'$–sentence φ' such that $MOD(\varphi) = MOD(\varphi')$. $\mathcal{L} \equiv \mathcal{L}'$ holds if $\mathcal{L} \leq \mathcal{L}'$ and $\mathcal{L}' \leq \mathcal{L}$. $\mathcal{L} < \mathcal{L}'$ holds if $\mathcal{L} \leq \mathcal{L}'$ but not $\mathcal{L} \equiv \mathcal{L}'$. For any vocabulary σ, $\mathcal{L} \leq_{\sigma} \mathcal{L}'$ means that for any $\mathcal{L}$–sentence φ of vocabulary σ there is an $\mathcal{L}'$–sentence φ' such that $MOD(\varphi) = MOD(\varphi')$. $\mathcal{L} \equiv_{\sigma} \mathcal{L}'$ and $\mathcal{L} <_{\sigma} \mathcal{L}'$ are defined accordingly.

We assume familiarity with standard computational complexity classes such as P, NP or PH. Recall that the Σ^p_m level of PH can be defined as $\bigcup_{k\in\omega} \Sigma_m TIME(n^k)$, where $\Sigma_m TIME(f)$ consists of problems solvable in time f by a Σ_m machine, i.e. an alternating machine which starts in an existential state and makes at most $(m-1)$ alternations between existential and universal states on any input. The same holds for Π^p_m and the dual notion of $\Pi_m TIME$.

The class $P^{NP[n^r]}$ consists of those problems which may be solved by a deterministic polynomial–time oracle machine which makes at most n^r queries to an NP oracle on an input of size n (note that when considering problems defined as classes of σ–models for some σ, we take the size of a given σ–model $\mathbf{M}$ to be equal to M, and not to the length of a standard code for $\mathbf{M}$, which may be polynomially larger). The nondeterministic exponential time class $NEXPTIME$ is defined as $\bigcup_{k\in\omega} NTIME(2^{n^k})$.

If $\mathcal{L}$ is a logic and $\mathcal{C}$ is a complexity class, we say that $\mathcal{L}$ captures $\mathcal{C}$ if for any σ and any class of σ–models $\mathcal{K}$, it holds that $\mathcal{K} \in \mathcal{C}$ if and only if $\mathcal{K} = MOD(\varphi)$ for some $\mathcal{L}$–sentence φ. If every $\mathcal{L}$–definable class $\mathcal{K}$ (of σ–models) is in $\mathcal{C}$, we say that model checking for $\mathcal{L}$ (over σ) is in $\mathcal{C}$. If every class in $\mathcal{C}$ (of σ–models) is $\mathcal{L}$–definable, $\mathcal{L}$ is said to capture at least $\mathcal{C}$ (over σ)[1].

A well-known result is that Σ^1_m captures Σ^p_m for any m ([F74],[Sto77]).

1.1 FM–truth definitions

We prove our main result using FM–truth definitions. The idea of FM–truth definitions was introduced in [MM01] and developed in [K04]. The present subsection discusses the basic definitions and results related to this concept.

Definition 1.1. A relation $R \subseteq \omega^n$ is *FM–represented* by the (first order) σ_0–formula $\varphi(\mathbf{x})$ if and only if: for any $\mathbf{a} \in \omega^n$, $\varphi(\mathbf{a})$ is true in almost all σ_0–models if $R(\mathbf{a})$ holds, and false in almost all σ_0–models if $R(\mathbf{a})$ does not hold. R is *FM–representable* if there is a formula which FM–represents it.

Theorem 1.2 ([MM01]; FM-representability theorem)**.** *$R \subseteq \omega^n$ is FM–representable if and only if it is of degree $\leq \mathbf{0}'$ (recursive with an RE oracle).*

The notion of FM–representability was intended as a finite model counterpart of the classical notions of *definability* (in a model, esp. in the standard model of arithmetic) or *representability* (in a theory) of arithmetical relations. One consequence of the FM–representability theorem is that, just as in the standard model of arithmetic or in arithmetical theories, also in finite models we may freely talk about operations and relations connected to the syntax of logics (such as "formula x is the result of preceding formula y with

[1] Note that it is more usual to define "$\mathcal{L}$ captures $\mathcal{C}$" and the other related notions over models with a built–in linear ordering, and not built–in arithmetic. In our framework, some logics (such as first order logic) may capture larger complexity classes than in the usual one. However, if a logic semantically contains deterministic transitive closure logic (DTC), it is able to define the arithmetical relations from the ordering, so its expressive power is the same in both frameworks.

an existential quantifier) — all these relations are decidable, so we simply use the formulae which FM–represent them. Care must be taken, however, since FM–representation works "asymptotically": a formula φ which FM–represents some relation will only tell us whether a given tuple $\mathbf{a}$ is in the relation if we look at the truth value of $\varphi(\mathbf{a})$ in a sufficiently large model.

Henceforth, $\ulcorner w \urcorner$ stands for the Gödel number of the string w (we assume that some appropriate gödelization has been carried out).

The concept of truth FM–truth definition is a finite model analogue of Tarski's notion of truth definiton:

Definition 1.3. Let $\mathcal{L}$ be a logic and σ be a vocabulary. We say that the σ–formula $Tr_{\mathcal{L},\sigma}(x)$ is an *FM–truth definition* for $\mathcal{L}$ over σ if and only if for every $\mathcal{L}$–sentence ψ of vocabulary σ,

$$\mathbf{M} \models \psi \equiv Tr_{\mathcal{L},\sigma}(\ulcorner \psi \urcorner)$$

holds for almost all σ–models $\mathbf{M}$.

