RANK HIERARCHIES FOR GENERALIZED QUANTIFIERS

H. JEROME KEISLER AND WAFIK BOULOS LOTFALLAH

Abstract. We show that for each n and m, there is an existential first order sentence which is NOT logically equivalent to a sentence of quantifier rank at most m in infinitary logic augmented with all generalized quantifiers of arity at most n. We use this to show the strictness of the quantifier rank hierarchies for various logics ranging from existential (or universal) fragments of first order logic to infinitary logics augmented with arbitrary classes of generalized quantifiers of bounded arity.

The sentence above is also shown to be equivalent to a first order sentence with at most $n+2$ variables (free and bound). This gives the strictness of the quantifier rank hierarchies for various logics with only $n+2$ variables. The proofs use the bijective Ehrenfeucht-Fraisse game and a modification of the building blocks of Hella.

1. Introduction

In the context of finite model theory and descriptive complexity, first order logic FO is known to have severe limitations. On one hand it cannot count, because very simple linear time queries such as parity are not first order definable. On the other hand, it lacks recursion, because it cannot express some simple nondeterministic logarithmic space queries such as reachability.

To overcome the first limitation, we can augment first order logic with the ability to count. If S is a set of natural numbers, the quantifier Q_S can be added to first order logic, with the following semantics:

$$\mathcal{A} \models (Q_S\, x)\varphi(x) \text{ iff } |\{a \in A : \mathcal{A} \models \varphi[a]\}| \in S.$$

The property of odd parity is simply expressed by $(Q_S\, x)(x = x)$, where S is the set of odd natural numbers.

To express a recursive query like reachability, we can augment first order logic with the transitive closure operator TC with the following semantics[1]:

> $\mathcal{A} \models (TC\ x_1, x_2, y, z)(\varphi(x_1, x_2), \alpha(y), \beta(z))$ iff for each pair (a, b) in A, if $(\mathcal{A}, a) \models \alpha[a]$ and $(\mathcal{A}, b) \models \beta[b]$, then (a, b) belongs to the transitive closure of the relation on A defined by $\varphi(x_1, x_2)$.

Reachability over the signature $\{E, s, t\}$ is expressed by the sentence

$$(TC\ x_1, x_2, y, z)(E(x_1, x_2), y = s, z = t).$$

Date: May 22, 2010.

[1]Usually Transitive Closure is defined by $[(TC\ x_1, x_2)\varphi(x_1, x_2)](y', z')$, which is the special case of the definition here, with $\alpha(y, y')$ being $y = y'$ and $\beta(z, z')$ being $z = z'$.

These different ways of extending first order logic are two instances of the generalized quantifiers introduced by Lindström in [Lin66]. A generalized quantifier Q is simply a class of (finite) models with a fixed finite signature ν_Q that is closed under isomorphism, and acts like an oracle in the interpretation of a formula. Q is n-ary, or has arity n, if n is the maximum number of arguments in the relations in ν_Q. In general, we can add a class $\mathbf{Q}$ of generalized quantifiers to first order logic to produce the logic $FO(\mathbf{Q})$ (defined in Section 2 below).

It should be noted here that if $\mathbf{Q}$ is the class of all generalized quantifiers, then any class of Boolean queries on finite models can be defined by a formula in $FO(\mathbf{Q})$. (Non-Boolean queries can also be expressed by turning them into Boolean queries over wider signatures, similar to the definition of TC above.) One would like to capture interesting classes of queries by augmenting FO with finitely many generalized quantifiers. However, this is not always possible. Hella proved in [Hel96] that if the class $\mathbf{Q}$ has bounded arity, then $FO(\mathbf{Q})$ cannot capture the language $PTIME$, or even the weaker language $DATALOG$. This suggests that the class of all generalized quantifiers is too large, but one should consider infinite class of generalized quantifiers, and that the class $\mathbf{Q}_n$ of all generalized quantifiers of arity at most n is interesting.

In this paper, $\mathbf{Q}$ *will always denote a class of generalized quantifiers of bounded arity, so that* $\mathbf{Q} \subseteq \mathbf{Q}_n$ *for some* n*.*

To capture different kinds of recursion, Moschovakis proposed in [Mos74] the addition of inductive operators to first order logic. This has been described as the most successful logic in computer science. Indeed, in the presence of order, Vardi [Var82] and Immerman [Imm86] showed that $FO(LFP)$ and $FO(PFP)$ (first order logic augmented with the least- and partial- fixed point operator) capture $PTIME$ and $PSPACE$, respectively.

The study of logics with inductive operators was greatly simplified when Kolaitis and Vardi noted in [KV92] that these logics are mere fragments of the logic $\mathcal{L}^{\omega}_{\infty\,\omega}$, whose formulas have infinite conjunctions and disjunctions but finitely many free and bound variables. $\mathcal{L}^{\omega}_{\infty\,\omega}$ was introduced by Barwise [Bar77] in the context of infinite models (see [EF99], [Imm99]).

The infinitary logic $\mathcal{L}_{\infty\,\omega}$ contains FO and allows infinite conjunction and disjunction, which enables inductive constructions as many times as needed. $\mathcal{L}_{\infty\,\omega}$, however, is too powerful for finite model theory, as it can express all possible queries over finite models. Thus, going backwards, we should restrict $\mathcal{L}_{\infty\,\omega}$ (or its extension $\mathcal{L}_{\infty\,\omega}(\mathbf{Q})$) by limiting either the quantifier rank m of the formulas, or the total number k of variables (free and bound) occurring in the formulas. Note that, in descriptive complexity theory, the quantifier rank may be considered as a kind of time resource, while the number of variables may be considered as a kind of time resource. This leads to the logics $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, m)$ and $\mathcal{L}^{k}_{\infty\,\omega}(\mathbf{Q})$, and their respective unions $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, \omega)$ and $\mathcal{L}^{\omega}_{\infty\,\omega}(\mathbf{Q})$. The following natural questions arise:

> Do the quantifier rank hierarchies of $\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)$ and $FO(\mathbf{Q}, m)$ collapse at some finite level?

In this paper we will give a negative answer to these questions when the set of quantifiers $\mathbf{Q}$ has bounded arity and contains at least the existential quantifier $\exists$. In particular, we show that for each m, $\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)$ is strictly contained in $\mathcal{L}_{\infty\omega}(\mathbf{Q}, m+1)$, and $FO(\mathbf{Q}, m)$ is strictly contained in $FO(\mathbf{Q}, m+1)$. We also show the strictness of the quantifier rank hierarchies with one or both of the following restrictions:

(1) The set of all variables occurring in formulas has cardinality at most k, where $k \geq n+2$ is fixed. This leads to the k-variable logics $\mathcal{L}^{k}_{\infty\omega}(\mathbf{Q}, m)$.
(2) Negation is allowed only on quantifier-free formulas. This leads to the quantifier-positive logics $\mathcal{L}^{+}_{\infty\omega}(\mathbf{Q}, m)$.

These results complement different hierarchy strictness results in the literature. In [Hel89] Hella showed the strictness of the hierarchy based on the maximum arity n of the quantifiers in $\mathbf{Q}$. Hella used an Ehrenfeucht-Fraïssé like game (called the bijective game) that characterizes the power of the logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$. Later on, he together with Luosto and Väänänen gave a cardinality argument in [HLV96] showing the strictness of the hierarchy based on the similarity type of the quantifier, which contains both the quantifier arity and width (= the maximum number of formulas bound by the quantifier). However, their result was not explicit, that is, it did not lead to specific queries in the levels of the hierarchies. In the unary case, Luosto gave in [Luo00] an explicit strict hierarchy result based on the width of unary quantifiers.

Our results here are "almost explicit". Fix a maximum arity n. We say that a sentence φ is **expressible** in a class of sentences $\mathcal{L}$ if φ is logically equivalent to some sentence $\psi \in \mathcal{L}$. For each m we give an existential first order sentence σ_m that is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$. We then use the sentences σ_m to show that the hierarchies are strict. To prove that σ_m is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$, we give a recursive construction of particular models $\mathcal{A}_m$ and $\mathcal{B}_m$ that disagree on σ_m but agree on each sentence of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$. The construction of these models uses building blocks that are similar to those of Hella in [Hel96].

After giving the basic definitions, including Hella's bijective game in Section 2, we introduce our recursive building strategy of the models $\mathcal{A}_m$ and $\mathcal{B}_m$ in Section 3. In Section 4, we introduce the operation $Array_n(\theta)$ defined on sentences θ. This leads to Section 5, where the main hierarchy results are obtained. We define a recursive sequence ρ_m of first order sentences such that ρ_m is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$, and use these sentences to show that the quantifier rank hierarchies of $FO(\mathbf{Q})$ and $\mathcal{L}_{\infty\omega}(\mathbf{Q})$ are strict at all finite levels. In Section 6 we strengthen the results by getting a version σ_m of ρ_m which has just $n+2$ variables. The sentence σ_m is an existential first order sentence, and is equivalent to a purely existential first order sentence

$\sigma'_m = (\exists y_1) \cdots (\exists y_\ell)\varphi(y_1, \ldots, y_\ell)$ where φ is quantifier-free and the number of variables is $\ell = (n+1)m + 1$. We then show that the quantifier rank hierarchies of $FO^k(\mathbf{Q})$ and $\mathcal{L}^k_{\infty\omega}(\mathbf{Q})$ are strict at all finite levels. In Section 7 we get analogous results for the logics $FO^+(\mathbf{Q})$ and $\mathcal{L}^+_{\infty\omega}(\mathbf{Q})$. Finally, In Section 8 we show that the logics $\mathcal{L}^k_{\infty\omega}(\mathbf{Q})$ and $\mathcal{L}^+_{\infty\omega}(\mathbf{Q})$ have sentences that are not equivalent to any sentences of finite quantifier rank.

