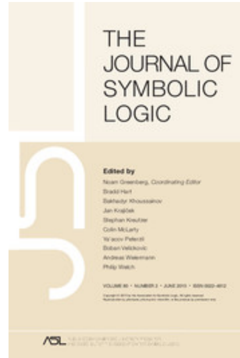
Sh:801



A dichotomy in classifying quantifiers for finite models

Saharon Shelah and Mor Doron

The Journal of Symbolic Logic / Volume 70 / Issue 04 / December 2005, pp 1297 - 1324
DOI: 10.2178/jsl/1129642126, Published online: 12 March 2014

Sh:801

THE JOURNAL OF SYMBOLIC LOGIC
Volume 70, Number 4, Dec. 2005

A DICHOTOMY IN CLASSIFYING QUANTIFIERS FOR FINITE MODELS

SAHARON SHELAH AND MOR DORON

Abstract. We consider a family $\mathfrak{U}$ of finite universes. The second order existential quantifier $Q_{\mathfrak{R}}$, means for each $U \in \mathfrak{U}$ quantifying over a set of $n(\mathfrak{R})$-place relations isomorphic to a given relation. We define a natural partial order on such quantifiers called interpretability. We show that for every $Q_{\mathfrak{R}}$, either $Q_{\mathfrak{R}}$ is interpretable by quantifying over subsets of U and one to one functions on U both of bounded order, or the logic $L(Q_{\mathfrak{R}})$ (first order logic plus the quantifier $Q_{\mathfrak{R}}$) is undecidable.

CONTENTS

§1. Introduction.

1.1. Background. In this work we continue [6], but it is self contained and the reader may read it independently. Our aim is to analyze and classify second order existential quantifiers in finite model theory. The quantifiers will be defined as follows:

(∗) Let U be a finite universe, and n a natural number. Let K be a class of n-place relations on U closed under permutations of U. Define Q_K to be the n-place existential quantifier ranging over the relations in K, i.e., the formula $(Q_K r)\varphi(r)$ holds iff $\varphi(R)$ holds for some $R \in K$.

Received December 18, 2003; accepted June 22, 2005.
Research supported by the United States-Israel Binational Science Foundation.
Publication no. 801 in Saharon Shelah's list.

0022-4812/05/7004-0014/$3.80

We will usually work on quantifiers of the form $Q_R = Q_{K_R}$ where R is a n-place relation over U and K_R is defined by: $K_R := \{R' \subseteq {}^nU : (U,R) \approx (U,R')\}$. We define below two partial orders on the class of such quantifiers, called: interpretability and expressibility. It will be interesting to consider the problem for classes K of n-place relations definable in some logic $\mathfrak{L}$, that is such that there exists a formula $\varphi(r) \in \mathfrak{L}$ (r is a n-place relation symbol) and $R \in K$ iff $(U,R) \models \varphi(r)$. In [4] the problem was solved for the case: K is definable in first order logic and U is infinite. It was shown that in this case Q_K is equivalent (in the sense of interpretability) to one of only four quantifiers: trivial (first order), monadic, quantifying over 1–1 functions or full second order. A revue paper is [1]. If we do not assume K to be first order definable but keep assuming U is infinite we get a classification of Q_K by equivalence relations. Formally in [5] it was shown:

THEOREM 1.1. *Let U be an infinite countable universe, and K be as in* $(*)$. *Then there exist a family E of equivalence relations on U, such that Q_K and Q_E are equivalent* (*each is interpretable by the other*).

We remark that if U is infinite not necessarily countable then the situation is more complicated, but if we assume $L = V$ then we have the same result. [6] deals with the case U is finite. Under this assumption we get a reasonable understanding of Q_R, we can "bound" it between two simple and close quantifiers (close meaning that the size of one is a polynomial in the size of the other). We say that Q_1 is "uniformly" interpretable (expressible) by Q_2 if the formulas used to interpret (express) are independent of U and depend on n alone. Let Q^{1-1}_λ be the existential quantifier ranging over 1–1 partial functions with domain $\leq \lambda$. Formally in the finite case we have:

THEOREM 1.2. *Let U be a finite universe, and R a n-place relation on U. Then there exist a natural number $\lambda = \lambda(R)$, and equivalence relation E on U such that uniformly we have*:

1. *Q_E and Q^{1-1}_λ are interpretable by Q_R.*
2. *If $|U| \geq \lambda^n$ then Q_R is expressible by $\{Q_E, Q^{1-1}_{\lambda^n}\}$.*
3. *If $|U| < \lambda^n$ then every binary relation on a subset $A \subseteq U$ with cardinality $\leq |U|^{1/2n}$ is interpretable by Q_R.*

In case (2) of the theorem if we want to have "interpretable" instead of "expressible" then the situation is more complicated and we deal with it in this paper. Since U is a "large" universe we check the "asymptotic behavior", that is we consider a class $\mathfrak{U}$ of finite universes with unbounded cardinality. For each $U \in \mathfrak{U}$ let $\mathfrak{R}[U] \subseteq {}^nU$ be a n-place relation on U. We will see that there is a dichotomy in the behavior of $Q_{\mathfrak{R}[U]}$, that relates to cases (1) and (2) of theorem 1.2. Formally we prove:

THEOREM 1.3. *Let $\mathfrak{R}$ be as above. Then exactly one of the following conditions holds*:

1. *$Q_{\mathfrak{R}[U]}$ is uniformly interpretable by 1–1 functions and unary relations both of bounded cardinality.*
2. *For each $m \in \mathbb{N}$, there exist $U \in \mathfrak{U}$ such that we can uniformly interpret number theory up to m, by $Q_{\mathfrak{R}[U]}$.*

We prove this theorem in sections 3 to 6. In section 3 we analyze the situation, and give a condition for the dichotomy. In section 4 we prove that if the condition of section 3 holds then part (2) of theorem 1.3 is satisfied. In section 5 we prove, for the binary case that if the condition does not hold then part (1) of the theorem is satisfied. In section 6 we prove the same for the n-place case. In section 2 we show that in the finite case we can not get a full understanding of $Q_{\mathfrak{K}}$ similar to what we have in the countable case (not even for expressibility).

1.2. Preliminaries.

Notation 1.4. 1. $\mathfrak{U}$ is a class of finite universes, possibly with repetitions. So formally: $\mathfrak{U} = \{U_i : i \in \mathfrak{I}\}$ for an index class $\mathfrak{I}$ and we allow $U_i = U_j$ for $i \neq j \in \mathfrak{I}$. We will usually not be so formal and will write $U \in \mathfrak{U}$ and it should be understood as $i \in \mathfrak{I}$ and $U = U_i$. We assume $sup\{|U| : U \in \mathfrak{U}\} = \aleph_0$.
2. $\mathfrak{K}$ is a function on $\mathfrak{U}$ and for all $U \in \mathfrak{U}$, $\mathfrak{K}[U]$ is a set of n-place relations on U (where $n = n(\mathfrak{K})$ is a natural number), closed under permutations of U. This means: if $R_1, R_2 \subseteq {}^nU$ and $(U, R_1) \approx (U, R_2)$ then $R_1 \in \mathfrak{K}[U] \Leftrightarrow R_2 \in \mathfrak{K}[U]$.
3. $\overline{\mathfrak{K}}$ is a sequence of such functions. We write $\overline{\mathfrak{K}} = (\mathfrak{K}_0, \dots, \mathfrak{K}_{lg(\overline{\mathfrak{K}})-1})$.
4. $\mathfrak{R}$ is a function on $\mathfrak{U}$ and for each $U \in \mathfrak{U}$, $\mathfrak{R}[U]$ is a n-place relation over U (where $n = n(\mathfrak{R})$ is a natural number).
5. r is a $n(\mathfrak{R})$-place relation symbol.
6. For all $U \in \mathfrak{U}$ if S is a n-place relation on U, and F is a m-place function on U, then s and f are a n-place relation symbol and a m-place function symbol respectively. We write $(U, S) \models s(\overline{a})$ iff $\overline{a} \in s$, and $(U, F) \models f(\overline{b}) = c$ iff $F(\overline{b}) = c$. (That is for all $c \in U, \bar{a} \in {}^nU, \bar{b} \in {}^mU$).
7. For all $U \in \mathfrak{U}$ and $n \in \omega$, $\bar{a} \in {}^nU$ is a sequence of n elements in U. We write: $\bar{a} = (a_0, \dots, a_{n-1})$, and $lg(\bar{a}) = n$.