Remark. Under our definition of logic, there are logics for which the notion of a formula with a free (first order) variable does not, strictly speaking, make sense. The formal way to deal with this problem is to think of such "formulae" as sentences of vocabulary $(\sigma + c)$, where c is a new individual constant. The details of a suitable redefinition of "FM–truth definition" are left to the reader.

For any logics $\mathcal{L}, \mathcal{L}'$, if there is an $\mathcal{L}'$–formula which is an FM–truth definition for $\mathcal{L}$ over vocabulary σ, we say that $\mathcal{L}'$ defines FM–truth for $\mathcal{L}$ over σ and write $\mathcal{L} \ll_\sigma \mathcal{L}'$.

Using the FM–representability theorem, one may prove a finite model version of the Gödel diagonal lemma, and derive from it a version of Tarski's famous theorem on the undefinability of truth: no (reasonable) logic closed under negation defines FM–truth for itself.

Theorem 1.4 ([MM01]; Tarski's theorem, finite version)**.** *If $\mathcal{L}$ is a logic closed under first order quantification, forming conjunctions with first order formulae, and negation, then for any σ it is not true that $\mathcal{L} \ll_\sigma \mathcal{L}$.*

When considering one of the usual logics met in finite model theory, one can often actually give an exact characterization of the logics for which it defines FM–truth. For our current purposes, we will only need such characterizations for the Σ^1_m and Π^1_m classes:

Theorem 1.5 ([K04]). *For any m, $\mathcal{L}$ and σ: $\mathcal{L} \ll_\sigma \Sigma^1_m$ if and only if there exists a number k such that model checking for $\mathcal{L}$ over σ is in $\Sigma_m TIME(n^k)$. Analogously for Π^1_m and $\Pi_m TIME$.*

The "only if" part of theorem 1.5 is essentially trivial. The proof of the "if" part, while also simple, requires a result stating that the difficulty of defining FM–truth for a logic depends only on its expressive power, not on the peculiarities of its syntax (theorem 3.2 in [K04]).

2 $P^{NP[n^r]}$ versus Σ^p_m

We turn now to a proof of (A). The key observation is that for any given r, all $P^{NP[n^r]}$ properties of finite models can be defined by second order sentences in a particular normal form:

Lemma 2.1. *Let σ be a vocabulary and let $r \geq 1$. Then any $P^{NP[n^r]}$ class of σ–models can be defined by a sentence of the form*

$$\exists R^r(\varphi(R)\&\neg\psi(R)),$$

where φ, ψ are Σ^1_1 formulae of vocabulary σ with R as the unique free variable.

This normal form result is very similar to results proved for logics capturing relativized logarithmic space classes by I. Stewart (theorem 3.3.1 and corollary 3.3.1 of [Ste93]) and G. Gottlob (theorem 4.9 and corollary 5.3 of [G97]). The proof is also almost identical to the one given by Gottlob. We present it here in some detail so that the reader may verify that it is constructive; we will need this fact in the next section (see the proof of lemma 3.3).

The class of SO sentences in the form given by the lemma will be denoted by SNF_r ("SNF" stands for *Stewart Normal Form*, a term coined in [G97] to refer to the normal form for logics capturing relativized $LOGSPACE$).

Proof. Let $\mathcal{K}$ be a $P^{NP[n^r]}$ class of σ–models. Thus, there is a polynomial–time deterministic oracle machine T and an NP language $L \in \{0,1\}^*$ such that $\mathcal{K} = \{\mathbf{M} : \mathbf{M}$ is a σ–model and T_1 accepts $\mathbf{M}$ using L as its oracle$\}$. Furthermore, T makes at most M^r oracle queries on input $\mathbf{M}$. Thus, the string $oans(\mathbf{M})$ of all oracle answers in the computation of T on input $\mathbf{M}$

(ordered chronologically) has length at most M^r. Recall that there is a canonical correspondence between binary strings of length M^r and r–ary relations over M.

We introduce two classes of $(\sigma + R^r)$–models (where R is a relational symbol not contained in σ). The classes $\mathcal{K}_1$ and $\mathcal{K}_2$ are defined as follows:

(1) Given $(\mathbf{M}, R)$, consider the machine T_R^+ which behaves as T except that before making its i–th oracle query (for $i = 1, \ldots, M^r$), it looks at the i–th bit of the string corresponding to R — and then makes a reqular query to L if the bit is 1, simulates a negative answer without querying if the bit is 0. $(\mathbf{M}, R)$ is in $\mathcal{K}_1$ iff: (a) all queries regularly made by T_R^+ on input $\mathbf{M}$ are answered positively, and (b) T_R^+ accepts $\mathbf{M}$.

(2) Given $(\mathbf{M}, R)$, consider the machine T_R^- which behaves as T except that before making its i–th oracle query (for $i = 1, \ldots, M^r$), it looks at the i–th bit of the string corresponding to R — and then makes a reqular query to L if the bit is 0, simulates a positive answer without querying if the bit is 1. $(\mathbf{M}, R)$ is in $\mathcal{K}_2$ iff all queries regularly made by T_R^+ on input $\mathbf{M}$ are answered negatively.