2. Basic Definitions

We always consider classes of finite models with a fixed **underlying signature** ν, which is a set of relation symbols. Given two formulas φ and ψ, $\varphi \models \psi$ will mean that every finite model of φ is a model of ψ. Two formulas φ, ψ will be called **equivalent** if $\models \varphi \leftrightarrow \psi$. Two sets of formulas will be called equivalent if every formula in one set is equivalent to some formula in the other set, and vice versa.

A generalized quantifier Q is simply a class of finite models over a finite signature $\nu_Q = \{R_1, \ldots, R_k\}$ that is closed under isomorphisms. The **arity** of Q is the maximum number of arguments of the relation symbols in ν_Q, and the **width** of Q is the length k of the signature ν_Q. The symbol Q can be used to build a formula of the form $\varphi(\mathbf{x}) = (Q\,\mathbf{y}_1, \ldots, \mathbf{y}_k)(\psi_1(\mathbf{x}, \mathbf{y}_1), \ldots, \psi_k(\mathbf{x}, \mathbf{y}_k))$, where all the variables in the variable strings $\mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_k$ are distinct, and each $\psi_i(\mathbf{x}, \mathbf{y}_i)$ has its free variables included in $\mathbf{x}, \mathbf{y}_i$. $\varphi(\mathbf{x})$ has its free variables included in $\mathbf{x}$, and is given the following semantics in a finite structure $\mathcal{A}$ with a tuple $\mathbf{a}$ interpreting the free variable tuple $\mathbf{x}$:

> $(\mathcal{A}, \mathbf{a}) \models \varphi(\mathbf{x})$ iff the ν_Q-model over the universe of $\mathcal{A}$ with each R_i interpreted by the relation $\{\mathbf{b} : (\mathcal{A}, \mathbf{a}) \models \psi_i(\mathbf{a}, \mathbf{b})\}$ is in the class Q.

We let $\mathbf{Q}_n$ denote the class of all quantifiers of arity at most n. Thus, the universal and existential quantifiers $\forall, \exists$, as well as the counting quantifiers $\exists^{\geq k}$ all belong to $\mathbf{Q}_1$. Note that different quantifiers $Q \in \mathbf{Q}$ can have different signatures ν_Q.

For any class of quantifiers $\mathbf{Q} \subseteq \mathbf{Q}_n$ and underlying signature ν, we let $\mathcal{L}_{\infty\omega}(\mathbf{Q})$ be the class of all formulas with finitely many (possibly zero) free variables built from atomic formulas over ν using negations, infinite conjunctions and disjunctions, as well as generalized quantifiers in $\mathbf{Q}$. We take the classical infinitary logic to be the logic $\mathcal{L}_{\infty\omega} = \mathcal{L}_{\infty\omega}(\{\exists\})$ with just the existential quantifier. (The logics $\mathcal{L}_{\infty\omega}(\{\exists\}), \mathcal{L}_{\infty\omega}(\{\forall\})$, and $\mathcal{L}_{\infty\omega}(\{\exists, \forall\})$ are equivalent, but have different quantifier-positive fragments). We will only consider classes $\mathbf{Q}$ of bounded arity, that is, $\mathbf{Q} \subseteq \mathbf{Q}_n$ for some n.

It is well known that the logic $\mathcal{L}_{\infty\omega}$ is too powerful for finite model theory, as it can express any class of finite models closed under isomorphism. One way to weaken this logic, and the logic $\mathcal{L}_{\infty\omega}(\mathbf{Q})$ in general, is to restrict the quantifier rank of its formulas. The **quantifier rank** of a formula φ in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n)$ is defined by induction on the complexity of φ as in [Hel89]:

- A quantifier-free formula has quantifier rank 0;
- The quantifier rank of $\neg\varphi$ is equal to the quantifier rank of φ;

- The quantifier rank of $\bigwedge_{i\in I}\varphi_i$, and also of $\bigvee_{i\in I}\varphi_i$, is the supremum of the quantifier ranks of $\varphi_i, i \in I$;
- the quantifier rank of $(Q\,\mathbf{y}_1,\ldots,\mathbf{y}_k)(\psi_1(\mathbf{x},\mathbf{y}_1),\ldots,\psi_k(\mathbf{x},\mathbf{y}_k))$ is 1 plus the maximum of the quantifier ranks of $\psi_1,\ldots,\psi_k$.

For a class $\mathbf{Q} \subseteq \mathbf{Q}_n$, we let $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, m)$ be the set of all formulas of $\mathcal{L}_{\infty\,\omega}(\mathbf{Q})$ of quantifier rank at most m. Also, let $\mathcal{L}_{\infty\,\omega}(\mathbf{Q},\omega) = \bigcup_m \mathcal{L}_{\infty\,\omega}(\mathbf{Q}, m)$. Note that for any $\mathbf{Q}$, $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, 0)$ is the set of quantifier-free formulas of $\mathcal{L}_{\infty\,\omega}$.

We define $FO(\mathbf{Q})$ to be the first order part of $\mathcal{L}_{\infty\,\omega}(\mathbf{Q})$, i.e. the set of all formulas of $\mathcal{L}_{\infty\,\omega}(\mathbf{Q})$ of finite length. Also, $FO(\mathbf{Q}, m)$ is the set of all formulas of $FO(\mathbf{Q})$ of quantifier rank at most m. Since all formulas in $FO(\mathbf{Q})$ have finite quantifier rank, we have $FO(\mathbf{Q}) = \bigcup_m FO(\mathbf{Q}, m)$. We take the classical first order logic to be $FO = FO(\{\exists\})$. Note that for any $\mathbf{Q}$, $FO(\mathbf{Q}, 0)$ is the set of quantifier-free first order formulas. And if the underlying signature is finite, $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, 0) = FO(\mathbf{Q}, 0)$.

The following proposition shows that, unless $\mathbf{Q}$ is infinite, every formula in the restricted logic $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, m)$ is equivalent to a formula in $FO(\mathbf{Q}, m)$.

Proposition 2.1. *If both the underlying signature and the class $\mathbf{Q}$ are finite, then for each m:*

(i) *For each finite set of variables $\mathbf{x}$, the set of formulas in $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, m)$ with at most $\mathbf{x}$ free is equivalent to a finite set of formulas of $FO(\mathbf{Q}, m)$.*

(ii) *$\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, m)$ is equivalent to $FO(\mathbf{Q}, m)$.*

Proof: Part (i) is proved by induction on m. Part (ii) then follows. ■ *2.1*

The logics $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}_n, m)$ are characterized by the n-**bijective** Ehrenfeucht-Fraïssé game introduced in [Hel89]. This game is played on two models by two players, Spoiler and Duplicator. Roughly speaking, Spoiler tries to prove that the two models look different, while Duplicator tries to prove that they look alike. By a game **position** we will mean a triple $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), m)$ where $|\mathbf{a}| = |\mathbf{b}|$, and m is a natural number that represents the number of moves yet to be played. The game is defined by recursion on m as follows.

When $m > 0$, the game proceeds from the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), m)$ according to the rules:

(1) Duplicator chooses a bijection $f : \mathcal{A} \longrightarrow \mathcal{B}$. (If $|\mathcal{A}| \neq |\mathcal{B}|$, then Spoiler wins.)
(2) Spoiler chooses an n-tuple $\mathbf{c}$ in $\mathcal{A}$.
(3) The game continues from the new position

$$((\mathcal{A}, \mathbf{ac}), (\mathcal{B}, \mathbf{b}f(\mathbf{c})), m-1),$$

where $f(\mathbf{c}) = f(c_1,\ldots,c_n) = (f(c_1),\ldots,f(c_n))$.

When $m = 0$ the game ends. $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), 0)$ is a **winning position** for Duplicator iff $(\mathcal{A}, \mathbf{a})$ and $(\mathcal{B}, \mathbf{b})$ satisfy the same atomic formulas. (In other words, $\mathbf{a}$, $\mathbf{b}$ are either empty or generate submodels of $\mathcal{A}$, $\mathcal{B}$ that are isomorphic with an isomorphism mapping $\mathbf{a}$ to $\mathbf{b}$.)

We write $(\mathcal{A}, \mathbf{a}) \equiv_n^m (\mathcal{B}, \mathbf{b})$ if Duplicator has a winning strategy in the n-bijective game starting from the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), m)$, and we write $\mathcal{A} \equiv_n^m \mathcal{B}$ if Duplicator has a winning strategy starting from the position $(\mathcal{A}, \mathcal{B}, m)$. By the **game** $\mathcal{A} \equiv_n^m \mathcal{B}$ we will mean the n-bijective game starting from the position $(\mathcal{A}, \mathcal{B}, m)$.

The importance of this game stems from the following result, which is proved in [Hel89, Hel96].

Proposition 2.2. $(\mathcal{A}, \mathbf{a}) \equiv_n^m (\mathcal{B}, \mathbf{b})$ *iff* $(\mathcal{A}, \mathbf{a})$ *and* $(\mathcal{B}, \mathbf{b})$ *agree on all formulas in* $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$. ■ *2.2*

3. The Building Strategy

The following definitions will use a modification of the building block introduced in [Hel96]. We first introduce some notation.