Definition 1.5. For all $\mathfrak{K}$ as in 1.4.2 we define the second order existential quantifier $Q_{\mathfrak{K}}$ to range over all relations in $\mathfrak{K}$. Formally we define the logic $L(Q_{\mathfrak{K}_1}, \dots, Q_{\mathfrak{K}_m})$ to be first order logic but we allow formulas of the form $(Q_{\mathfrak{K}_i} r)\varphi(r)$ (r is a $n(\mathfrak{K}_i)$-place relation symbol) for all $1 \leq i \leq m$. Satisfaction is defined only for models with universe $U \in \mathfrak{U}$ as follows: $\models (Q_{\mathfrak{K}_i} r)\varphi(r)$ iff there exists $R^0 \in \mathfrak{K}_i[U]$ such that $(U, R^0) \models \varphi(r)$.

Definition 1.6. We say that $\mathfrak{K}$ (or $Q_{\mathfrak{K}}$) is definable in some logic $\mathfrak{L}$ iff there exists a formula $\varphi(r) \in \mathfrak{L}$ (r is a $n(\mathfrak{K})$-place relation symbol) such that for all $U \in \mathfrak{U}$ and $R \subseteq {}^{n(\mathfrak{K})}U$:

$$(U, R) \models \varphi(r) \Longleftrightarrow R \in \mathfrak{K}[U].$$

Notation 1.7. For $\mathfrak{R}$ as in 1.4.4 we denote by $Q_{\mathfrak{R}}$ the quantifier $Q_{\mathfrak{K}_{\mathfrak{R}}}$ where $\mathfrak{K} = \mathfrak{K}_{\mathfrak{R}}$ is defined by:

$$\mathfrak{K}[U] := \{R^1 \subseteq {}^{n(\mathfrak{R})}U : (U, R^1) \approx (U, \mathfrak{R}[U])\}.$$

Definition 1.8. 1. We say that $Q_{\mathfrak{K}_1}$ is interpretable by $Q_{\mathfrak{K}_2}$ and write $Q_{\mathfrak{K}_1} \leq_{int} Q_{\mathfrak{K}_2}$ if there exist $k^* \in \omega$ and first order formulas:

$$\varphi_k(\overline{x}, \overline{r}) = \varphi_k(x_0, \dots, x_{n(\mathfrak{K}_1)-1}, r_0, \dots, r_{m-1})$$

for $k < k^*$ (each r_l is a $n(\mathfrak{K}_2)$-place relation symbol) and the following holds:

Sh:801

(∗) For all $U \in \mathfrak{U}$ and $R \in \mathfrak{K}_1[U]$ there exists $k < k^*$ and $R_0, \dots, R_{m-1} \in \mathfrak{K}_2[U]$ such that $(U, R_0, \dots, R_{m-1}) \models (\forall \overline{x})[R(\overline{x}) \equiv \varphi_k(\overline{x}, r_0, \dots, r_{m-1})]$.

2. We say that $Q_{\mathfrak{K}_1}$ is expressible by $Q_{\mathfrak{K}_2}$ and write $Q_{\mathfrak{K}_1} \leq_{exp} Q_{\mathfrak{K}_2}$ if there exist $k^* \in \omega$ and formulas in the logic $L(Q_{\mathfrak{K}_2})$:

$$\varphi_k(\overline{x}, \overline{r}) = \varphi_k(x_0, \dots, x_{n(\mathfrak{K}_1)-1}, r_0, \dots, r_{m-1})$$

for $k < k^*$ (each r_l is a $n(\mathfrak{K}_2)$-place relation symbol) and (∗) holds.
3. In (1) and (2) if $k^* = 1$ we write $Q_{\mathfrak{K}_1} \leq_{1-int} Q_{\mathfrak{K}_2}$ and $Q_{\mathfrak{K}_1} \leq_{1-exp} Q_{\mathfrak{K}_2}$ respectively.
4. We write $Q_{\mathfrak{K}_1} \equiv_{int} Q_{\mathfrak{K}_2}$ if $Q_{\mathfrak{K}_1} \leq_{int} Q_{\mathfrak{K}_2}$ and $Q_{\mathfrak{K}_2} \leq_{int} Q_{\mathfrak{K}_1}$. $\equiv_{exp}$ is defined in the same way.
5. We define $Q_{\mathfrak{K}} \leq_{int} \{Q_{\mathfrak{K}_0}, \dots, Q_{\mathfrak{K}_{l-1}}\}$ as in (1) only in (∗) for each $0 \leq j \leq m-1$, R_j may belong to some $\mathfrak{K}_i$ for $0 \leq i \leq l-1$ with $n(\mathfrak{K}_i) = n(R_j)$. We write $Q_{\overline{\mathfrak{K}}} = \{Q_{\mathfrak{K}_0}, \dots, Q_{\mathfrak{K}_{lg(\overline{\mathfrak{K}})-1}}\}$ when $\overline{\mathfrak{K}} = (\mathfrak{K}_0, \dots, \mathfrak{K}_{lg(\overline{\mathfrak{K}})-1})$. In the same way we define for $\leq_{exp}$.
6. We define $Q_{\overline{\mathfrak{K}^1}} \leq_{int} Q_{\overline{\mathfrak{K}^2}}$ if $Q_{\mathfrak{K}^1_i} \leq_{int} Q_{\overline{\mathfrak{K}^2}}$ for all $i < lg(\overline{\mathfrak{K}^1})$ again when $\overline{\mathfrak{K}^1} = (\mathfrak{K}^1_0, \dots, \mathfrak{K}^1_{lg(\overline{\mathfrak{K}^1})-1})$. In the same way we define for $\leq_{exp}$.

The following two lemmas are straightforward.

Lemma 1.9. 1. *$\leq_{int}$ and $\leq_{exp}$ are partial orders, and hence $\equiv_{int}$ and $\equiv_{exp}$ are equivalence relations on the class of quantifiers of the form $Q_{\mathfrak{K}}$.*
2. *$Q_{\overline{\mathfrak{K}^1}} \leq_{int} Q_{\overline{\mathfrak{K}^2}}$ implies $Q_{\overline{\mathfrak{K}^1}} \leq_{exp} Q_{\overline{\mathfrak{K}^2}}$.* ⊣

Lemma 1.10. *Let $\mathfrak{L}$ be some logic and assume $\overline{\mathfrak{K}^1}, \overline{\mathfrak{K}^2}$ are definable in $\mathfrak{L}$ (that is every $\mathfrak{K}^l_i$ is, see definition 1.6) and $Q_{\overline{\mathfrak{K}^1}} \leq_{exp} Q_{\overline{\mathfrak{K}^2}}$ then:*

1. *There exists a computable function that attach to every formula in $\mathfrak{L}(Q_{\overline{\mathfrak{K}^1}})$ an equivalent formula in $\mathfrak{L}(Q_{\overline{\mathfrak{K}^2}})$.*
2. *The set of valid sentences in $\mathfrak{L}(Q_{\overline{\mathfrak{K}^1}})$ is recursive from the set of valid sentences in $\mathfrak{L}(Q_{\overline{\mathfrak{K}^2}})$.* ⊣

So $\leq_{exp}$ gives a hierarchy on on logics of the form $\mathfrak{L}(Q_{\overline{\mathfrak{K}}})$, i.e., under the assumptions of lemma 1.10 the expressive power of $\mathfrak{L}(Q_{\overline{\mathfrak{K}^2}})$ is at least as strong as that of $\mathfrak{L}(Q_{\overline{\mathfrak{K}^1}})$.