Consider now the second order sentence $\gamma := \exists R^r(\varphi(R)\&\neg\psi(R))$ (where $\varphi(R)$ is φ, now treated as a formula of vocabulary σ with R as a free variable; similarly for $\psi(R)$). Certainly, γ has the required form, so it remains to check that it defines $\mathcal{K}$.

It is fairly easy to see that $\mathcal{K}_1$ is in NP and $\mathcal{K}_2$ is in *co-*NP. Let φ, ψ be Σ_1^1 sentences of vocabulary $(\sigma + R)$ such that $\mathcal{K}_1 = MOD(\varphi)$, $\mathcal{K}_2 = MOD(\neg\psi)$.

Consider now the second order sentence $\gamma := \exists R^r(\varphi(R)\&\neg\psi(R))$ (where $\varphi(R)$ is φ, now treated as a formula of vocabulary σ with R as a free variable; similarly for $\psi(R)$). Certainly, γ has the required form, so it remains to check that it defines $\mathcal{K}$.

Let $\mathbf{M} \in \mathcal{K}$ and let $R_0 \subseteq M^r$ be the relation corresponding to $oans(\mathbf{M})$ (if $oans(\mathbf{M})$ is shorter than M^r, let R_0 be any relation corresponding to a string which has $oans(\mathbf{M})$ as an initial substring). Then $(\mathbf{M}, R_0)$ is in $\mathcal{K}_1$ and $\mathcal{K}_2$, so $\mathbf{M} \models \gamma$.

Let $\mathbf{M} \models \gamma$ and let $R \subseteq M^r$ be such that $(\mathbf{M}, R) \models \varphi$, $(\mathbf{M}, R) \models \neg\psi$. Then $(\mathbf{M}, R)$ is in both $\mathcal{K}_1$ and $\mathcal{K}_2$, so, by an inductive argument, the string corresponding to R either simply is $oans(\mathbf{M})$ or contains $oans(\mathbf{M})$ as an initial substring. In any case, T_R^+ works on input $\mathbf{M}$ exactly as T does, and since $(\mathbf{M}, R) \in \mathcal{K}_1$, we know that T_R^+ accepts $\mathbf{M}$. Then so does T, and therefore $\mathbf{M}$ is in $\mathcal{K}$

Thus, the proof that γ defines $\mathcal{K}$ is completed. $\square$

Once we have the above lemma, the proof of (A) presents no further difficulty. Fix m and assume that the hypothesis of (A) holds for m. It follows from this assumption and from theorem 1.5 that for any vocabulary σ, $\Sigma^1_1 \ll_\sigma \Sigma^1_m$, $\Sigma^1_1 \ll_\sigma \Pi^1_m$.

Fix r and a vocabulary σ. Let $Sat_\Sigma(x)$ (*resp.* $Sat_\Pi(x)$) be a Σ^1_m (*resp.* Π^1_m) FM–truth definition for Σ^1_1 over the vocabulary $(\sigma + R^r)$. Consider the following formula $Tr(x)$:

$$\exists\ulcorner\varphi\urcorner\exists\ulcorner\psi\urcorner(x = \ulcorner\exists R^r(\varphi(R)\&\neg\psi(R))\urcorner \,\&\, \exists R^r(Sat_\sigma(\ulcorner\varphi\urcorner)\&\neg Sat_\Pi(\ulcorner\psi\urcorner))).$$

$Tr(x)$ is clearly equivalent to a Σ^1_m formula, and one may easily verify that it is an FM–truth definition for SNF_r over σ.

Thus, again by theorem 1.5, for every r and σ there exists a number k_r such that $P^{NP[n^r]} \subseteq \Sigma_m TIME(n^{k_r})$ over σ. In particular, such k_r exists for the vocabulary $(\sigma_0 + P^1)$ where P^1 is some chosen unary predicate. Note that models of this vocabulary can be identified with binary words. Thus, $P^{NP[n^r]} \subseteq \Sigma_m TIME(n^{k_r})$ as classes of languages over $\{0, 1\}$.

The inclusion $P^{NP[n^r]} \subseteq \Pi_m TIME(n^{k_r})$ follows from the closure of $P^{NP[n^r]}$ under complementation.

3 $P^{NP[n^r]}$ versus $\Sigma^p_m \cap \Pi^p_m$

The proof of (B) is quite similar to the proof of (A), although it requires some more care. We introduce a logic $\mathfrak{L}_r$ which captures $P^{NP[n^r]}$, and show that if $\Sigma^p_m \cap \Pi^p_m$ contains $NTIME(f)$ for a superpolynomial time–constructible f, then for every vocabulary σ, $\mathfrak{L}_r \ll_\sigma \Delta^1_m$ (i.e. there is a Σ^1_m FM–truth definition for $\mathfrak{L}_r$ which is equivalent to a Π^1_m formula). (B) will then follow by Tarski's theorem.

Remark. Δ^1_m is not necessarily a logic, even under the liberal notion of section 1, so it is a slight abuse of notation to write $\mathfrak{L}_r \ll_\sigma \Delta^1_m$. Nevertheless, it remains true that if there is a Δ^1_m FM–truth definition for some $\mathcal{L}$ over some vocabulary σ, then there is a $\Sigma^p_m \cap \Pi^p_m$ property which is not $\mathcal{L}$–definable.