Definition 3.1. *A signature* ν *is called* n***-adequate*** *if* ν *contains at least a binary relation symbol* E *and an* $(n+1)$*-ary relation symbol* R*.*

We assume throughout this section that $n > 0$ and the underlying signature ν is n-adequate. Let $\mathcal{A}$ and $\mathcal{B}$ be finite models with signature ν. Also let $X = C \cup D$, with $C = \{c_1, \ldots, c_{n+1}\}$ and $D = \{d_1, \ldots, d_{n+1}\}$, where all the c_i's and d_i's are distinct, i.e. $|X| = 2n + 2$. Let G be the set

$$G = X \cup (A \times C) \cup (B \times D).$$

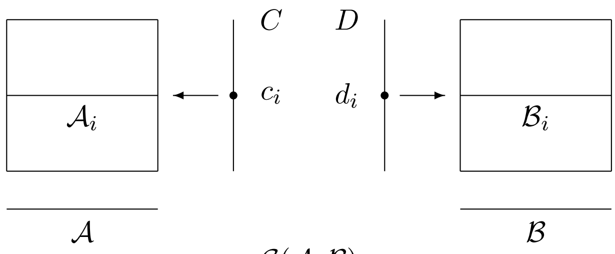
For each $i \in \{1, \ldots, n+1\}$, let

$$A_i = A \times \{c_i\}, \quad B_i = B \times \{d_i\}, \quad G_i = \{c_i, d_i\} \cup A_i \cup B_i.$$

Note that G_i is the set of elements of G that involve i, and the sets G_i form a partition of G. Let $\mathcal{A}_i$ be the copy of $\mathcal{A}$ with universe A_i such that $a \mapsto (a, c_i)$ is an isomorphism, and $\mathcal{B}_i$ be the copy of $\mathcal{B}$ with universe B_i such that $b \mapsto (b, d_i)$ is an isomorphism.

Definition 3.2. *Let* $\mathcal{G}(\mathcal{A}, \mathcal{B})$ *be the model with universe* G *and signature* ν *such that:*

(i) *For each* $i \in \{1, \ldots, n+1\}$*,* $\mathcal{A}_i$ *and* $\mathcal{B}_i$ *are submodels of* $\mathcal{G}(\mathcal{A}, \mathcal{B})$*.*
(ii) *For each* $i \in \{1, \ldots, n+1\}$*,* $a \in A_i$*, and* $b \in B_i$*,* $E(c_i, a)$ *and* $E(d_i, b)$ *hold in* $\mathcal{G}(\mathcal{A}, \mathcal{B})$*.*
(iii) *No atomic formulas hold in* $\mathcal{G}(\mathcal{A}, \mathcal{B})$ *except for equalities and those given in (1) and (2).*



$\mathcal{G}(\mathcal{A}, \mathcal{B})$

Intuitively, $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is the union of $n+1$ copies of $\mathcal{A}$ and $n+1$ copies of $\mathcal{B}$, which is also equipped with an "edge" relation $E(x, y)$, that connects each $c_i \in C$ to each element of the corresponding copy of $\mathcal{A}$, and similarly for D and $\mathcal{B}$. In the terminology of [JM01] and [KL04], c_i is the root of a cone over A_i, and d_i is the root of a cone over B_i, in the model $\mathcal{G}(\mathcal{A}, \mathcal{B})$.

With some extra work, it may be possible to dispense with the extra relation symbol E by coding with the $(n+1)$-ary relation R. However, this would make the proofs harder to follow, so we chose instead to include E in our notion of an adequate vocabulary.

Definition 3.3. *We now define two new models $\mathcal{G}^+(\mathcal{A}, \mathcal{B})$ and $\mathcal{G}^-(\mathcal{A}, \mathcal{B})$, which have the same universe and signature as $\mathcal{G}(\mathcal{A}, \mathcal{B})$. $\mathcal{G}^+(\mathcal{A}, \mathcal{B})$ is built from $\mathcal{G}(\mathcal{A}, \mathcal{B})$ by replacing R by $R \cup R^+$ where R^+ is the relation on X given by:*

$R^+(a_1, \ldots, a_{n+1})$ *iff* $|U \cap \{c_i, d_i\}| = 1$ *for each* i*, and* $|U \cap D|$ *is even,*

where $U = \{a_1, \ldots, a_{n+1}\}$. $\mathcal{G}^-(\mathcal{A}, \mathcal{B})$ is defined similarly but replacing R by $R \cup R^-$ where:

$R^-(a_1, \ldots, a_{n+1})$ *iff* $|U \cap \{c_i, d_i\}| = 1$ *for each* i*, and* $|U \cap D|$ *is odd.*

We define $\chi(x)$ to be the special formula $(\exists y)E(x, y) \wedge \neg(\exists y)E(y, x)$, which says that x is a source for the graph E. The next lemma is easily verified.

Lemma 3.4. *For every pair of models $\mathcal{A}, \mathcal{B}$, the set X is defined by $\chi(x)$ in both of the models $\mathcal{G}^+(\mathcal{A}, \mathcal{B})$ and $\mathcal{G}^-(\mathcal{A}, \mathcal{B})$.* ∎ *3.4*

The set X can also be defined by the simpler formula $\neg(\exists y)E(y, x)$, but later on it will be useful to know that X is defined by the particular formula $\chi(x)$.

For each $k \in \{1, \ldots, n+1\}$, define the bijection f_k on X by:

$$f_k(x) = \begin{cases} d_k & \text{if } x = c_k \\ c_k & \text{if } x = d_k \\ x & \text{otherwise} \end{cases},$$

i.e. f_k just swaps c_k with d_k.

The following lemma is easily proved.

Lemma 3.5. *The bijection f_k is an isomorphism from (X, R^+) to (X, R^-).* ■ *3.5*

We now prove our key lemma.

Lemma 3.6. *Suppose $\mathcal{A} \equiv_n^m \mathcal{B}$. Then*

$$\mathcal{G}^+(\mathcal{A}, \mathcal{B}) \equiv_n^{m+1} \mathcal{G}^-(\mathcal{A}, \mathcal{B}).$$

Proof: We write $\mathcal{G}^+ = \mathcal{G}^+(\mathcal{A}, \mathcal{B})$ and $\mathcal{G}^- = \mathcal{G}^-(\mathcal{A}, \mathcal{B})$. Duplicator uses the identity bijection $\iota : G \longrightarrow G$ for her first move. Spoiler selects n elements $e_1, \ldots, e_n \in G$. Since these elements belong to at most n of the sets G_i, there must exist $k \in \{1, \ldots, n+1\}$ such that G_k is disjoint from $\{e_1, \ldots, e_n\}$. Hold k fixed for the rest of the proof.

Suppose first that $m = 0$. In this case one can check directly that $(\mathcal{G}^+, e_1, \ldots, e_n)$ and $(\mathcal{G}^-, e_1, \ldots, e_n)$ satisfy the same atomic formulas, so Duplicator wins the game.

Now suppose $m > 0$. By hypothesis, $\mathcal{A}_k \equiv_n^m \mathcal{B}_k$. In the last m moves Duplicator plays the auxiliary game $\mathcal{A}_k \equiv_n^m \mathcal{B}_k$ and uses her winning bijections for that game to give her moves for the main game. Each of these moves will be a bijection of G that agrees with f_k on X, matches the auxiliary game on $A_k \cup B_k$, and is the identity function on the rest of G. Below is the detailed description of Duplicator's strategy.

Suppose that at some stage $0 < j \leq m$, the two games have the positions

$$((\mathcal{G}^+, \mathbf{s}), (\mathcal{G}^-, \mathbf{t}), j), \quad ((\mathcal{A}_k, \mathbf{u}), (\mathcal{B}_k, \mathbf{v}), j).$$

In the auxiliary game, let Duplicator's next winning move be h_j, which is a bijection from A_k to B_k. We note that since Duplicator is using a winning strategy in the auxiliary game, and equalities are atomic formulas, we must have $h_j(\mathbf{u}) = \mathbf{v}$.

Then in the main game, we define Duplicator's next move to be the bijection g_j on G such that:

$$g_j(x) = \begin{cases} f_k(x) & \text{if } x \in X \\ h_j(x) & \text{if } x \in A_k \\ h_j^{-1}(x) & \text{if } x \in B_k \\ x & \text{otherwise.} \end{cases}$$

Let Spoiler's next move in the main game be $\mathbf{r} = (r_1, \ldots, r_n)$, an n-tuple in G. Let a be an arbitrary element of A_k. In the auxiliary game, take Spoiler's next move to be $\mathbf{w} = (w_1, \ldots, w_n)$, an n-tuple in A_k, where

$$w_i = \begin{cases} r_i & \text{if } r_i \in A_k \\ g_j(r_i) & \text{if } r_i \in B_k \\ a & \text{otherwise.} \end{cases}$$

To show that Duplicator wins the main game, we must show that at the final position

$$((\mathcal{G}^+, \mathbf{s}), (\mathcal{G}^-, \mathbf{t}), 0),$$

$(\mathcal{G}^+, \mathbf{s})$ and $(\mathcal{G}^-, \mathbf{t})$ satisfy the same atomic formulas. We know that Duplicator wins the auxiliary game at the final position

$$((\mathcal{A}_k, \mathbf{u}), (\mathcal{B}_k, \mathbf{v}), 0).$$

That is, $(\mathcal{A}_k, \mathbf{u})$ and $(\mathcal{B}_k, \mathbf{v})$ satisfy the same atomic formulas. Moreover, Duplicator's last move h_1 is a bijection from A_k to B_k, and $h_1(\mathbf{u}) = \mathbf{v}$. It follows from the definition of g_1 that $g_1(\mathbf{s}) = \mathbf{t}$.

Consider an atomic formula $\alpha(\mathbf{s}')$ with parameters from $\mathbf{s}$ such that $\mathbf{s}'$ is the list of all parameters occurring in $\alpha(\mathbf{s}')$. Let $\mathbf{t}' = g_1(\mathbf{s}')$. We show that $\alpha(\mathbf{s}')$ holds in $\mathcal{G}^+$ iff $\alpha(\mathbf{t}')$ holds in $\mathcal{G}^-$. We suppose first that $\alpha(\mathbf{s}')$ holds in $\mathcal{G}^+$ and prove that $\alpha(\mathbf{t}')$ holds in $\mathcal{G}^-$. We must consider several cases.