1.3. Summation of previous results. We will use the following results. Proofs can be found in [6].

Definition 1.11. 1. let λ be a function from $\mathfrak{U}$ to $\mathbb{N}$ such that $\lambda[U] \leq |U|/2$. Define $\mathfrak{K}^{mon}_\lambda$ by $\mathfrak{K}^{mon}_\lambda[U] := \{A \subseteq U : |A| = \lambda[U]\}$. We denote $Q_{\mathfrak{K}^{mon}_\lambda}$ by Q^{mon}_λ.
2. For λ as above define $\mathfrak{K}^{mon}_{\leq\lambda}$ by $\mathfrak{K}^{mon}_{\leq\lambda}[U] := \bigcup\{\mathfrak{K}^{mon}_\mu : \mu \leq \lambda\}$. We denote $Q_{\mathfrak{K}^{mon}_{\leq\lambda}}$ by $Q^{mon}_{\leq\lambda}$.
3. For λ as above define $\mathfrak{K}^{1-1}_\lambda$ by $\mathfrak{K}^{1-1}_\lambda[U] := \{f : U \to U : |Dom(f)| = \lambda[U], f \text{ one to one}\}$. We denote $Q_{\mathfrak{K}^{1-1}_\lambda}$ by Q^{1-1}_λ.
4. For λ as above define $\mathfrak{K}^{1-1}_{\leq\lambda}$ by $\mathfrak{K}^{1-1}_{\leq\lambda}[U] := \bigcup\{\mathfrak{K}^{1-1}_\mu : \mu \leq \lambda\}$. We denote $Q_{\mathfrak{K}^{1-1}_{\leq\lambda}}$ by $Q^{1-1}_{\leq\lambda}$.

Sh:801

5. Let λ and μ be functions from $\mathfrak{U}$ to $\mathbb{N}$. Define $\mathfrak{K}^{eq}_{\lambda,\mu}$ as follows: $\mathfrak{K}^{eq}_{\lambda,\mu}[U]$ is the collection of all equivalence relations on subsets of U with exactly $\lambda[U]$ classes, and the size of each class is $\mu[U]$. We denote $Q_{\mathfrak{K}^{eq}_{\lambda,\mu}}$ by $Q^{eq}_{\lambda,\mu}$.
6. Let λ and μ be as in (5). Define $\mathfrak{K}^{eq}_{\leq\lambda,\leq\mu}$ as follows: $\mathfrak{K}^{eq}_{\leq\lambda,\leq\mu}$ is the collection of all equivalence relations on subsets of U with at most $\lambda[U]$ classes, and the size of each is at most $\mu[U]$. We denote $Q_{\mathfrak{K}^{eq}_{\leq\lambda,\leq\mu}}$ by $Q^{eq}_{\leq\lambda,\leq\mu}$.

Remark 1.12. It is an easy fact that $Q^{mon}_{\lambda} \equiv_{int} Q^{mon}_{\leq\lambda}$ and $Q^{1-1}_{\lambda} \equiv_{int} Q^{1-1}_{\leq\lambda}$ so we will usually not distinguish between them.

LEMMA 1.13. *Let λ be a function from $\mathfrak{U}$ to $\mathbb{N}$, and $\mathfrak{E}$ a binary relation on $\mathfrak{U}$ such that for all $U \in \mathfrak{U}$, $\mathfrak{E}[U]$ is an equivalence relation with at least $\lambda[U]$ classes each of which has at least $\lambda[U]$ elements (and possibly smaller classes). Then $Q^{eq}_{\lambda,\lambda} \leq_{int} Q_{\mathfrak{E}}$.*

PROOF. Straight foreword. The interpreting formula is

$$\varphi(x, y, s_0, s_1, s_2) := s_0(x, y) \wedge \neg s_1(x, y) \wedge s_2(x, y).$$

(See [6] for similar proofs). ⊣

THEOREM 1.14. *For every $\mathfrak{R}$ there exists a function $\lambda_0 = \lambda_0(\mathfrak{R})$ from $\mathfrak{U}$ to $\mathbb{N}$ and a relation $\mathfrak{R}_1$ with $n = n(\mathfrak{R}) = n(\mathfrak{R}_1)$ and $|Dom(\mathfrak{R}_1[U])| \leq \lambda_0[U] + n$ for all $U \in \mathfrak{U}$, such that $Q_{\mathfrak{R}} \equiv_{int} \{Q_{\mathfrak{R}_1}, Q^{mon}_{\lambda_0}\}$.*

The interpretation is done uniformly, that is the formulas used are independent of $\mathfrak{R}$ (depend on $n(\mathfrak{R})$ alone).

THEOREM 1.15. *For every $\mathfrak{R}$ there exists a function $\lambda_1 = \lambda_1(\mathfrak{R})$ from $\mathfrak{U}$ to $\mathbb{N}$ such that uniformly: $Q_{\mathfrak{R}} \equiv_{int} \{Q^{mon}_{\lambda_0}, Q^{1-1}_{\lambda_1}, Q_{\mathfrak{R}_1}, Q_{\mathfrak{E}}\}$, where: $n = n(\mathfrak{R}) = n(\mathfrak{R}_1)$, λ_0 is given by* 1.14, *and for all $U \in \mathfrak{U}$, $|Dom(\mathfrak{R}_1[U])| \leq n \cdot \lambda_1[U]$ and $\mathfrak{E}[U]$ is an equivalence relation on U.*

Remark 1.16. In the proof of theorem 3.6 we can assume without loss of generality that for all $U \in \mathfrak{U}$, $|\mathfrak{R}[U]| \leq \lambda_0[U] + n(\mathfrak{R})$, this is true since we can interpret $\mathfrak{R}_1$ instead of $\mathfrak{R}$. Similarly using 1.15 we can assume $|\mathfrak{R}[U]| \leq \lambda_1[U] \cdot n(\mathfrak{R})$. Here we have an equivalence relation $\mathfrak{E}$ that can change the bounds but the change will not be significant. Note also that $Q^{1-1}_{\lambda_1} \equiv_{int} Q^{1-1}_{n\cdot\lambda_1}$ (for all $n \in \omega$). So in the simple case of the dichotomy (theorem 5.2) we prove $Q_{\mathfrak{R}} \leq_{int} \{Q^{mon}_{\lambda_0}, Q^{1-1}_{\lambda_1}\}$ but in the proof we will not pay attention to the size of the sets and functions we use.

§2. Limitations on the classification of $Q_{\mathfrak{K}}$ in the finite.

In this section we show that unlike the countable case in which we had an understanding of $Q_{\mathfrak{K}}$ by equivalence relations, in the finite case there are classes of relations we can not express.