The logic $\mathfrak{L}_r$ we will use is defined as follows. For any vocabulary σ, the $\mathfrak{L}_r$–sentences over σ are ordered pairs $\langle T_1, T_2\rangle$, where:

- T_1 is a (code of a) deterministic oracle machine, equipped with a polynomial time clock and a query counter which prohibits T_1 from asking more than n^r oracle queries on inputs of size n;
- T_2 is a (code of a) nondeterministic Turing machine equipped with a polynomial time clock.

The semantics is straightforward: $\mathbf{M} \models \langle T_1, T_2 \rangle$ iff T_1 accepts $\mathbf{M}$ when using the language recognized by T_2 as its oracle.

Obviously, $\mathfrak{L}_r$ captures exactly $P^{NP[n^r]}$. Note that it follows from theorem 1.5 that already if the hypothesis of (A) holds for a given m, we have $\mathfrak{L}_r \ll_\sigma \Sigma^1_m$, $\mathfrak{L}_r \ll_\sigma \Pi^1_m$ for every σ. To obtain $\mathfrak{L}_r \ll_\sigma \Delta^1_m$, however, we will need the (apparently) stronger assumption that $\Sigma^p_m \cap \Pi^p_m$ contains a superpolynomial nondeterministic time class.

We will also need a simple but important observation on FM–representing computable relations (or, more generally, recursively enumerable relations). Given a relation $R \subseteq \omega^n$ and a formula $\varphi_R(\mathbf{x})$ which FM–represents it, call φ_R *decent* if it has the property that: (a) for any $\mathbf{a} \in \omega^n$, if there exists a model $\mathbf{M}$ such that $\mathbf{M} \models \varphi_R(\mathbf{a})$, then $\mathbf{a} \in R$, and (b) if $\mathbf{M} \models \varphi_R(\mathbf{a})$, then also $\mathbf{M'} \models \varphi_R(\mathbf{a})$ whenever $M' > M$. Thus, a decent formula is one which "never falsely claims" that some tuple is in the relation it FM–represents, and additionally "never withdraws such a claim" when passing to larger models. The observation is:

Proposition 3.1. *For any recursively enumerable relation $R \subseteq \omega^n$, there is a decent formula which FM–represents it.*

Proof. This fact is implicitly contained in the proof of the FM–representability theorem in [MM01]; we give an explicit argument nonetheless. If $R \subseteq \omega^n$ is RE, then it is defined in $\mathbb{N}$ by a Σ^0_1 formula, i.e. an FO–formula of the form $\exists \mathbf{y} \psi(\mathbf{x}, \mathbf{y})$ where ψ consists of a string of *bounded* quantifiers followed by a quantifier–free matrix. It is well–known that we may assume w.l.o.g. that the quantifiers in ψ are bounded by variables and not by complex terms (see e.g. [HP93]).

The matrix of ψ is a boolean combination of equalities between some polynomials in $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$, where $\mathbf{z}$ is the tuple of variables which are quantified in ψ. Therefore, since the z's are bounded from above by the x's and/or y's, there is a polynomial $p(\mathbf{x}, \mathbf{y})$ such that for any choice of tuples $\mathbf{a}, \mathbf{b}$ to interpret $\mathbf{x}, \mathbf{y}$ respectively, the truth of $\psi(\mathbf{a}, \mathbf{b})$ does not depend on any number greater than $p(\mathbf{a}, \mathbf{b})$.

Let $\varphi_R(\mathbf{x})$ be $\exists \mathbf{y} \xi(\mathbf{x}, \mathbf{y})$, where ξ is ψ modified in the following way: the original quantifier prefix of ψ is left unchanged, and the matrix is replaced by the conjunction of an FO–formula expressing "there exists $p(\mathbf{x}, \mathbf{y})$" and a formula which arises from the original matrix by eliminating all the complex terms (i.e. substituting $\exists w(+(x_1, x_2, w) \,\&\, w = x_3)$ for $x_1 + x_2 = x_3$ etc.). It is not difficult to see that φ_R is a decent formula which FM–represents R. $\square$

Actually, one can show that being FM–represented by some decent formula is exactly equivalent to being RE. We leave out the (easy) proof of this fact as we do not need it.

We now prove a lemma which states that if the hypothesis of (B) holds for m, then there exist Σ^1_m and Π^1_m FM–truth definitions for Σ^1_1 which are, in a sense, well–behaved.

Lemma 3.2. *If $\Sigma^p_m \cap \Pi^p_m$ contains $NTIME(f)$ for some superpolynomial time–constructible f, then for any vocabulary σ there exists a Σ^1_m formula $Sat^+_\Sigma(x)$ and a Π^1_m formula $Sat^+_\Pi(x)$ such that:*

- *$Sat^+_\Sigma(x)$ and $Sat^+_\Pi(x)$ are FM–truth definitions for Σ^1_1 over σ;*
- *there is a computable function which assigns to a Σ^1_1 sentence ψ over σ a number $bound(\psi)$ such that both Sat^+_Σ and Sat^+_Π recognize the truth value of ψ correctly in all σ–models of cardinality greater than $bound(\psi)$.*

Proof. Fix σ. From now on, any model $\mathbf{M}$ appearing in the proof is a σ–model.

We may assume w.l.o.g. that $f(n)$ is of the form $n^{g(n)}$, where g is some computable function satisfying $lim_{n \to \infty} g(n) = \infty$; if not, then we may find such g for which $n^{g(n)} \leq f(n)$ — here it is not required for $n^{g(n)}$ to be time–constructible, computability will suffice.

Clearly, there exists a computable function h such that $g(h(n)) \geq n$ for all n.