Case 1: $\mathbf{s}'$ does not meet G_k $(= \{c_k, d_k\} \cup A_k \cup B_k)$. In this case, $\mathbf{t}' = \mathbf{s}'$ and $\mathcal{G}^-$ also satisfies $\alpha(\mathbf{s}')$.
Case 2: $\mathbf{s}' \subseteq A_k$. Then $\mathcal{A}_k$ satisfies $\alpha(\mathbf{s}')$ and $\mathbf{t}' = h_1(\mathbf{s}')$, so $\mathcal{B}_k$ satisfies $\alpha(\mathbf{t}')$ and hence $\mathcal{G}^-$ satisfies $\alpha(\mathbf{t}')$.
Case 3: $\mathbf{s}' \subseteq B_k$. Similar to Case 2.
Case 4: $\mathbf{s}' \subseteq X$ and $\alpha(\mathbf{s}')$ is $R(\mathbf{s}')$. Then $R^+(\mathbf{s}')$ holds and $\mathbf{t}' = f_k(\mathbf{s}')$, so by Lemma 3.5, $R^-(\mathbf{t}')$ holds and $\mathcal{G}^-$ satisfies $R(\mathbf{t}')$.
Case 5: $\mathbf{s}' = (c_k, a)$ where $a \in A_k$. Then $\mathbf{t}' = (d_k, b)$ where $b = h_1(a)$, so $E(\mathbf{s}')$ holds in $\mathcal{G}^+$ and $E(\mathbf{t}')$ holds in $\mathcal{G}^-$.
Case 6: $\mathbf{s}' = (d_k, b)$ where $b \in B_k$. Similar to Case 5.

The other direction is proved by a similar argument. Therefore Duplicator wins the main game $\mathcal{G}^+ \equiv_n^{m+1} \mathcal{G}^-$. ■ *3.6*

4. The $Array_n$ Operation

In this section we will introduce the operation $Array_n$ on $\mathcal{L}_{\infty\omega}$, which can climb up one step in the generalized quantifier rank ladder. For a first order sentence θ that distinguishes between two models $\mathcal{A}$ and $\mathcal{B}$, $Array_n(\theta)$ will distinguish between the models $\mathcal{G}^+(\mathcal{A}, \mathcal{B})$ and $\mathcal{G}^-(\mathcal{A}, \mathcal{B})$.

We assume throughout this section that $n > 0$ and the underlying signature ν is n-adequate.

Definition 4.1. *Let θ be a formula in $\mathcal{L}_{\infty\omega}$. Define $(\theta \restriction z)$ to be the formula obtained from θ by first replacing each bound occurrence of z by a variable z' that does not occur in θ, and then relativizing all of its quantifiers to the set $\{y : E(z, y)\}$, i.e. everywhere replacing $(\exists y)\varphi(\mathbf{x}, y)$ by $(\exists y)(E(z, y) \wedge \varphi(\mathbf{x}, y))$.*

Note that in the above definition, the variable z may have already been free in θ.

Recall that $\chi(x)$ is the formula $(\exists y)E(x, y) \wedge \neg(\exists y)E(y, x)$, which says that x is a source for E.

Definition 4.2. *Let θ be a sentence in $\mathcal{L}_{\infty\omega}$. Define $Array_n(\theta)$ to be the sentence*

$$(\exists x_1)\cdots(\exists x_{n+1})\left[R(x_1,\ldots,x_{n+1})\wedge\bigwedge_{i=1}^{n+1}[\chi(x_i)\wedge(\theta\restriction x_i)]\right].$$

Informally, $(\theta \restriction z)$ says that z points to a (nonempty) submodel that satisfies θ, and $Array_n(\theta)$ says:

"There are $n+1$ elements related by R,
which are sources for E and point to θ-submodels."

$Array_n(\theta)$ is a sentence in $\mathcal{L}_{\infty\omega}$ with the same underlying signature ν. If θ is first order, then $Array_n(\theta)$ is also first order. We also have a bound on the quantifier rank of $Array_n(\theta)$.

Lemma 4.3. *If $m>0$ and $\theta\in\mathcal{L}_{\infty\omega}(\{\exists\},m)$, then*

$$Array_n(\theta)\in\mathcal{L}_{\infty\omega}(\{\exists\},m+n+1).$$

Proof: This follows easily from the definition of $Array_n(\theta)$. ■ *4.3*

Lemma 4.4. *Let θ and φ be sentences in $\mathcal{L}_{\infty\omega}$. If θ is equivalent to φ, then $Array_n(\theta)$ is equivalent to $Array_n(\varphi)$.*

Proof: Suppose θ is equivalent to φ. Since $\chi(x)$ implies that the set $\{y : E(x,y)\}$ is nonempty, the sentence

$$(\forall x)[\chi(x)\rightarrow[(\theta\restriction x)\leftrightarrow(\varphi\restriction x)]]$$

holds in all models. The lemma follows from this. ■ *4.4*

The above lemma would fail if we replaced $\chi(x)$ by the simpler formula $\neg(\exists y)E(y,x)$. If we defined $Array_n$ with the formula $\neg(\exists y)E(y,x)$ in place of $\chi(x)$, and took θ to be $(\exists y)y=y$ and φ to be $\neg(\exists y)\neg y=y$, then θ would be equivalent to φ but $Array_n(\theta)$ would not be equivalent to $Array_n(\varphi)$.

The power of the $Array_n$ operation stems from the following:

Lemma 4.5. *Let $\mathcal{A}$ and $\mathcal{B}$ be two finite models, and let θ be a first order sentence.*

(i) *If $\mathcal{A}\models\theta$ then $\mathcal{G}^+(\mathcal{A},\mathcal{B})\models Array_n(\theta)$.*
(ii) *If $\mathcal{B}\not\models\theta$ then $\mathcal{G}^-(\mathcal{A},\mathcal{B})\not\models Array_n(\theta)$.*

Proof: (i) In $\mathcal{G}^+(\mathcal{A},\mathcal{B})$ the $(n+1)$-tuple $(c_1,\ldots,c_{n+1})$ belongs to R^+ because $|\{c_1,\ldots,c_{n+1}\}\cap D|=0$ which is even, and each c_i points to a copy of $\mathcal{A}$.

(ii) In $\mathcal{G}^-(\mathcal{A},\mathcal{B})$, every $(n+1)$-tuple in R^- has at least one term that points to a copy of $\mathcal{B}$, which is not a model of θ. ■ *4.5*

Lemma 4.6. *Let $m>0$ and let θ be a sentence in $\mathcal{L}_{\infty\omega}$. If θ is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n,m)$, then $Array_n(\theta)$ is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n,m+1)$.*

Proof: By Proposition 2.2, there must be two models $\mathcal{A}, \mathcal{B}$ such that $\mathcal{A} \models \theta$, $\mathcal{B} \not\models \theta$, and $\mathcal{A} \equiv_n^m \mathcal{B}$. By Lemma 4.5,

$$\mathcal{G}^+(\mathcal{A}, \mathcal{B}) \models Array_n(\theta) \text{ and } \mathcal{G}^-(\mathcal{A}, \mathcal{B}) \not\models Array_n(\theta).$$

Also, using Lemma 3.6, we get

$$\mathcal{G}^+(\mathcal{A}, \mathcal{B}) \equiv_n^{m+1} \mathcal{G}^-(\mathcal{A}, \mathcal{B}).$$

By Proposition 2.2 again, $Array_n(\theta)$ is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m+1)$. ■ *4.6*

We remark that there is a natural way to extend the $Array_n$ operation to all sentences of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n)$. However, this would require a discussion of relativized generalized quantifiers, and will not be needed here.

5. The Quantifier Rank Hierarchies

It will be convenient to have a short way to say that a sentence is not expressible in a logic $\mathcal{L}$. Given a logic $\mathcal{L}$, we let $[\mathcal{L}]$ denote the set of all sentences (of some $\mathcal{L}_{\infty\omega}(\mathbf{Q})$) that are **expressible** in $\mathcal{L}$. Then $\varphi \notin [\mathcal{L}]$ says that φ is not expressible in $\mathcal{L}$, i.e., no sentence of $\mathcal{L}$ is equivalent to φ.

From Lemmas 4.3 and 4.6, we get:

Theorem 5.1. *Suppose $n > 0$ and the underlying signature ν is n-adequate. For each m there is a first order sentence of quantifier rank $m(n+1)+1$ that is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$, that is, there is a sentence*

$$\rho_m \in FO(\{\exists\}, m(n+1)+1) \setminus [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)].$$

Proof: Let ρ_0 be any first order sentence of quantifier rank 1 that is not equivalent to a quantifier-free formula, that is, $\rho_0 \in FO(\{\exists\}, 1) \setminus [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, 0)]$. For definiteness, we'll take $\rho_0 = (\exists x_1)E(x_1, x_1)$. For each m, let

$$\rho_{m+1} = Array_n(\rho_m).$$

Proceeding by induction on m, we use Lemma 4.3 to show that $\rho_m \in FO(\{\exists\}, m(n+1)+1)$, and use Lemma 4.6 to show that $\rho_m \notin [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)]$. ■ *5.1*

For each j, let $(\exists\, x_1, \ldots, x_j)$ be the j-ary existential quantifier

$$\{(A, S) : (A, S) \models (\exists x_1) \cdots (\exists x_j) S(x_1, \ldots, x_j)\},$$

and let $(\forall\, x_1, \ldots, x_j)$ be the corresponding j-ary universal quantifier.

Note that if the n-ary existential quantifier belongs to $\mathbf{Q}$, then the preceding proof shows that ρ_m is expressible in $FO(\mathbf{Q}, 2m+1)$.

We can now prove the first of our main results.