DEFINITION 2.1. For all $n \in \omega$ define $\mathfrak{K}_n$ by: $\mathfrak{K}_n[U] := \{R : R \subseteq {}^nU\}$ for all $U \in \mathfrak{U}$.

LEMMA 2.2. *For all $n \in \omega$: $Q_{\mathfrak{K}_{n+1}} \not\leq_{exp} Q_{\mathfrak{K}_n}$.*

PROOF. Suppose $Q_{\mathfrak{K}_{n+1}} \leq_{exp} Q_{\mathfrak{K}_n}$, and assume that the formulas used for expressing are $\varphi_k(\bar{x}, r_0, \ldots, r_{m_k-1})$ for $k < k^*$. Put $m = \max\{m_k : k < k^*\} \cup \{k^*\}$ and let $U \in \mathfrak{U}$. Then by these formulas we can express at most $m \cdot |\mathfrak{K}[U]|^m$ different relations. Since $|\mathfrak{K}_n[U]| = 2^{|U|^n}$, if we choose U such that $|U| > m$, we get $|U|^n(|U| - m) > \log_2(m)$, hence $2^{|U|^n(|U|-m)} > m$, and hence $2^{|U|^{n+1}} > m \cdot 2^{|U|^n \cdot m}$.

So the maximal number of different expressible relations is smaller than $|\mathfrak{K}_{n+1}[U]|$, a contradiction. $\dashv$

We have that for $n \geq 2$, $Q_{\mathfrak{K}_n}$ is not expressible by equivalence relations, unlike the countable case (see 1.1). Moreover we have:

LEMMA 2.3. *For all $n \geq 2$:*

1. $Q_{\mathfrak{K}_n} \not\leq_{exp} Q^{1\text{-}1}$.
2. $Q_{\mathfrak{K}_n} \not\leq_{exp} Q^{eq}$.

PROOF. We prove (1). again suppose $Q_{\mathfrak{K}_n} \leq_{exp} Q^{1\text{-}1}$, and we use the notations of the previous proof. Note that $\mathfrak{K}^{1\text{-}1}[U] = |U|!$, and for $|U|$ large enough we have $|U|! < 2^{|U| \cdot log(|U|) \cdot c}$ where c is some constant. Moreover for all $n \geq 2$, for $|U|$ large enough we have $|U| \cdot \log(|U|) \cdot cm < |U|^n$. So we get: $m \cdot |\mathfrak{K}^{1\text{-}1}[U]|^m < 2^{|U|^n}$ which means the number of relations expressible is smaller than $|\mathfrak{K}_n[U]|$, a contradiction.

The proof of (2) is similar using: $\mathfrak{K}^{eq}[U] \leq |U|^{|U|} \leq 2^{|U| \cdot \log(|U|) \cdot c}$. $\dashv$

We get that in the finite case even for $n(\mathfrak{K}) = 2$, we can not express every $Q_{\mathfrak{K}}$ by 1–1 functions and equivalence relations.

§3. Primary analysis.

From here on, unless said otherwise, we assume that $\mathfrak{R}$ is fixed and $\lambda_i = \lambda_i(\mathfrak{R})$ for $i \in \{0, 1\}$ (see 1.14 and 1.15).

In this section we start the analysis of $Q_{\mathfrak{R}}$. For each universe U we define a natural number k which is the maximal size, in some sense, of an equivalence relation on U interpretable by $\mathfrak{R}[U]$. The size of k is an indicator of the degree of "complexity" of $\mathfrak{R}$. We will show that there is a dichotomy, ether $\mathfrak{R}$ is very "complex" or it is "simple". This is made precise below.

DEFINITION 3.1. Let $\tau = \{f_0, \dots, f_{m_1}, s_0, \dots, s_{m_2}, c_0, \dots, c_{m_3}\}$ be a vocabulary, that is f_i are $n(f_i)$-place function symbols, s_i are $n(s_i)$-place relation symbols and c_i are individual constants. Define:

1. for all $U \in \mathfrak{U}$ a model for τ on U is
$$M = (U, f_0^M, \dots, f_{m_1}^M, s_0^M, \dots, s_{m_2}^M, c_0^M, \dots, c_{m_3}^M),$$
where f_i^M are $n(f_i)$-place **partial** functions.
2. a model for τ on $\mathfrak{U}$ denoted by $\mathfrak{M}$ is a function from $\mathfrak{U}$ such that for all $U \in \mathfrak{U}$, $\mathfrak{M}[U]$ is a model for τ on U. Note that the function $U \mapsto (U, \mathfrak{R}[U])$ is a model for $\{r\}$ on $\mathfrak{U}$, we will not be as formal and say that $\mathfrak{R}$ is.
3. Assume $r \in \tau$. We say that $\mathfrak{M}$ expands (or is an expansion of) $\mathfrak{R}$ if for all $U \in \mathfrak{U}$, $r^{\mathfrak{M}[U]} = \mathfrak{R}[U]$. More generally:
4. Let $\tau \subseteq \tau'$ be vocabularies, and let $\mathfrak{M}$ and $\mathfrak{M}'$ be models on $\mathfrak{U}$ for τ and τ' respectively. We say that $\mathfrak{M}'$ expands $\mathfrak{M}$ if $\mathfrak{M}'|\tau = \mathfrak{M}$. That means for all $U \in \mathfrak{U}$ and $f_i \in \tau$, $f_i^{\mathfrak{M}'[U]} = f_i^{\mathfrak{M}[U]}$, and similarly for relation symbols and constants.
5. We call τ **simple** if it is finite and all the relation and function symbols are unary.
6. We call M a simple model for τ on U if:
   (a) τ is simple.
   (b) M is a model for τ on U.
   (c) For all $i \leq m_1$, f_i^M is a one to one function and $|Dom(f_i^M)| \leq \lambda_1[U]$.
   (d) For all $i \leq m_2$, $|s_i^M| \leq \lambda_0[U]$.

Sh:801

7. We call $\mathfrak{M}$ a simple model for τ on $\mathfrak{U}$ if for all $U \in \mathfrak{U}$, $\mathfrak{M}[U]$ is a simple model for τ on U.
8. Let $U \in \mathfrak{U}$ and R be $n(\mathfrak{R})$-place relation on U. We call M a simple expansion of R on U for vocabulary τ if:
   (a) $r \in \tau$.
   (b) M is a model for τ on U.
   (c) $r^M = R$.
   (d) The restriction of M to $\tau \setminus \{r\}$ is a simple model for $\tau \setminus \{r\}$ on U. In particular $\tau \setminus \{r\}$ is a simple vocabulary.
9. We call $\mathfrak{M}$ a simple expansion of $\mathfrak{R}$ (on $\mathfrak{U}$) for vocabulary τ, if for all $U \in \mathfrak{U}$, $\mathfrak{M}[U]$ is a simple expansion of $\mathfrak{R}[U]$ for τ on U.

DEFINITION 3.2. Let τ be a finite vocabulary, and Δ a set of formulas in τ. Let M be a model for τ on U, $m \in \omega$, $A \subseteq U$, and $\bar{a} \in {}^m U$. Define:

1. The Δ-type of $\bar{a}$ over A in M is:

$$tp_\Delta(\bar{a}, A, M) := \{\varphi(\bar{x}, \bar{b}) : \varphi(\bar{x}, \bar{y}) \in \Delta, lg(\bar{x}) = m, \bar{b} \in {}^{<\omega}A, M \models \varphi(\bar{a}, \bar{b})\}.$$

2. $S^m_\Delta(A, M) := \{tp_\Delta(\bar{a}, A, M) : \bar{a} \in {}^m U\}$. if $M = (U, R)$ we write $S^m_\Delta(A, R)$, and similarly in (1).
3. If $p \in S^m_\Delta(A, M)$, $\overline{a'} \in {}^m U$ and $\varphi(\bar{x}, \bar{b}) \in p \Rightarrow M \models \varphi(\overline{a'}, \bar{b})$, then we say that $\overline{a'}$ realizes p, in particular $\bar{a}$ realizes $tp_\Delta(\bar{a}, A, M)$.