It is well–known that there exists a Turing machine T which solves the problem:

"given input $(\mathbf{M}, \ulcorner \psi \urcorner)$, where ψ is a Σ^1_1 sentence,
is it true that $\mathbf{M} \models \psi$?"

in nondeterministic time $M^{lh(\psi)}$. So, the machine $\tilde{T}$ which, given input $(\mathbf{M}, \ulcorner \psi \frown 0^i \urcorner)$, disregards the string 0^i and simulates T on $(\mathbf{M}, \ulcorner \psi \urcorner)$ actually works in $NTIME(M^{g(lh(\psi^\frown 0^i))})$, and hence in $NTIME(f)$, when restricted to inputs of the form $(\mathbf{M}, \ulcorner\psi^\frown 0^{h(lh(\psi))-lh(\psi)}\urcorner)$.

Since the function h is computable, then by proposition 3.1 there is a decent formula $\varphi_h(x,y)$ which FM–represents its graph. Consider now a machine T^* which on input $(\mathbf{M}, \ulcorner\psi^\frown 0^i\urcorner)$ does the following:

- using the time–constructibility of f, keep a clock for, say, $2f$; if a computation tries to use more than this allotted amount of time, stop and reject;
- check whether $\mathbf{M} \models (\varphi_h(lh(\psi), i + lh(\psi))$; if not, reject;
- else work as $\tilde{T}$ on the input and accept iff $\tilde{T}$ does.

Clearly, for all sufficiently large $\mathbf{M}$, the machine T^* accepts $(\mathbf{M}, \ulcorner\psi^\frown 0^i\urcorner)$ if and only if $i = h(lh(\psi)) - lh(\psi)$ and $\mathbf{M} \models \psi$. Moreover, T^* works in $NTIME(O(f)) \subseteq NTIME(f) \subseteq \Sigma^p_m \cap \Pi^p_m$, so there is a Σ^1_m formula $\varphi^*_\Sigma(x)$ — equivalent to a Π^1_m formula $\varphi^*_\Pi(x)$ — which is true of w in $\mathbf{M}$ iff T^* accepts $(\mathbf{M}, w)$.

Let $Sat^+_\Sigma(x)$ be

$$\exists w(\text{“}w = \ulcorner x^\frown z\urcorner \text{ where } z \text{ is a string of zeroes”}$$

$$\& \ \varphi_h(lh(x), lh(w)) \ \& \ \varphi^*_\Sigma(w)),$$

and let $Sat^+_\Pi(x))$ be defined analogously using φ^*_Π. It is not hard to see that $Sat^+_\Sigma(x)$ and $Sat^+_\Pi(x)$ are FM–truth definitions for Σ^1_1 over σ. It therefore remains to check that given ψ, we can compute $bound(\psi)$ such that $Sat^+_\Sigma(x)$ (or, equivalently, $Sat^+_\Pi(x)$) works properly for ψ in all models larger than $bound(\psi)$.

Before we describe the algorithm, we note that checking whether $\mathbf{M} \models (\varphi_h(lh(\psi), i + lh(\psi))$ on input $(\mathbf{M}, \ulcorner\psi^\frown 0^i\urcorner)$ requires (deterministic) time M^l for some fixed l independent of ψ. By assumption, f is superpolynomial, so there is a c such that for all $n \geq c$, $f(n) \geq n^l$.

Now compute $bound(\psi)$ on input ψ as follows. Find the smallest model $\mathbf{M}'$ in which there is a w such that $\mathbf{M}' \models \chi(w)$, where $\chi(w)$ is

$$\text{“}w = \ulcorner\psi^\frown z\urcorner \text{ where } z \text{ is a string of zeroes” } \& \ \varphi_h(lh(\psi), lh(w)).$$

Since we may assume that both the conjuncts in χ are decent, this w will be equal to $\ulcorner\psi^\frown 0^{h(lh(\psi))-lh(\psi)}\urcorner$ and will also satisfy χ in all models larger than $\mathbf{M}'$. Let $bound(\psi) := max(M', c)$.

Given any $(\mathbf{M}, w)$ with $M \geq bound(\psi)$, the machine T^* will recognize that w is of the proper length in time $M^l \leq f(M)$, and check whether $\tilde{T}$ accepts $(\mathbf{M}, w)$ in time $f(M)$. Thus, T^* will not exceed its time limit $2f(M)$, so it will correctly answer whether $\mathbf{M} \models \psi$, which proves that $bound(\psi)$ is indeed large enough. $\square$

We need one more observation: two important constructions associated with the logic $\mathfrak{L}_r$ are effective and can therefore be FM–represented by decent formulae.

Lemma 3.3. *There exist computable functions:*

(a) neg, which assigns to (the Gödel number of) an $\mathfrak{L}_r$–sentence ξ (the Gödel number of) the "negation" of ψ, i.e. an $\mathfrak{L}_r$–sentence $neg(\xi)$ which is true in exactly those models in which ξ is false;

(b) snf, which assigns to (the Gödel number of) an $\mathfrak{L}_r$–sentence ξ (the Gödel number of) a Σ^1_2 sentence $snf(\xi)$ in SNF_r such that ξ and $snf(\xi)$ are equivalent.

Proof. For (a), let $\xi = \langle T_1, T_2 \rangle$. Then $neg(\ulcorner \xi \urcorner)$ is $\ulcorner \langle \overline{T_1}, T_2 \rangle \urcorner$, where $\overline{T_1}$ is T_1 with accepting and rejecting states interchanged.