Theorem 5.2. *Suppose $n > 0$ and the underlying signature ν is n-adequate. Let $\mathbf{Q}$ be a class of generalized quantifiers of arity at most n (i.e. $\mathbf{Q} \subseteq \mathbf{Q}_n$). Assume that for some $j > 0$, $\mathbf{Q}$ contains at least one of the quantifiers $(\exists\, x_1, \ldots, x_j)$ or $(\forall\, x_1, \ldots, x_j)$. Then:*

(i) *The finite quantifier rank hierarchy of the logic* $\mathcal{L}_{\infty\omega}(\mathbf{Q})$ *is strict, i.e. for each* m,

$$[\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)] \subsetneq [\mathcal{L}_{\infty\omega}(\mathbf{Q}, m+1)].$$

(ii) *The quantifier rank hierarchy of the logic* $FO(\mathbf{Q})$ *is strict, i.e. for each* m,

$$[FO(\mathbf{Q}, m)] \subsetneq [FO(\mathbf{Q}, m+1)].$$

Proof: Suppose that (i) fails, so for some m,

$$[\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)] = [\mathcal{L}_{\infty\omega}(\mathbf{Q}, m+1)].$$

It follows by induction that the quantifier rank hierarchy in $\mathcal{L}_{\infty\omega}(\mathbf{Q})$ collapses. Thus,

$$[\mathcal{L}_{\infty\omega}(\mathbf{Q}, m(n+1)+1)] = [\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)].$$

By Lemma 5.1 there is a sentence

$$\rho_m \in FO(\{\exists\}, m(n+1)+1) \setminus [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)].$$

$\mathbf{Q}$ is a subset of $\mathbf{Q}_n$, so $\rho_m \notin [\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)]$. If the variables $x_2, \ldots, x_j$ do not occur in a formula φ, then the formula $(\exists x_1)\varphi$ is equivalent to each of the formulas $(\exists x_1, \ldots, x_j)\varphi$ and $\neg(\forall x_1, \ldots, x_j)\neg\varphi$. Since $\mathbf{Q}$ contains either the quantifier $(\exists\, x_1, \ldots, x_j)$ or the quantifier $(\forall\, x_1, \ldots, x_j)$, it follows that ρ_m is expressible in $FO(\mathbf{Q}, m(n+1)+1)$. But

$$[FO(\mathbf{Q}, m(n+1)+1)] \subseteq [\mathcal{L}_{\infty\omega}(\mathbf{Q}, m(n+1)+1)] = [\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)],$$

so $\rho_m \in [\mathcal{L}_{\infty\omega}(\mathbf{Q}, m)]$ and we have a contradiction.

The proof of (ii) is similar. ∎ *5.2*

In Theorem 5.2, $\mathbf{Q}$ could in particular be just $\{\exists\}$ or the whole $\mathbf{Q}_n$. This theorem concerns finite quantifier ranks. Sentences of quantifier rank ω will be considered in Section 8. Note that the quantifier rank hierarchies collapse beyond ω, because every class of finite models can be defined by a sentence of $\mathcal{L}_{\infty\omega}$ of quantifier rank ω. When the underlying signature is finite, every class can even be defined by a countable disjunction of first order sentences. On the other hand, every formula in $FO(\mathbf{Q})$ has finite quantifier rank.

6. Bounding the Number of Variables

We now obtain hierarchy results similar to those of Theorem 5.2 but within the fragments of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n)$ and $FO(\mathbf{Q}_n)$ whose formulas have at most a fixed finite number of variables (free and bound).

Definition 6.1. *For a formula* φ *in* $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n)$, *let* $Var(\varphi)$ *be the set of distinct variables (free and bound) occurring in* φ. *The number of variables of* φ *is simply* $|Var(\varphi)|$.

Given a natural number k *and a logic* $\mathcal{L} \subseteq \mathcal{L}_{\infty\omega}(\mathbf{Q}_n)$, *the* ***k-variable fragment*** *of* $\mathcal{L}$ *is the set* $\mathcal{L}^k$ *of all formulas* $\varphi \in \mathcal{L}$ *with* $|Var(\varphi)| \leq k$. *We also define* $\mathcal{L}^\omega = \bigcup_k \mathcal{L}^k$.

In particular, for each set of quantifiers $\mathbf{Q} \subseteq \mathbf{Q}_n$ we will consider the k-variable fragments

$$\mathcal{L}^k_{\infty\,\omega}(\mathbf{Q}),\ \mathcal{L}^k_{\infty\,\omega}(\mathbf{Q},m),\ FO^k(\mathbf{Q}),\ FO^k(\mathbf{Q},m).$$

It is obvious that $FO(\mathbf{Q}) = FO^{\omega}(\mathbf{Q})$, and for each m, $FO(\mathbf{Q},m) = FO^{\omega}(\mathbf{Q},m)$. From Proposition 2.1, we know that if both $\mathbf{Q}$ and the underlying signature ν are finite, then for each m, there exists a k (that depends on m, ν, and $\mathbf{Q}$), such that

$$[\mathcal{L}_{\infty\,\omega}(\mathbf{Q},m)] = [\mathcal{L}^{\omega}_{\infty\,\omega}(\mathbf{Q},m)] = [FO(\mathbf{Q},m)] = [FO^k(\mathbf{Q},m)].$$

The k-variable fragment of first order logic is FO^k. We note that there is no fixed k such that for all m, the sentences ρ_m of Lemma 4.6 is expressible in FO^k. The reason is that when passing from ρ_m to $\rho_{m+1} = Array_n(\rho_m)$, the number of variables in $(\rho_m \restriction x_i)$ is one more than the number of variables in ρ_m.

The next result is an analogue of Theorem 5.1 which gives sentences $\sigma_m \in FO^{n+2}$ that do the same thing as the sentences ρ_m.

Theorem 6.2. *Suppose $n > 0$ and the underlying signature ν is n-adequate. Let $k = n+2$. For each m there is a first order sentence of quantifier rank $mk+1$ that has at most k variables and is not expressible in $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}_n, m)$, that is, there is a sentence*

$$\sigma_m \in FO^k(\{\exists\}, mk+1) \setminus [\mathcal{L}_{\infty\,\omega}(\mathbf{Q}_n, m)].$$

Proof: Let ρ_m be the sentences defined in the proof of Theorem 5.1,

$$\rho_0 = (\exists x_1)E(x_1,x_1), \quad \rho_{m+1} = Array_n(\rho_m).$$

We first define $\sigma_0 = \rho_0 = (\exists x_1)E(x_1,x_1)$, which belongs to $FO^k(\{\exists\},1)$ but not to $[\mathcal{L}_{\infty\,\omega}(\mathbf{Q}_n,0)]$. Next, we define auxiliary formulas $\eta_m(z)$ in $FO^k(\{\exists\}, mk+1)$, which will play the role of the relativized formula $(\rho_m \restriction x_i)$ in the sentence $\rho_{m+1} = Array_n(\rho_m)$.

We define

$$\eta_0(z) \,:\, (\exists x_1)[E(z,x_1) \wedge \neg E(x_1,z) \wedge E(x_1,x_1)],$$

and define $\eta_{m+1}(z)$ inductively by

$$(\exists x_1)\cdots(\exists x_{n+1})$$
$$\left[R(x_1,\ldots,x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} (E(z,x_i) \wedge \neg E(x_i,z) \wedge (\exists z)(z = x_i \wedge \eta_m(z)))\right].$$

Note how the variable z was reused. The idea is that once we know that $E(z,x_i)$ holds for all i, we can just forget about z, and concentrate on each x_i. We now define σ_{m+1} as follows:

$$\sigma_{m+1} : (\exists x_1)\cdots(\exists x_{n+1})\left[R(x_1,\ldots,x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} (\exists z)(z = x_i \wedge \eta_m(z))\right].$$

This formula can be compared with the defining formula for ρ_{m+1}, which is

$$(\exists x_1)\cdots(\exists x_{n+1})\left[R(x_1,\ldots,x_{n+1})\wedge\bigwedge_{i=1}^{n+1}[\chi(x_i)\wedge(\rho_m\restriction x_i)]\right]$$

where $\chi(z)$ is the formula $(\exists y)E(z,y)\wedge\neg(\exists y)E(y,z)$.

By induction on m, we can easily see that both $\eta_m(z)$ and σ_m belong to $FO^k(\{\exists\},mk+1)$.

We note that for each m, z is the only free variable in the formula $\eta_m(z)$. If x is any other variable, we define $\eta_m(x)$ to be the formula obtained from $\eta_m(z)$ by replacing each occurrence of x by a new variable x', and then replacing each free occurrence of z by x. Thus the formula

$$x=z\rightarrow[\eta_m(x)\leftrightarrow\eta_m(z)]$$

is logically valid.

Claim 6.3. *Let $\mathcal{A}$ be a finite model with signature ν. Suppose that the interpretation of E in $\mathcal{A}$ is transitive, that is,*

$$\mathcal{A}\models(\forall x)(\forall y)(\forall z)[[E(x,y)\wedge E(y,z)]\rightarrow E(x,z)].$$

Then for each m,

(i) $\mathcal{A}\models(\forall z)(\forall y)[E(z,y)\rightarrow[\eta_m(y)\leftrightarrow(\eta_m(y)\restriction z)]]$,

(ii) $\mathcal{A}\models(\forall z)[\eta_m(z)\leftrightarrow(\eta_m(z)\restriction z)]$.