DEFINITION 3.3. Let τ be a finite vocabulary and Δ a set of formulas in τ.

1. For all $U \in \mathfrak{U}$, $A \subseteq U$ and M a model for τ on U, define an equivalence relation $E = E^{\Delta, M}_{A, U}$ (we usually write $E^{\Delta, M}_A$ where U is understood) on U by:

$$E := \{(x', x'') \in {}^2U : tp_\Delta(x', A, M) = tp_\Delta(x'', A, M)\}.$$

2. Let $U \in \mathfrak{U}$, $m \in \omega$ and E an equivalence relation on U. We call E m-big, if E has at least m equivalence classes of size at least m. If E is not m-big we say it is m-small.
3. Let $\mathfrak{M}$ be a model for τ over on $\mathfrak{U}$, define a function from $\mathfrak{U}$ to $\mathbb{N}$, $k_\Delta = k_{\Delta, \mathfrak{M}}$ as follows: $k_\Delta[U]$ is the maximal number k such that there exists $A \subseteq U$, $|A| \leq \lambda_0[U]$, and $E^{\Delta, \mathfrak{M}[U]}_A$ is k-big.

LEMMA 3.4. *Let $\mathfrak{M}$ be a simple expansion of $\mathfrak{R}$ for a vocabulary τ, and Δ a finite set of formulas in τ. Then*: $\{Q_{\mathfrak{R}}, Q^{mon}_{\lambda_0}, Q^{1-1}_{\lambda_1}\} \geq_{int} Q^{eq}_{k_{\Delta, \mathfrak{M}}, k_{\Delta, \mathfrak{M}}}$.

PROOF. For all $U \in \mathfrak{U}$, let $A_U \subseteq U$ be the subset the existence of which is promised by 3.3.3. Let s' be an unary relation symbol. Define a simple vocabulary $\tau' := \tau \cup \{s'\}$, and a formula in τ':

$$\psi(x', x'') := (\forall \bar{b}) \bigwedge_{\varphi(x, \bar{y}) \in \Delta} \{s'(\bar{b}) \rightarrow [\varphi(x', \bar{b}) \equiv \varphi(x'', \bar{b})]\}$$

(where $(\forall \bar{b})$ stands for $\forall b_0 \ldots \forall b_{lg(\bar{b})-1}$, and $s'(\bar{b})$ stands for $\bigwedge_{i < lg(\bar{b})-1} s'(b_i)$). Let $\mathfrak{M}'$ be the simple expansion of $\mathfrak{M}$ for τ' defined by $s'^{\mathfrak{M}'[U]} := A_U$, for all $U \in \mathfrak{U}$. Then for all $U \in \mathfrak{U}$ and $a, b \in U$:

$$a E^{\Delta, \mathfrak{M}[U]}_{A_U} b \iff \mathfrak{M}'[U] \models \psi(a, b).$$

Define $\mathfrak{E}$ by $\mathfrak{E}[U] = E^{\Delta,\mathfrak{M}[U]}_{A_U,U}$. Since $\mathfrak{M}'$ is a simple expansion of $\mathfrak{N}$ we have $Q_{\mathfrak{E}} \leq_{int} \{Q_{\mathfrak{N}}, Q^{mon}_{\lambda_0}, Q^{1-1}_{\lambda_1}\}$ when the interpreting formula is ψ. Now by 1.13 we have $Q_{\mathfrak{E}} \geq_{int} Q^{eq}_{k_{\Delta,\mathfrak{M}},k_{\Delta,\mathfrak{M}}}$, so by transitivity of $\leq_{int}$ we are done. ⊣

LEMMA 3.5. *Let n be a natural number no larger than $n(\mathfrak{N})$. Let τ be a simple vocabulary, and $\mathfrak{M}$ a simple expansion of $\mathfrak{N}$ for $\tau \cup \{r\}$. Let Δ be a finite set of formulas in $\tau \cup \{r\}$, of the form $\varphi(x, \bar{y})$ such that $lg(\bar{y}) \leq n$. Let $U \in \mathfrak{U}$ and $k = k_{\Delta,\mathfrak{M}}[U]$. Then there exists $A \subseteq U$ such that:*

1. *$|A| \leq nk$.*
2. *If $\varphi(x, \bar{y}) \in \Delta$ and $\bar{a} \in {}^{lg(\bar{y})}U$ are a formula and parameters, then the formula $\varphi(-, \bar{a})$ divides every equivalence class of $E^{\Delta,\mathfrak{M}[U]}_{A}$ into two parts one of which has no more than $(k+1) \cdot 2^{m^*}$ elements, where $m^* = |\Delta|(k+1)^{n+1} \cdot n^n$.*
3. *There exists at most k types, $p \in S^1_\Delta(A, \mathfrak{M}[U])$, realized by at least $k \cdot 2^{m^*}$ elements of U each.*

PROOF. Define a natural number m_l by downward induction on $l \leq k+1$: $m_{k+1} = 0$, $m_l = |\Delta|(n(l+1))^n + m_{l+1}$. By induction on $l \leq k+1$ we try to build a set $A_l \subseteq U$ such that $|A_l| \leq n * l$, and there exists at least l types $p \in S^1_\Delta(A_l, \mathfrak{M}[U])$ realized by at least $(k+1) * 2^{m_l}$ elements each. If we succeed then the existence of A_{k+1} is a contradiction to the definition of k. (We assume here that $|A_{k+1}| \leq \lambda_0[U]$, but without loss of generality we can assume that as $|A_{k+1}|$ is bounded. see also remark 1.16). Let $l_0 < k+1$ be such that we have built $A_0, \dots, A_{l_0}$ but we can not build A_{l_0+1}. Put $A = A_{l_0}$. Clearly A satisfies (1). We prove (2).

Put $M := \mathfrak{M}[U]$. Let $\langle B_i : i \leq l_0 \rangle$ be am enumeration of equivalence classes of $E^{\Delta,M}_A$ with at least $(k+1) * 2^{m_{l_0}}$ elements. (Note that there are exactly l_0 such classes since l_0 is maximal). Let $\varphi(x, \bar{y}) \in \Delta$ and $\bar{a} \in {}^{lg(\bar{y})}U$ be some formula and parameters. The relation $E^{\Delta}_{A \cup \bar{a}}$ divides every class B_i to at most $2^{|\Delta| * (|A|+n)^n}$ parts. Hence by the pigeon hole principle at least one of those parts has at least $\frac{|B_i|}{2^{|\Delta| * (|A|+n)^n}} \geq \frac{|B_i|}{2^{|\Delta| * (n(l_0+1))^n}} \geq (k+1) * 2^{m_{l_0+1}}$ elements. If for some i there are more than one part with more than $(k+1) * 2^{m_{l_0+1}}$ elements then define $A_{l_0+1} = A \cup \bar{a}$ and we get:

1. $|A_{l_0+1}| \leq |A_{l_0}| + |\bar{a}| \leq n * l_0 + n \leq n(l_0+1)$.
2. There exists at least $l_0 + 1$ types $p \in S^1_\Delta(A_{l_0+1}, \mathfrak{M}[U])$ realized by at least $(k+1) \cdot 2^{m_{l_0+1}}$ elements each.