For (b), we just need to make sure that the construction in the proof of lemma 2.1 can be carried out in an effective way. We sketch the argument. Given $\ulcorner \xi \urcorner = \ulcorner \langle T_1, T_2 \rangle \urcorner$, we may effectively construct NP–machines $T_{\mathcal{K}_1}$, $T_{\mathcal{K}_2}$ which recognize $\mathcal{K}_1$ and the complement of $\mathcal{K}_2$, respectively. Moreover, we can compute a time bound for both of these machines: if T_1 has an n^{k_1} clock and T_2 has an n^{k_2} clock, then $T_{\mathcal{K}_1}$ essentially runs as T_1 except that it will have to simulate T_2 at most n^r times (the exact number depends on the additional input relation R) on strings of length $\leq n^{k_1}$ (since T_1 will not have the time to write any longer strings on its oracle tape). So a rough estimate of the time needed by $T_{\mathcal{K}_1}$ is $n^{k_1} + n^{r+k_1 \cdot k_2}$. Some additional time is needed for looking at the bits of R, so we may take $n^{r+k_1 \cdot k_2+1}$ as a suitable upper bound. The same bound will do for $T_{\mathcal{K}_2}$.

Once $T_{\mathcal{K}_1}$, $T_{\mathcal{K}_2}$, and their time bounds are known, we can use them to compute the Σ^1_1 sentences φ, ψ which describe the action of $T_{\mathcal{K}_1}$ and $T_{\mathcal{K}_2}$, respectively. We then set $\gamma := \exists R^r(\varphi(R) \& \neg\psi(R))$ and $snf(\ulcorner \xi \urcorner) := \ulcorner \gamma \urcorner$. $\square$

By proposition 3.1, there exist decent formulae which represent the graphs of neg and snf. Choose some such formulae and call them $neg(x, y)$ and $snf(x, y)$, respectively.

We are now ready to give a proof of (B). Fix m and assume that the hypothesis of (B) holds for m. Fix r and σ. Let $Sat^+_\Sigma(x)$ and $Sat^+_\Pi(x)$ be the equivalent Σ^1_m and Π^1_m FM–truth definitions for Σ^1_1 over $(\sigma + R^r)$ given by lemma 3.2. Let *bound* be the function from that lemma appropriate for Sat^+_Σ and Sat^+_Π.

Recall that *bound* is a computable function. Thus, by a minor modification of the proof of proposition 3.1, there is a first order σ_0–formula $large(x)$ which satisfies the following three requirements for any given Σ^1_1 sentence ψ over $(\sigma + R^r)$:

- $\mathbf{M} \models large(\ulcorner\psi\urcorner)$ implies $M \geq bound(\psi)$;
- $\mathbf{M} \models large(\ulcorner\psi\urcorner)$ for all sufficiently large $\mathbf{M}$;
- if $\mathbf{M} \models large(\ulcorner\psi\urcorner)$ and $M' \geq M$, then $\mathbf{M}' \models large(\ulcorner\psi\urcorner)$.

Consider now the formula $good(x, x', y, y', \ulcorner\varphi\urcorner, \ulcorner\psi\urcorner, \ulcorner\varphi'\urcorner, \ulcorner\psi'\urcorner)$:

$$\begin{gathered} neg(x, x') \mathbin{\&} snf(x, y) \mathbin{\&} snf(x', y') \\ \mathbin{\&} y = \ulcorner\exists R^r(\varphi(R) \& \neg\psi(R))\urcorner \mathbin{\&} y' = \ulcorner\exists R^r(\varphi'(R) \& \neg\psi'(R))\urcorner \\ \mathbin{\&} large(\ulcorner\varphi\urcorner) \mathbin{\&} large(\ulcorner\psi\urcorner) \mathbin{\&} large(\ulcorner\varphi'\urcorner) \mathbin{\&} large(\ulcorner\psi'\urcorner). \end{gathered}$$

Let $Tr_\Sigma(x)$ be:

$$\exists x' \ldots \exists \ulcorner\psi'\urcorner(good(x, x', \ldots, \ulcorner\psi'\urcorner) \mathbin{\&} \exists R^r(Sat^+_\Sigma(\ulcorner\varphi\urcorner) \& \neg Sat^+_\Pi(\ulcorner\psi\urcorner))),$$

and let $Tr_\Pi(x)$ be:

$$\exists x \ldots \exists \ulcorner\psi'\urcorner(good(x, x', \ldots, \ulcorner\psi'\urcorner) \mathbin{\&} \neg\exists R^r(Sat^+_\Sigma(\ulcorner\varphi'\urcorner) \& \neg Sat^+_\Pi(\ulcorner\psi'\urcorner))).$$

It is not hard to see that Tr_Σ is equivalent to a Σ^1_m formula, that Tr_Π is equivalent to a Π^1_1 formula, and that both Tr_Σ and Tr_Π are FM–truth definitions for $\mathfrak{L}_r$ over σ. Let us then check that Tr_Σ and Tr_Π are equivalent.