Proof of Claim 6.3: (i) We argue by induction on m. The result is clear for $m=0$. Let $\mathcal{A}$ be a finite model with signature ν. Work in $\mathcal{A}$ and assume the result for m. By renaming bound occurrences of y in $\eta_m(z)$, we see that $\eta_{m+1}(y)$ is equivalent to

$$(\exists x_1)\cdots(\exists x_{n+1})$$

$$\left[R(x_1,\ldots,x_{n+1})\wedge\bigwedge_{i=1}^{n+1}\{E(y,x_i)\wedge\neg E(x_i,y)\wedge(\exists z)[z=x_i\wedge\eta_m(z)]\}\right].$$

$(\eta_{m+1}(y)\restriction z)$ is equivalent to each of the following formulas:

$$(\exists x_1)\cdots(\exists x_{n+1})\left[\bigwedge_{i=1}^{n+1}E(z,x_i)\wedge R(x_1,\ldots,x_{n+1})\wedge\right.,$$

$$\left.\bigwedge_{i=1}^{n+1}\{E(y,x_i)\wedge\neg E(x_i,y)\wedge(\exists z')[E(z,z')\wedge z'=x_i\wedge(\eta_m(z')\restriction z)]\}\right],$$

$$(\exists x_1)\cdots(\exists x_{n+1})\left[\bigwedge_{i=1}^{n+1}E(z,x_i)\wedge R(x_1,\ldots,x_{n+1})\wedge\right.$$

$$\left.\bigwedge_{i=1}^{n+1}\{E(y,x_i)\wedge\neg E(x_i,y)\wedge E(z,x_i)\wedge(\eta_m(x_i)\restriction z)\}\right],$$

$$(\exists x_1)\cdots(\exists x_{n+1})$$
$$\left[R(x_1,\ldots,x_{n+1})\wedge\bigwedge_{i=1}^{n+1}\{E(y,x_i)\wedge\neg E(x_i,y)\wedge E(z,x_i)\wedge(\eta_m(x_i)\restriction z)\}\right].$$

Now assume $E(z,y)$. Then since E is transitive, we have

$$(\forall x)[E(y,x)\leftrightarrow[E(y,x)\wedge E(z,x)]].$$

Therefore $(\eta_{m+1}(y)\restriction z)$ holds if and only if

$$(\exists x_1)\cdots(\exists x_{n+1})$$
$$\left[R(x_1,\ldots,x_{n+1})\wedge\bigwedge_{i=1}^{n+1}\{E(y,x_i)\wedge\neg E(x_i,y)\wedge(\eta_m(x_i)\restriction z)\}\right].$$

By inductive hypothesis, this holds if and only if

$$(\exists x_1)\cdots(\exists x_{n+1})$$
$$\left[R(x_1,\ldots,x_{n+1})\wedge\bigwedge_{i=1}^{n+1}\{E(y,x_i)\wedge\neg E(x_i,y)\wedge\eta_m(x_i)\}\right],$$

which is clearly equivalent to $\eta_{m+1}(y)$. This completes the induction.

(ii) The formula $(\eta_{m+1}(z)\restriction z)$ is built by replacing the bound occurrences of z by a new variable z' and then relativizing all quantifiers to $\{u : E(z,u)\}$. Thus $(\eta_{m+1}(z)\restriction z)$ is

$$(\exists x_1)\cdots(\exists x_{n+1})\left[\bigwedge_{i=1}^{n+1}E(z,x_i)\wedge R(x_1,\ldots,x_{n+1})\wedge\right.$$
$$\left.\bigwedge_{i=1}^{n+1}\{E(z,x_i)\wedge\neg E(x_i,z)\wedge(\exists z')[E(z,z')\wedge z'=x_i\wedge(\eta_m(z')\restriction z)]\}\right].$$

By part (i), the subformula

$$E(z,x_i)\wedge\neg E(x_i,z)\wedge(\exists z')[E(z,z')\wedge z'=x_i\wedge(\eta_m(z')\restriction z)]$$

is equivalent to each of the following:

$$E(z,x_i)\wedge\neg E(x_i,z)\wedge(\exists z')[E(z,z')\wedge z'=x_i\wedge\eta_m(z')],$$
$$E(z,x_i)\wedge\neg E(x_i,z)\wedge(\exists z')[E(z,x_i)\wedge z'=x_i\wedge\eta_m(z')],$$
$$E(z,x_i)\wedge\neg E(x_i,z)\wedge(\exists z')[z'=x_i\wedge\eta_m(z')],$$
$$E(z,x_i)\wedge\neg E(x_i,z)\wedge(\exists z)[z=x_i\wedge\eta_m(z)].$$

Therefore $(\eta_{m+1}(z)\restriction z)$ is equivalent to

$$(\exists x_1)\cdots(\exists x_{n+1})\left[\bigwedge_{i=1}^{n+1}E(z,x_i)\wedge R(x_1,\ldots,x_{n+1})\wedge\right.$$

$$\left. \bigwedge_{i=1}^{n+1} \{E(z,x_i) \wedge \neg E(x_i,z) \wedge (\exists z)[z = x_i \wedge \eta_m(z)]\} \right] .$$

This is clearly equivalent to

$$(\exists x_1) \cdots (\exists x_{n+1})$$
$$\left[R(x_1,\ldots,x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} \{E(z,x_i) \wedge \neg E(x_i,z) \wedge (\exists z)[z = x_i \wedge \eta_m(z)]\} \right],$$

which by definition is $\eta_{m+1}(z)$. ∎ *6.3*

The proof of the theorem will be complete once we show that $\sigma_m \notin [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)]$. We will do this by producing a sequence of pairs of models $\mathcal{A}_m, \mathcal{B}_m$ such that $\mathcal{A}_m \equiv_n^m \mathcal{B}_m$, but $\mathcal{A}_m \models \sigma_m$ and $\mathcal{B}_m \not\models \sigma_m$.

We start by choosing two models $\mathcal{A}_0$ and $\mathcal{B}_0$ in which the interpretation of E is transitive, such that $\mathcal{A}_0 \models \sigma_0$ and $\mathcal{B}_0 \not\models \sigma_0$, but $\mathcal{A}_0 \equiv_n^0 \mathcal{B}_0$. This is easy, since the relation $\equiv_n^0$ holds between any two relational models, but for definiteness, we take $\mathcal{A}_0$ and $\mathcal{B}_0$ to have a universe $\{c\}$ of size 1. Also, in $\mathcal{A}_0$ the only atomic formula that holds is $E(c,c)$, while in $\mathcal{B}_0$ no atomic formula holds. Next, we recursively define

$$\mathcal{A}_{m+1} = \mathcal{G}^+(\mathcal{A}_m, \mathcal{B}_m) \text{ and } \mathcal{B}_{m+1} = \mathcal{G}^-(\mathcal{A}_m, \mathcal{B}_m).$$

The following claim is proved by an easy induction on m:

Claim 6.4. *For each m, both $\mathcal{A}_m$ and $\mathcal{B}_m$ have a transitive interpretation of E, and moreover satisfy the following formula for each $i \le n+1$:*

$$[R(x_1,\ldots,x_{n+1}) \wedge E(z,x_i)] \rightarrow \neg E(x_i,z).$$

∎ *6.4*

It follows from Lemma 3.6 that $\mathcal{A}_m \equiv_n^m \mathcal{B}_m$, so by Proposition 2.2, $\mathcal{A}_m$ and $\mathcal{B}_m$ agree on $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$. From Lemma 4.5 and the recursive definition of ρ_m, we know that $\mathcal{A}_m \models \rho_m$ and $\mathcal{B}_m \not\models \rho_m$. So it will be enough to show that for each m, $\sigma_m \leftrightarrow \rho_m$ holds in both $\mathcal{A}_m$ and $\mathcal{B}_m$. We first get a connection between $\eta_m(z)$ and $\chi(z)$.

Claim 6.5. *For each m, both $\mathcal{A}_{m+1}$ and $\mathcal{B}_{m+1}$ satisfy* $(\forall z)[\eta_m(z) \rightarrow \chi(z)]$.

Proof of Claim 6.5: One can show by induction that in both $\mathcal{A}_{m+1}$ and $\mathcal{B}_{m+1}$, an element a satisfies $\chi(z)$ if and only if there is a directed path of length $m+1$ starting at a. Also, if a satisfies $\eta_m(z)$ then there is such a directed path. ∎ *6.5*

We now get a connection between $\eta_m(z)$ and the relativized formula $(\rho_m \restriction z)$.

Claim 6.6. *For each m, both $\mathcal{A}_{m+1}$ and $\mathcal{B}_{m+1}$ satisfy the sentence*

$$(\forall z)[\eta_m(z) \leftrightarrow [\chi(z) \wedge (\rho_m \restriction z)]].$$

Proof of Claim 6.6: We argue by induction on m. The result is clear for $m = 0$. Suppose $m > 0$ and assume the result for $m - 1$. To prove the result for m, we work in either $\mathcal{A}_{m+1}$ or $\mathcal{B}_{m+1}$.

Suppose that $\chi(z) \wedge (\rho_m \restriction z)$. Since $(\rho_m \restriction z)$, z points to a model $\mathcal{C}$ of ρ_m. Then there are elements $x_1, \ldots, x_{n+1}$ of $\mathcal{C}$ such that

$$\mathcal{C} \models R(x_1, \ldots, x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} [\chi(x_i) \wedge (\rho_{m-1} \restriction x_i)].$$

Since $\chi(z)$, $\mathcal{C}$ is a copy of either $\mathcal{A}_m$ or $\mathcal{B}_m$. By inductive hypothesis,

$$\mathcal{C} \models R(x_1, \ldots, x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} \eta_{m-1}(x_i).$$

Therefore

$$R(x_1, \ldots, x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} [E(z, x_i) \wedge (\eta_{m-1}(x_i) \restriction z)].$$

By Claim 6.3,

$$R(x_1, \ldots, x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} [E(z, x_i) \wedge \eta_{m-1}(x_i)].$$

By Claim 6.4,

$$R(x_1, \ldots, x_{n+1}) \wedge \bigwedge_{i=1}^{n+1} [E(z, x_i) \wedge \neg E(x_i, z) \wedge \eta_{m-1}(x_i)],$$

and $\eta_m(z)$ follows.