This is a contradiction to the maximality of l_0. Now assume toward contradiction that $\varphi(-, \bar{a})$ divides some B_i into two parts, both larger than $(k+1) * 2^{m^*}$ (note that $m^* \geq m_{l_0}$ so there is no need to check classes smaller than $(k+1) * 2^{m^*}$). Then $E_{A \cup \bar{a}}$ divides each part into at most $2^{|\Delta| * (n(l_0+1))^n}$ classes and hence each part contains an equivalence class of $E_{A \cup \bar{a}}$ with at least $\frac{(k+1) * 2^{m^*}}{2^{|\Delta| * (n(l_0+1))^n}} \geq \frac{(k+1) * 2^{m_{l_0}}}{2^{|\Delta| * (n(l_0+1))^n}} = (k+1) * 2^{m_{l_0+1}}$ elements, so B_i contains two such classes and this, as we saw, is a contradiction. To prove (3) we note that $l_0 + 1 \leq k+1$, and $m^* \geq m_{l_0} \geq m_{l_0+1}$ hence the existence of $k+1$ classes with $k * 2^{m^*}$ elements contradicts the maximality of l_0. ⊣

THEOREM 3.6. *Let τ be a simple vocabulary, and Δ a finite set of formulas in $\tau \cup \{r\}$. Then one of the following conditions hold:*

1. *There exists a sequence of worlds: $\langle U_i \in \mathfrak{U} : i \in \omega \rangle$, and a sequence of natural numbers: $\langle n_i : i \in \omega \rangle$ such that $n_i \longrightarrow \infty$ and there exists a simple vocabulary*

τ', a formula $\varphi(x,y)$ in $\tau' \cup \{r\}$ and a simple expansion $\mathfrak{M}'$ of $\mathfrak{N}$ for $\tau' \cup \{r\}$, such that for all $i \in \omega$: $\{(x,y) \in {}^2U_i : \mathfrak{M}'[U_i] \models \varphi(x,y)\}$ is an n_i-big equivalence relation on U_i.

2. *There exists a natural number k^* such that for all $U \in \mathfrak{U}$ and M — a simple expansion of $\mathfrak{N}[U]$ for $\tau \cup \{r\}$ on U, there exist $A = A_U^{\Delta,M} \subseteq U$ such that $|A| \leq k^*$, $E_{A,U}^{\Delta,M}$ is k^*-small, and for every formula $\varphi(x,\bar{y}) \in \Delta$ and parameters $\bar{a} \in {}^{lg(\bar{y})}U$, $\varphi(-,\bar{a})$ divides each equivalence class of $E_A^{\Delta,M}$ into two parts one of which has less than k^* elements.*

PROOF. Define $\mathbb{M}$ to be the class of all simple expansions of $\mathfrak{N}$ for $\tau \cup \{r\}$ on $\mathfrak{U}$. For all $U \in \mathfrak{U}$ define:

$$k_\Delta^{max}[U] = \max\{k_{\Delta,\mathfrak{M}}[U] : \mathfrak{M} \in \mathbb{M}\}.$$

Next we assume that $sup\{k_\Delta^{max}[U] : U \in \mathfrak{U}\} = \aleph_0$ and show that condition (1) is satisfied. Let $\langle U_i \in \mathfrak{U} : i \in \omega \rangle$ be a sequence of universes such that $n_i = k_\Delta^{max}[U_i] \longrightarrow \infty$. Define a simple vocabulary $\tau' = \tau \cup \{s'\}$ (s' an unary relation symbol). We now define $\mathfrak{M}$. For all $i \in \omega$ denote by M_i the model for $\tau \cup \{r\}$ on U, for which the maximum in the definition of $k_\Delta^{max}[U_i]$ is obtained. Define $\mathfrak{M}[U_i]|\tau \cup \{r\} := M_i$. By the definition of $k_{\Delta,\mathfrak{M}}$ there exists a subset $A_i \subseteq U_i$ such that $E_{A_i}^{\Delta,M_i}$ is a n_i-big equivalence relation on U. Let $s'^{\mathfrak{M}[U_i]} = A_i$. That defines $\mathfrak{M}$ (obviously the definition on universes not among the U_i is irrelevant). We define $\varphi(x,y)$ to be the formula interpreting $E_{A_i}^{\Delta,M_i}$ (see 3.4) namely:

$$\varphi(x,y) := (\forall \bar{b}) \bigwedge_{\psi(x,\bar{z}) \in \Delta} \{s'(\bar{b}) \rightarrow [\psi(x,\bar{b}) \equiv \psi(y,\bar{b})]\}.$$

It is clear that condition (1) is satisfied.

We now assume that $\{k_\Delta^{max}[U] : U \in \mathfrak{U}\}$ is bounded and let k by its bound. we show that condition (2) is satisfied. Let $n := max\{lg(\bar{y}) : \varphi(x,\bar{y}) \in \Delta\}$. We define $k^* = \max\{(k+1) * 2^{|\Delta|(k+1)^{n+1} * n^n}, n(k+1)\}$. Now let $U \in \mathfrak{U}$, and M a simple expansion of $\mathfrak{N}[U]$ on U for vocabulary $\tau \cup \{r\}$. Let $A \subseteq U$ be the subset the existence of which is promised by the previous lemma. Then all the demands of (2) are clear from the previous lemma and the fact $k \geq k_\Delta^{max}[U] \geq k_{\Delta,\mathfrak{M}}[U]$. ⊣

§4. The complicated case of the dichotomy. In this section we assume that $\mathfrak{U}$ and $\mathfrak{N}$ satisfy condition (1) in 3.6, that is we can uniformly interpret an arbitrarily large equivalence relation. We show that in this case we can interpret bounded number theory in the logic $L(Q_\mathfrak{N})$. It follows that the logic $L(Q_\mathfrak{N})$ is undecidable.

We make use of the following:

LEMMA 4.1. *Let E be an n^2-big equivalence relation on a universe U. Then we can uniformly (that is using formulas independent of U and E) interpret the model $(\{0,1,\dots,n-1\}; 0, S, +, *)$ using a finite number of isomorphic copies of E.*

PROOF. We will not go in to details, as similar results are known from "history" (see [3]). In short, using a fixed number of of isomorphic n-big equivalence relations (actually two relations are enough), we can uniformly interpret any graph on n vertexes. Now the model $(\{0,1,\dots,n-1\}; 0, S, +, *)$ can be easily shown to be interpretable (again uniformly) by a graph with n^2 vertexes. ⊣

From this we conclude:

COROLLARY 4.2. *In theorem* 3.6 *if condition* (1) *is satisfied then we can uniformly interpret number theory bounded by* n_i *using a finite number of isomorphic copies of* $R[U_{i^2}]$. ⊣

We can now prove our undecidability result:

THEOREM 4.3. *In theorem* 3.6 *if condition* (1) *is satisfied then the logic* $L(Q_{\mathfrak{R}})$ *is undecidable, i.e., the set of logically valid sentences in* $L(Q_{\mathfrak{R}})$ *is not recursive.*

PROOF. Again we will not go into details. Let Ψ be a finite subset of the axioms of number theory large enough, so that we can code Turing machines in any model of Ψ. Let ψ^* be the conjunction of the axioms in Ψ. Now the set of sentences ψ in the vocabulary of number theory such that $\psi \wedge \psi^*$ has a finite model is not recursive. For each sentences ψ in the vocabulary of number theory let φ_ψ be the "translation" of $\psi \wedge \psi^*$ to the vocabulary $\{r_1, \dots, r_m\}$ under the interpreting formulas obtained from the previous corollary. We then have that $\psi \wedge \psi^*$ has a finite model iff $(Q_{\mathfrak{R}} r_1) \dots (Q_{\mathfrak{R}} r_m)\varphi_\psi$ has a model in the context of the logic $L(Q_{\mathfrak{R}})$. Hence the logic $L(Q_{\mathfrak{R}})$ can not be decidable. ⊣

§5. The simple case of the dichotomy. In this section we will interpret $Q_{\mathfrak{R}}$ when $\mathfrak{R}$ is "simple" that is when condition (1) in theorem 3.6 is not satisfied. We will show that in this case there exists a simple model on $\mathfrak{U}$ in which it is possible to interpret $\mathfrak{R}$ by a first order formula. In fact we prove $Q_{\mathfrak{R}} \leq_{int} \{Q^{mon}_{\lambda_0}, Q^{1-1}_{\lambda_1}\}$ so we get a full understanding of $Q_{\mathfrak{R}}$.