It may be assumed that neither $Tr_\Sigma(w)$ nor $Tr_\Pi(w)$ ever holds if w is not the Gödel number of an $\mathfrak{L}_r$–sentence. So let $\mathbf{M}$ be a σ–model and let $\ulcorner\xi\urcorner \in M$ for some $\mathfrak{L}_r$–sentence ξ. If there is no tuple $\langle x', \ldots, \ulcorner\psi'\urcorner\rangle$ such that $\mathbf{M} \models good(\ulcorner\xi\urcorner, x', \ldots, \ulcorner\psi'\urcorner)$, then we have $\mathbf{M} \not\models Tr_\Sigma(\ulcorner\xi\urcorner)$ and $\mathbf{M} \not\models Tr_\Pi(\ulcorner\xi\urcorner)$.

Otherwise, by the decency of the formulae neg and snf, it must be the case that $x' = neg(\ulcorner\xi\urcorner)$, that $y = snf(\ulcorner\xi\urcorner) = \ulcorner\exists R^r(\varphi(R) \& \neg\psi(R))\urcorner$, and that $y' = snf(neg(\ulcorner\xi\urcorner)) = \ulcorner\exists R^r(\varphi'(R)\&\neg\psi'(R))\urcorner$. Moreover, by the choice of the formula $large$, M is large enough for Sat^+_Σ and Sat^+_Π to correctly recognize, given any $R \subseteq M^r$, the truth value of φ, ψ, φ', and ψ' in $(\mathbf{M}, R)$. Thus, $\mathbf{M} \models Tr_\Sigma(\ulcorner\xi\urcorner)$ iff $\mathbf{M} \models \exists R^r(\varphi(R)\&\neg\psi(R))$, that is, iff $\mathbf{M} \models \xi$. Similarly, $\mathbf{M} \models Tr_\Pi(\ulcorner\xi\urcorner)$ iff $\mathbf{M} \not\models \exists R^r(\varphi'(R)\&\neg\psi'(R))$, that is, again, iff $\mathbf{M} \models \xi$. So, also in this case $\mathbf{M} \models Tr_\Sigma(\ulcorner\xi\urcorner)$ iff $\mathbf{M} \models Tr_\Pi(\ulcorner\xi\urcorner)$, which proves that Tr_Σ and Tr_Π are indeed equivalent.

Hence, for any choice of r and σ, we have found a Δ^1_m FM–truth definition for $\mathfrak{L}_r$ over σ. By the finite version of Tarski's theorem, for any r and over any σ which contains an individual constant, it holds that $\mathfrak{L}_r$ is strictly less expressive that Δ^1_m, so $P^{NP[n^r]}$ is properly contained in $\Sigma^p_m \cap \Pi^p_m$. It follows that $P^{NP[n^r]} \subsetneq \Sigma^p_m \cap \Pi^p_m$ as classes of languages over $\{0,1\}$.

4 Concluding remarks

I. There are no obstacles to extending the methods of sections 2 and 3 to $P^{\Sigma^p_i[n^r]}$ instead of just $P^{NP[n^r]}$. Thus, for any i, we obtain the following result:

Theorem 4.1. *Let $m \geq i+1$. Then:*

(A) If there exists k such that $\Sigma^p_i \subseteq \Sigma_m TIME(n^k) \cap \Pi_m TIME(n^k)$, then for every r there exists k_r such that $P^{\Sigma^p_i[n^r]} \subseteq \Sigma_m TIME(n^{k_r}) \cap \Pi_m TIME(n^{k_r})$;

(B) If there exists a superpolynomial time–constructible function f such that $\Sigma_i TIME(f) \subseteq \Sigma^p_m \cap \Pi^p_m$, then additionally $P^{\Sigma^p_i[n^r]} \subsetneq \Sigma^p_m \cap \Pi^p_m$.

II. Recall that $LOGSPACE^{NP}$ is $P^{NP[O(\log n)]}$ ([W90],[BH91]). Thus, for example, if there exists k such that $NP \subseteq \Sigma_m TIME(n^k) \cap \Pi_m TIME(n^k)$, then there also is some k' such that $LOGSPACE^{NP} \subseteq \Sigma_m TIME(n^{k'}) \cap \Pi_m TIME(n^{k'})$. Obviously, appropriate analogues of part (B) of theorem 0.1, and of theorem 4.1, also hold.

III. Define $(\Sigma^1_1)^{\leq r}$ as the subclass of Σ^1_1 consisting of those formulae in which the existential second order quanitifiers have arity at most r. The hierarchy $\langle(\Sigma^1_1)^{\leq r}\rangle_{r\in\omega}$ is sometimes referred to as the Σ^1_1 arity hierarchy. This hierarchy is known to be strict if we take into account vocabularies of arbitrary arity ([A83]), but its strictness over a uniform vocabulary remains an open problem.

Observe that if model checking for FO over any fixed vocabulary is contained in $NTIME(n^k)$ — the number k may depend on σ — then there is a Σ^1_1 FM–truth definition for FO over any vocabulary σ. It is a routine task to transform these definitions into Σ^1_1 FM–truth definitions for $(\Sigma^1_1)^{\leq r}$, for any r and over any σ. We therefore have:

Proposition 4.2. *If for any σ there is k such that model checking for FO over σ is in $NTIME(n^k)$, then for any σ and r there is k_r such that model checking for $(\Sigma^1_1)^{\leq r}$ over σ is in $NTIME(n^{k_r})$ (and thus the Σ^1_1 arity hierarchy is infinite — hence, by [F75], strict — over any σ).*

Note that the hypothesis of the proposition is satisfied if $LOGSPACE \subseteq NTIME(n^k)$ for some k, in particular if NP contains $DSPACE(f)$ for any space–constructible function f which dominates log.