Now suppose $\eta_m(z)$. By Claim 6.5, $\chi(z)$ holds, so again z points to a model $\mathcal{C}$ which is a copy of $\mathcal{A}_m$ or $\mathcal{B}_m$. By reversing the steps in the preceding paragraph, we can show that $(\rho_m \restriction z)$. ■ *6.6*

To complete the proof of Theorem 6.2, we show that for each m, $\sigma_m \leftrightarrow \rho_m$ holds in both $\mathcal{A}_m$ and $\mathcal{B}_m$. The case $m = 0$ is trivial, since $\sigma_0 = \rho_0$. For $m + 1$, this follows easily from Claim 6.6 and the defining formulas for ρ_{m+1} and σ_{m+1}. ■ *6.2*

We now get the following theorem, which parallels Theorem 5.2 with a similar proof:

Theorem 6.7. *Suppose $n > 0$ and the underlying signature ν is n-adequate. Let $n > 0$ and let $\mathbf{Q} \subseteq \mathbf{Q}_n$ be a class of generalized quantifiers of arity at most n. Also, assume that for some $j > 0$, $\mathbf{Q}$ contains at least one of the quantifiers $(\exists\, x_1, \ldots, x_j)$ or $(\forall\, x_1, \ldots, x_j)$. Then:*

(i) *For each $k \geq n + 2$, the finite quantifier rank hierarchy of the logic $\mathcal{L}^k_{\infty\omega}(\mathbf{Q})$ is strict, i.e. for each m,*

$$[\mathcal{L}^k_{\infty\omega}(\mathbf{Q}, m)] \subsetneq [\mathcal{L}^k_{\infty\omega}(\mathbf{Q}, m+1)].$$

(ii) *The finite quantifier rank hierarchy of the logic* $\mathcal{L}^{\omega}_{\infty\,\omega}(\mathbf{Q})$ *is strict, i.e. for each* m,

$$[\mathcal{L}^{\omega}_{\infty\,\omega}(\mathbf{Q}, m)] \subsetneq [\mathcal{L}^{\omega}_{\infty\,\omega}(\mathbf{Q}, m+1)].$$

(iii) *For each* $k \geq n+2$*, the quantifier rank hierarchy of the logic* $FO^k(\mathbf{Q})$ *is strict, i.e. for each* m,

$$[FO^k(\mathbf{Q}, m)] \subsetneq [FO^k(\mathbf{Q}, m+1)].$$

■ *6.7*

7. The Quantifier-positive Rank Hierarchies

In this section we consider the quantifier rank hierarchies inside the quantifier-positive fragments of $\mathcal{L}_{\infty\,\omega}(\mathbf{Q}, n)$ defined below. The term "quantifier-positive" stems from the fact that negation must only have a quantifier-free scope, so every occurrence of a quantifier is positive.

Definition 7.1. *Given a logic* $\mathcal{L} \subseteq \mathcal{L}_{\infty\,\omega}(\mathbf{Q}, n)$*, the* ***quantifier-positive fragment*** $\mathcal{L}^+$ *of* $\mathcal{L}$ *is the set of all formulas* $\varphi \in \mathcal{L}$ *in which no occurrence of a quantifier is in the scope of a negation.*

In particular, we will consider the quantifier-positive fragments

$$\mathcal{L}^{+}_{\infty\,\omega}(\mathbf{Q}),\ \mathcal{L}^{+}_{\infty\,\omega}(\mathbf{Q}, m),\ FO^+(\mathbf{Q}),\ FO^+(\mathbf{Q}, m)$$

and the k-variable quantifier-positive fragments

$$\mathcal{L}^{k+}_{\infty\,\omega}(\mathbf{Q}),\ \mathcal{L}^{k+}_{\infty\,\omega}(\mathbf{Q}, m),\ FO^{k+}(\mathbf{Q}),\ FO^{k+}(\mathbf{Q}, m).$$

It is important in this section that the underlying logic $\mathcal{L}_{\infty\,\omega}(\mathbf{Q})$ has both the conjunction and the disjunction connectives. Thus each of the above fragments is closed under conjunction and disjunction.

Each sentence $\rho \in FO^+ = FO^+(\{\exists\})$ is equivalent to a **purely existential sentence**, that is, a sentence of the form

$$(\exists x_1) \cdots (\exists x_\ell)\varphi(x_1, \ldots, x_\ell)$$

where φ is quantifier-free. The logics $\mathcal{L}^{k+}_{\infty\,\omega}$ and $\mathcal{L}^{\omega+}_{\infty\,\omega} = \bigcup_k \mathcal{L}^{k+}_{\infty\,\omega}$ have been studied in the literature under the names $\exists\mathcal{L}^{k}_{\infty\,\omega}$ and $\exists\mathcal{L}^{\omega}_{\infty\,\omega}$, respectively, and are characterized by the so-called existential pebble game. The importance of these logics stems from the fact that the latter contains the language *DATALOG* (see [KV95]).

We are going to show that each of the above fragments has a strict finite quantifier rank hierarchy. We can do this using the same sentences σ_m as we used in Section 6.

Proposition 7.2. *Suppose* $n > 0$ *and the underlying signature* ν *is* n*-adequate. Let* $k = n+2$*. The sentences* σ_m *from Theorem 6.2 belong to* $FO^{k+}(\{\exists\}, mk+1)$*. Moreover,* σ_m *is equivalent to a purely existential first order sentence*

$$\sigma'_m = (\exists y_1) \ldots (\exists y_\ell)\varphi(y_1, \ldots, y_\ell)$$

where φ *is quantifier-free and* $\ell = m(n+1)+1$.

Proof: It is clear from the definition that for each m the formula $\eta_m(z)$ belongs to FO^+, and hence that σ_m belongs to FO^+. The equivalent sentence σ'_{m+1} is obtained by first replacing $(\exists z)[x = x_i \wedge \eta_m(z)]$ by $\eta_m(x_i)$, then replacing $x_1, \ldots, x_{n+1}$ by new variables in the formula σ_{m+1}, and moving all the quantifiers to the front. ∎ *7.2*

Note that the number $m(n+1)+1$ in Proposition 7.2 is close to optimal—the sentence σ_m is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$, but any purely existential sentence

$$(\exists y_1)\ldots(\exists y_{mn})\varphi(y_1, \ldots, y_{mn})$$

is expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$.

If $\mathbf{Q}$ contains the n-ary existential quantifier, then by merging the existential quantifiers we see that σ_m is expressible in $FO^+(\mathbf{Q}, m + [m/n] + 1)$.

We now get the following theorem, which again parallels Theorem 5.2 with a similar proof:

Theorem 7.3. *Suppose $n > 0$ and the underlying signature ν is n-adequate. Let $\mathbf{Q} \subseteq \mathbf{Q}_n$ be a class of generalized quantifiers of arity at most n. Also, assume that for some $j > 0$, $\mathbf{Q}$ contains at least one of the quantifiers $(\exists\, x_1, \ldots, x_j)$ or $(\forall\, x_1, \ldots, x_j)$. Then:*

(i) *The finite quantifier rank hierarchies of the logics $\mathcal{L}^{+}_{\infty\omega}(\mathbf{Q})$, $\mathcal{L}^{k+}_{\infty\omega}(\mathbf{Q})$, and $\mathcal{L}^{\omega+}_{\infty\omega}(\mathbf{Q})$ are strict*

(ii) *The quantifier rank hierarchies of the logics $FO^+(\mathbf{Q})$ and $FO^{k+}(\mathbf{Q})$ are strict.* ∎ *7.3*

8. Beyond Finite Quantifier Ranks

In this section we show that each of the hierarchies in Theorems 5.2 (i), 6.7 (i), (ii), and 7.3 (i) has sentences at level ω. To do this, it is enough to show that there are sentences $\tau \in \mathcal{L}^k_{\infty\omega}$ and $\tau^+ \in \mathcal{L}^{k+}_{\infty\omega}$, both of quantifier rank ω, that are not expressible in the logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, \omega)$ of sentences of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n)$ with finite quantifier rank. We will build τ and τ^+ out of the sentences σ_m that were constructed in Section 6.

We first take up the finite variable case.

Theorem 8.1. *Suppose $n > 0$ and the underlying signature ν is n-adequate. Let $k = n+2$. Then there is a sentence $\tau \in \mathcal{L}^k_{\infty\omega} \backslash [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, \omega)]$ of quantifier rank ω.*

Proof: For each m, let σ_m be the sentence constructed in Section 6. There is a first order sentence $\pi_m \in FO^2$ with two variables that says that the graph E has a directed path of length m, that is, there is a sequence of elements $y_0, \ldots, y_m$ such that $E(y_k, y_{k+1}) \wedge \neg E(y_{k+1}, y_k)$ for each $k < m$. Let τ be the sentence

$$\tau : \bigvee_m [\sigma_m \wedge \neg\pi_{m+1}].$$

It is clear that $\tau \in \mathcal{L}^k_{\infty\omega}$ with quantifier rank ω. Let $\mathcal{A}_m$ and $\mathcal{B}_m$ be the models constructed in Section 6. It was shown in the proof of Theorem 6.2 that $\mathcal{A}_m$ and $\mathcal{B}_m$ satisfy the same sentences of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$, and that $\mathcal{A}_m \models \sigma_m$ and $\mathcal{B}_m \not\models \sigma_m$. It is clear from the construction that both $\mathcal{A}_m$ and $\mathcal{B}_m$ satisfy π_j for all $j \leq m$ and also satisfy $\neg\pi_{m+1}$. Moreover, σ_m implies π_j for all $j \leq m$. It follows that for all m, $\mathcal{A}_m \models \tau$ and $\mathcal{B}_m \not\models \tau$. Therefore τ is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, \omega)$. ∎ *8.1*

We will give two examples of sentences which do the job in the quantifier-positive fragment $\mathcal{L}^{k+}_{\infty\omega}$. In first example the underlying signature has countably many proposition symbols in addition to the relation symbols E and R. In the second example the underlying signature has just two symbols in addition to E and R, a binary predicate F and a unary predicate U.