5.1. Formalizing the assumptions and the main theorem. In the rest of the paper we assume that $\mathfrak{U}$ and $\mathfrak{R}$ do not satisfy condition (1) in theorem 3.6. (Note that this condition is independent of Δ). Hence from that theorem we get the following:

1. For every simple vocabulary τ, and Δ a finite set of formulas in $\tau \cup \{r\}$, there exists a number $k_1^* = k_1^*(\Delta)$ and a function that assigns to every $U \in \mathfrak{U}$ and M — a simple expansion of $\mathfrak{R}[U]$ for $\tau \cup \{r\}$ on U, a set $A = A_U^{\Delta,M} \subseteq U$ such that condition (2) in theorem 3.6 is satisfied, that is:
   (∗) $|A| \leq k_1^*$, $E_A^{\Delta,M}$ is k_1^*-small, and for every formula $\varphi(x,\overline{y}) \in \Delta$ and parameters $\overline{a} \in {}^{lg(\bar{y})}U$, $\varphi(-,\overline{a})$ divides each equivalence class of $E_A^{\Delta,M}$ into two parts one of which has at most k_1^* elements.
2. For every simple vocabulary τ, and every formula $\varphi(x,y)$ in $\tau \cup \{r\}$, there exists a natural number $k_2^* = k_2^*(\varphi)$ such that:
   (∗∗) If $\mathfrak{M}$ is a simple expansion of $\mathfrak{R}$ for $\tau \cup \{r\}$ and $U \in \mathfrak{U}$, then the interpretation of $\varphi(x,y)$ in $\mathfrak{M}[U]$ (that is $\{(x,y) \in {}^2U : \mathfrak{M}[U] \models \varphi(x,y)\}$) is **not** a k_2^*-big equivalence relation.

Remark 5.1. We can increase $k_1^*(\Delta)$, meaning if $m \geq k_1^*(\Delta)$ than m satisfies (∗) (for the same function $A_U^{\Delta,M}$). Hence:

1. If we are given a function $\Delta \mapsto m(\Delta)$ then without loss of generality (by changing the definition of k_1^*) we may assume: $k_1^*(\Delta) \geq m(\Delta)$ for all Δ.
2. If $\Delta \subseteq \Delta'$ then without loss of generality (by redefining k_1^* by induction on $|\Delta|$) we may assume $k_1^*(\Delta') \geq k_1^*(\Delta)$.

We now formalize the main theorem of this section.

Sh:801

THEOREM 5.2. *There exists a simple vocabulary τ, and a first order formula $\varphi(x_0, \dots, x_{n(\mathfrak{R})-1})$ in τ, and there exists $\mathfrak{M}$ a simple expansion of $\mathfrak{R}$ for $\tau \cup \{r\}$ on $\mathfrak{U}$ such that for all $U \in \mathfrak{U}$:*

$$\mathfrak{M}[U] \models (\forall \bar{x})[r(\bar{x}) \equiv \varphi(\bar{x})].$$

COROLLARY 5.3. $Q_{\mathfrak{R}} \leq_{int} \{Q_{\lambda_1}^{1-1}, Q_{\lambda_0}^{mon}\}$.

PROOF. Straight from the theorem when the interpreting formula is φ. ⊣

In the rest of the paper we will prove theorem 5.2.

5.2. Proof of the main theorem in the binary case. We prove theorem 5.2 under the assumption $n(\mathfrak{R}) = 2$. Δ will be a finite set of formulas with at most 2 free variables in the vocabulary $\{r\}$. In other words $\tau = \emptyset$. Hence the set $A_U^{\Delta,M}$ and the relation $E_A^{\Delta,M}$ are independent of M, and depend on Δ alone so they will be denoted by A_U^{Δ} and E_A^{Δ}.

DEFINITION 5.4. Let Δ be as above and $U \in \mathfrak{U}$. Let $k^* = k_1^*(\Delta)$ and $A = A_U^{\Delta}$, define:

1. For all $\varphi(x, y) \in \Delta$ and $y_0 \in U$:

$$Minority_{\Delta}(y_0, \varphi) := \{x_0 \in U : |\{x \in U : xE_A^{\Delta}x_0 \wedge \varphi(x, y_0) \equiv \varphi(x_0, y_0)\}| \leq k^*\}.$$

2. $S = S^{\Delta}$ is the binary relation on U given by:

$$x_0 S y_0 \Leftrightarrow x_0 \in \bigcup_{\varphi(x,y)\in\Delta} Minority_{\Delta}(y_0, \varphi).$$

LEMMA 5.5. *Let Δ be as above. We use the notations of the previous definition and also $k_2^* = k_2^*(\psi)$ where $\psi(x', x'') := (\forall \bar{b}) \bigwedge_{\varphi(x,\bar{y})\in\Delta}\{s(\bar{b}) \rightarrow [\varphi(x', \bar{b}) \equiv \varphi(x'', \bar{b})]\}$ (s an unary relation symbol). Then:*

(1) $|\{x : |x/E_A^{\Delta}| \leq 2 \cdot k^*\}| \leq l^*$, *where* $l^* = k^* \cdot 2^{|\Delta|k^*+1}$.

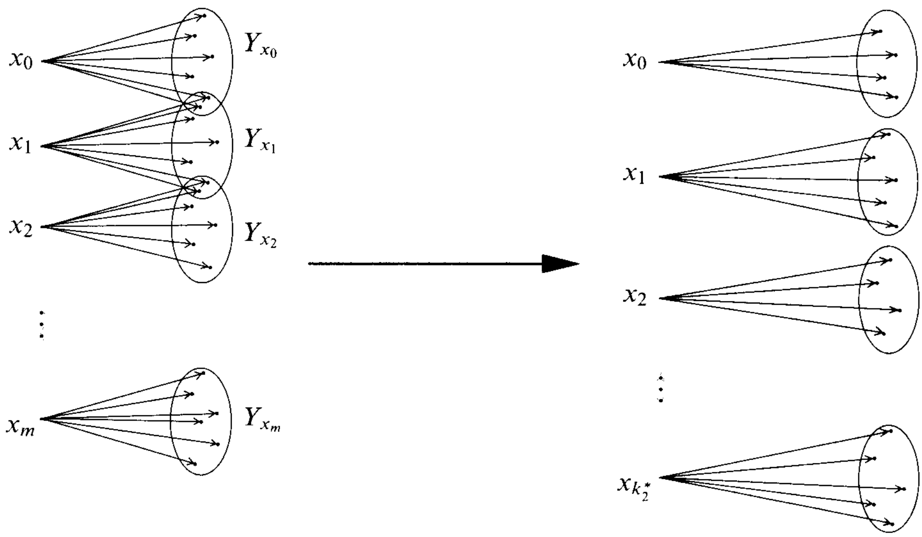
(2) *For all* $y \in U$: $|\{x : xSy\}| \leq |\Delta| \cdot (k^*)^2 + l^*$.