Again, this can be generalized to arity hierarchies of the higher prenex classes of SO. Let $(\Sigma^1_m)^{\leq r}$ denote the subclass of Σ^1_m consisting of formulae in which the second order quantifiers in the initial existential quantifier block have arity at most r (the arity of the other relational quantifiers is arbitrary). We have:

Proposition 4.3. *If there is k such that $\Pi^p_{m-1} \subseteq \Sigma_m TIME(n^k)$, then for σ and r there is k_r such that model checking for $(\Sigma^1_m)^{\leq r}$ over σ is in $NTIME(n^{k_r})$.*

Note again that the hypothesis will hold e.g. if Σ^p_m contains $\Pi_{m-1}TIME(f)$ for a time–constructible superpolynomial f.

IV. As pointed out in the introduction, our theorem 0.1 has as its consequence the (already known) result that for any r, $P^{NP[n^r]} \subsetneq NEXPTIME$. It is perhaps worth noting that the largest syntactically defined fragments of SO which can be separated from $NEXPTIME$ by a similar argument are the closures of Σ^1_1 under second order quantifiers of bounded arity. For, let $SO^{\leq r}(\Sigma^1_1)$ denote the closure of Σ^1_1 under boolean connectives, first order quantification, and second order quantification over relations of arity at most r. Then we easily get:

Proposition 4.4. *If $NP \subseteq DSPACE(n^k)$ for some k, then for any σ there exists k' such that all $SO^{\leq r}(\Sigma^1_1)$–definable classes of σ–models are in $DSPACE(n^{k'})$,*

from which it follows that $SO^{\leq r}(\Sigma^1_1)$ cannot define all $NEXPTIME$ classes of models.

V. We proved part (B) of theorem 0.1 by showing that if the hypothesis of this part holds, then there is a Δ^1_m FM–truth definition for the logic $\mathfrak{L}_r$ (over any vocabulary). Compared to the construction of Σ^1_m and Π^1_m FM–truth definitions in the proof of part (A), the construction of a Δ^1_m definition required some more subtlety — and an additional assumption.

This suggests the following, possibly interesting, question: given a vocabulary σ, is there any natural characterization of the logics for which Δ^1_m can define FM–truth over σ? Clearly, for any such logic $\mathcal{L}$ there must exist some k such that model checking for $\mathcal{L}$ over σ is contained in $\Sigma_m TIME(n^k) \cap \Pi_m TIME(n^k)$. However, there appears to be no good reason to suspect that this is sufficient, as the Σ^1_m and Π^1_m FM–truth definitions whose existence follows by theorem 1.5 from the existence of such a number k need not be equivalent.

References

[A83] M. Ajtai, Σ^1_1 *Sentences on Finite Structures*, in **Annals of Pure and Applied Logic** 24(1983), pp. 1–48.

[BFFT01] H. Buhrman, S. Fenner, L. Fortnow and L. Torenvliet, *Two Oracles that Force a Big Crunch*, in **Computational Complexity** 10(2001), pp. 93–116.

[BH91] S. R. Buss and L. Hay, *On Truth-Table Reducibility to SAT*, in **Information and Computation** 91(1991), pp. 86–102.

[BT94] H. Buhrman and L. Torenvliet, *On the Cutting Edge of Relativization: The Resource Bounded Injury Method*, in **Proc. International Colloquium on Automata, Languages and Programming 1994**, Springer–Verlag 1994, pp. 263–273.

[F74] R. Fagin, *Generalized First-Order Spectra and Polynomial-Time Recognizable Sets*, in **Complexity of Computation, SIAM-AMS Proceedings** 7(1974), pp. 43–73.

[F75] R. Fagin, *A Spectrum Hierarchy*, in **Zeitschrift für Mathematische Logik und Grundlagen der Mathematik** 21(1975), pp. 123–134.

[FLZ92] B. Fu, H. Li, and Y. Zhong, *Some Properties of Exponential Time Complexity Classes*, in **Proc. Structure in Complexity Theory Conference 1992**, IEEE Computer Society Press 1992, pp. 50–57.

[G97] G. Gottlob, *Relativized Logspace and Generalized Quantifiers over Finite Ordered Structures*, **Journal of Symbolic Logic** 62(1997), pp. 545–574.

[HP93] P. Hájek and P. Pudlák, **Metemathematics of First–Order Arithmetic**, Springer–Verlag 1993.

[K04] L. A. Kołodziejczyk, *Truth Definitions in Finite Models*, in **Journal of Symbolic Logic** 69(2004), pp. 183–200.

[M96] S. Mocas, *Separating Classes in the Exponential-Time Hierarchy From Classes in PH*, in **Theoretical Computer Science** 158(1996), pp. 221–231.

[MM01] M. Mostowski, *On Representing Concepts in Finite Models*, in **Mathematical Logic Quarterly** 47(2001), pp. 513–523.

[Ste93] I. A. Stewart, *Logical Characterization of Bounded Query Classes I: Logspace Oracle Machines*, in **Fundamenta Informaticae** 18(1993), pp. 65–92.

[Sto77] L. Stockmeyer, *The Polynomial-Time Hierarchy*, in **Theoretical Computer Science** 3(1977), pp. 1–22.

[W90] K. W. Wagner, *Bounded Query Classes*, in **SIAM Journal on Computing** 19(1990), pp. 833–846.