Theorem 8.2. *Suppose $n > 0$ and the underlying signature ν is n-adequate and contains countably many proposition symbols $P_0, P_1, \ldots$. Let $k = n+2$. Then there is a sentence $\tau^+ \in \mathcal{L}^{k+}_{\infty\omega} \setminus [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, \omega)]$ of quantifier rank ω.*

Proof: Let $\nu_0 = \nu \setminus \{P_0, P_1, \ldots\}$ be the signature obtained by removing each proposition symbol P_i from ν. The sentences σ_m constructed in Section 6 use only the smaller signature ν_0. Let τ^+ be the sentence

$$\tau^+ : \bigvee_m [\sigma_m \wedge P_m].$$

Then $\tau^+ \in \mathcal{L}^{k+}_{\infty\omega}$ with quantifier rank ω. For each model $\mathcal{A}$ with signature ν_0 and each m, let $\mathcal{A}(P_m)$ be the model with signature ν which is built from $\mathcal{A}$ by making P_m true and making P_j false for all $j \neq m$. Let $\mathcal{A}_m$ and $\mathcal{B}_m$ be the models constructed in Section 6 in the smaller signature ν. Then $\mathcal{A}_m(P_m)$ and $\mathcal{B}_m(P_m)$ satisfy the same sentences of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$ in the larger signature ν, and $\mathcal{A}_m(P_m) \models \sigma_m \wedge P_m$ and $\mathcal{B}_m(P_m) \not\models \sigma_m$. Moreover, for all $j \neq m$, $\mathcal{B}_m(P_m) \not\models P_j$. Therefore for all m, $\mathcal{A}_m(P_m) \models \tau^+$ and $\mathcal{B}_m(P_m) \not\models \tau^+$. This shows that τ^+ is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, \omega)$. ∎ *8.2*

Theorem 8.3. *Suppose $n > 0$ and the underlying signature ν is n-adequate and contains at least another binary predicate symbol F and unary predicate symbol U. Let $k = n+2$. Then there is a sentence $\tau'^+ \in \mathcal{L}^{k+}_{\infty\omega} \setminus [\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, \omega)]$ of quantifier rank ω.*

Proof: Let $\nu_1 = \nu \setminus \{F, U\}$ be the signature obtained by removing the symbols F, U from ν. As in the preceding proof, the sentences σ_m constructed in Section 6 use only the smaller signature ν_1. We first define a first order positive existential sentence π'_n with two variables that says that the relation F has a directed path with n edges that begins at an element of U and ends at an element x such that $E(x, x)$. Formally, we inductively define formulas $\varphi'_n(x)$ and $\varphi'_n(y)$ by:

$$\varphi'_0(x) : E(x, x), \quad \varphi'_0(y) : E(y, y),$$

$$\varphi'_{n+1}(x) : (\exists y)[F(x,y) \wedge \varphi'_n(y)], \quad \varphi'_{n+1}(y) : (\exists x)[F(y,x) \wedge \varphi'_n(x)].$$

We then define

$$\pi'_n : (\exists x)[U(x) \wedge \varphi'_n(x)].$$

Let τ'^{+} be the sentence

$$\tau'^{+} : \bigvee_{m=1}^{\infty} [\sigma_m \wedge \pi'_m].$$

Then $\tau'^{+} \in \mathcal{L}^{k+}_{\infty\omega}$ with quantifier rank ω. Let $\mathcal{A}_m$ and $\mathcal{B}_m$ be the models constructed in Section 6 in the smaller signature ν_1. Recall that for each m, $\mathcal{A}_m \models \sigma_m$ and $\mathcal{B}_m \not\models \sigma_m$. Moreover, $\mathcal{G}(\mathcal{A}_m, \mathcal{B}_m)$ has universe

$$G = X \cup (A_m \times C) \cup (B_m \times D)$$

where

$$C = \{c_1, \ldots, c_{n+1}\}, \quad D = \{d_1, \ldots, d_{n+1}\}, \quad X = C \cup D.$$

For each $c \in C$, the c-th copy of $\mathcal{A}_m$ in $\mathcal{G}(\mathcal{A}_m, \mathcal{B}_m)$ has universe $A_m \times \{c\}$, and for each $d \in D$ the d-th copy of $\mathcal{B}_m$ in $\mathcal{G}(\mathcal{A}_m, \mathcal{B}_m)$ has universe $B_m \times \{d\}$. We expand $\mathcal{A}_m$ and $\mathcal{B}_m$ to models $\mathcal{A}'_m$ and $\mathcal{B}'_m$ in the signature ν as follows. $\mathcal{A}'_0$ and $\mathcal{B}'_0$ are the expansions of $\mathcal{A}_0$ and $\mathcal{B}_0$ in which F is empty and $U = \{c\}$. $\mathcal{A}'_{m+1}$ and $\mathcal{B}'_{m+1}$ are the expansions of $\mathcal{A}_{m+1}$ and $\mathcal{B}_{m+1}$ where U is the set X in the construction $\mathcal{G}(\mathcal{A}_m, \mathcal{B}_m)$, and F is the set of all pairs (x, y) such that one of the following holds:

- $x = (x', c)$ and $y = (y', c)$ belong to the same copy of $\mathcal{A}_m$, and $\mathcal{A}'_m \models F(x', y')$,
- $x = (x', d)$ and $y = (y', d)$ belong to the same copy of $\mathcal{B}_m$, and $\mathcal{B}'_m \models F(x', y')$,
- $x \in C$, $y = (y', c)$ belongs to a copy of $\mathcal{A}_m$, and $\mathcal{A}'_m \models U(y')$, or
- $x \in D$, $y = (y', d)$ belongs to a copy of $\mathcal{B}_m$, and $\mathcal{B}'_m \models U(y')$.

It is easily shown by induction on m that F is included in E, that is, the sentence

$$\forall x \forall y (F(x,y) \to E(x,y))$$

holds in both $\mathcal{A}'_m$ and $\mathcal{B}'_m$. The proof of Lemma 3.6 can be slightly modified to show that $\mathcal{A}'_m$ and $\mathcal{B}'_m$ satisfy the same sentences of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, m)$ in the whole signature ν.

One can show by an easy induction that for each m, the sentence π'_m holds in $\mathcal{A}'_m$ (and holds in $\mathcal{B}'_m$ if $m > 0$), but for all $j \neq m$ the sentence π'_j fails in both $\mathcal{A}'_m$ and $\mathcal{B}'_m$. Therefore $\mathcal{A}'_m \models \sigma_m \wedge \pi'_m$, and for all j, $\mathcal{B}'_m \not\models \sigma_j \wedge \pi'_j$. Therefore for all m, $\mathcal{A}'_m \models \tau'^{+}$ and $\mathcal{B}'_m \not\models \tau'^{+}$, so τ'^{+} is not expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_n, \omega)$. ■ *8.3*

References:

[Bar77] J. Barwise, On Moschovakis closure ordinals, **Journal of Symbolic Logic**, vol. 42 (1977), pp. 292-296.

[EF99] H-D. Ebbinghaus and J. Flum, Finite Model Theory, Second Edition, **Springer Verlag** (1999).

[Hel89] L. Hella, Definability hierarchies of generalized quantifiers, **Annals of Pure and Applied Logic**, vol. 43 (1989), pp. 235-271.

[Hel96] L. Hella, Logical hierarchies in PTIME, **Information and Computation**, vol. 129 (1996), pp. 1-19.

[HLV96] L. Hella, K. Luosto, and J. Väänänen, The Hierarchy Theorem for Generalized Quantifiers. **Journal of Symbolic Logic**, vol. 61 (1996), pp. 802-817.

[Imm86] N. Immerman, Relational queries computable in polynomial time, **Information & Control**, vol. 68 (1986), pp. 86-104.

[Imm99] N. Immerman, Descriptive Complexity, **Springer Verlag**, New York (1999).

[JM01] D. Janin and J. Marcinkowski, Toolkit for first order extensions of monadic games, **Proceedings of STACS**, Springer, LNCS (2001).

[KL04] H. J. Keisler and W. B. Lotfallah, First order quantifiers in monadic second order logic, **J. Symbolic Logic**, vol. 69 (2004), pp. 118-136.

[KV92] P. G. Kolaitis and M. Y. Vardi, Fixpoint logic vs. infinitary logic in finite-model theory, **Proceedings of the Seventh Annual IEEE Symposium on Logic in Computer Science**, (1992), pp. 46-57.

[KV95] P. G. Kolaitis and M. Y. Vardi, On the expressive power of Datalog: tools and a case study, **Journal of Computer and System Sciences** vol. 51, iss. 1 (1995), pp. 110-134.

[Lin66] P. Lindström, First order predicate logic with generalized quantifiers, **Theoria**, vol. 32 (1966), pp. 186195.

[Luo00] K. Luosto, Hierarchies of monadic generalized quantifiers, **The Journal of Symbolic Logic**, vol. 65 (3) (2000), pp. 1241-1263.

[Mos74] Y. Moschovakis, Elementary induction on abstract structures, **North Holland** (1974).

[Var82] M. Vardi, Complexity of relational query languages, **Proceedings of 14th ACM Symposium on Theory of Computing**, (1982), pp. 137-146.

Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison WI 53706

E-mail address: keisler@math.wisc.edu

Department of Mathematics and Actuarial Science, The American University in Cairo, AUC Avenue, P.O. Box 74, New Cairo 11835, Egypt

E-mail address: lotfallah@aucegypt.edu, lotfalla@uwalumni.com