(3) $\left|\left\{x : |\{y : xSy\}| > 2^{|\Delta|(k^*+k_2^*)} \cdot k_2^* + l^*\right\}\right| \leq |\Delta| \cdot (k_2^* \cdot k^*)^2 \cdot 2^{|\Delta|(k^*+k_2^*)} + l^*$.

PROOF. (1): The number of types $p \in S_{\Delta}^1(A, \mathfrak{R}[U])$ is no larger than $2^{|\Delta||A|}$ since for every formula in Δ there are at most two free variables. We also have $|A| \leq k^*$. So the number of equivalence classes of E_A^{Δ} is no larger than $2^{|\Delta|k^*}$ and (1) follows directly.

(2): Let $x, y \in U$. Assume $|x/E_A^{\Delta}| > 2 \cdot k^*$. For all $\varphi \in \Delta$ we have $x/E_A^{\Delta} \cap Minority_{\Delta}(y, \varphi) \leq k^*$. Hence $|\{x' : xE_A^{\Delta}x' \wedge x'S^{\Delta}y\}| \leq |\Delta| \cdot k^*$. The number of equivalence classes of E_A^{Δ} which are larger than $2 \cdot k^*$ is also no larger than k^*. Hence we get: $|\{x : |x/E_A^{\Delta}| > 2 \cdot k^* \wedge xS^{\Delta}y\}| \leq |\Delta| \cdot (k^*)^2$. To this we add at most l^* elements from "small classes" and (2) follows.

(3): We write $m = |\Delta| \cdot (k_2^* \cdot k^*)^2 \cdot 2^{|\Delta|(k^*+k_2^*)}$. First we disregard all the elements of $\{x : |x/E_A| \leq 2 \cdot k^*\}$ and using (1) we decrease the bounds by l^*. So seeking a contradiction we assume that there are different $\{x_0, \dots, x_m\}$ so that for each $i \leq m$ there exists different $\{y_0^i, \dots, y^i_{2^{|\Delta|(k^*+k_2^*)} \cdot k_2^*}\}$ with $x_i S y_j^i$. Using (2) (with the bounds in (2) also decreased by l^*) there exists a subset of $\{x_0, \dots, x_m\}$ with at least k_2^* elements such that the elements of $Y_{x_i} := \{y_j^i : x_i S y_j^i\}$ are pairwise disjoint (see figure). Without loss of generality we assume that this set is $\{x_0, \dots, x_{k_2^*}\}$. For



every x_i the sets of Y_{x_i} satisfy at most $2^{|\Delta|(k^*+k_2^*)}$ different types $p \in S^1_\Delta(A \cup \{x_i : i \leq k_2^*\}, \mathfrak{R}[U])$. Hence there are more than k_2^* elements of Y_{x_i} that satisfy the same type (again we assume those are the first elements). In conclusion we get $\{x_i\}_{i=0}^{k_2^*}$ and $\{y_j^i\}_{i,j=0,\dots,k_2^*}$ without repetitions such that the type $tp_\Delta(y_j^i, A \cup \{x_i : i \leq k_2^*\}, \mathfrak{R}[U])$ is independent of j, and $x_{i_1} S y_j^{i_2} \Leftrightarrow i_1 = i_2$ holds for all $i_1, i_2, j \leq k_2^*$. So ψ with s taken to represent $A \cup \{x_i : i \leq k_2^*\}$ interprets a k_2^*-big equivalence relation on U. This is a contradiction to the definition of k_2^*.

⊣

LEMMA 5.6. *There exist a simple vocabulary* τ, *and a finite set of formulas* Φ *in* τ, *and a simple model* $\mathfrak{M}$ *for* τ *on* $\mathfrak{U}$, *such that for all* $U \in \mathfrak{U}$ *and* $x, x', y, y' \in U$ *if* $tp_\Phi((x,y), \emptyset, \mathfrak{M}[U]) = tp_\Phi((x',y'), \emptyset, \mathfrak{M}[U])$ *then* $(U, \mathfrak{R}[U]) \models r(x,y) \equiv r(x',y')$.

PROOF. We simultaneously define τ and its interpretation $\mathfrak{M}[U]$ for some $U \in \mathfrak{U}$. Φ will be the set of atomic formulas in τ with terms of the form $x, f(x), c, f(c)$ (function composition is not allowed). For brevity we write: $M := \mathfrak{M}[U]$ and $R := \mathfrak{R}[U]$. Let $\Delta := \{r(x,y)\}$. Using the notations of 5.5 we define:

$$A^* = A \cup \{x : |x/E_A^\Delta| \leq 2 \cdot k^*\} \cup \left\{x : \left|\left\{y : xS^\Delta y\right\}\right| > 2^{|\Delta|(k^*+k_2^*)} \cdot k_2^* + l^*\right\}.$$

By 5.5 $|A^*|$ is uniformly bounded (that is the bound is independent of U). τ will contain: private constants for all the elements of A^* ($\{c_x : x \in A^*\}$, $c_x^M := x$), and unary relation symbols for the equivalence classes of $E_{A^*}^\Delta$ ($\{s_{x/E_{A^*}^\Delta} : x \in U\}$, $s_{x/E_{A^*}^\Delta}^M := x/E_{A^*}^\Delta$). Note that the number of such classes is also uniformly bounded. Now we look at $S^\Delta | U \setminus A^*$ this is a digraph with (uniformly) bounded degree, that is for all $x \in U \setminus A^*$, $|\{y \notin A^* : xS^\Delta y\}|$ is bounded by $2^{|\Delta|(k^*+k_2^*)} \cdot k_2^* + l^*$ and for all $y \in U \setminus A^*$, $|\{x \notin A^* : xS^\Delta y\}|$ is bounded by $|\Delta| * (k^*)^2 + l^*$. Hence we can divide $S^\Delta | U \setminus A^*$ into $\langle S_m : m < m^* \rangle$ with: $\bigcup_{m<m^*} S_m = S^\Delta | U \setminus A^*$ and for all $m < m^*$, S_m is a digraph with degree 1, that is a one to one partial function on $U \setminus A^*$. To see this inductively apply Hall's theorem to the elements of the largest degree. Note

Sh:801

that m^* is uniformly bounded, in fact it is bounded by the sum of the two bounds mentioned above. We add to τ, unary function symbols $\{f_m : m < m^*\}$ and define $f_m^M := S_m$.

Let $\langle B_i : i < i^* \rangle$ be an enumeration of $\{x/E_A^\Delta \setminus A^* : |x/E_A^\Delta| > 2 \cdot k^*\}$. Note that $i^* \leq k^*$. For all $y \in U$ and $i < i^*$ there is a truth value t_i^y that is the value the formula $r(-, y)$ gets for the **majority** of the elements of B_i. Since we deal with "big" classes (that is with more than $2 \cdot k^*$ elements) we get: for all $y \in U$, $i < i^*$ and $x \in B_i$, $R(x, y) = t_i^y \Leftrightarrow \neg x S^\Delta y$. We divide each B_i into 2^{i^*} parts according to the truth values, $t_i^y : i < i^*$. This means that for each part, the value of the vector $\langle t_i^y : i < i^* \rangle$ is independent of y. For all $i < i^*$, we denote these parts by $\langle B_j^i : j < 2^{i^*} \rangle$. We add to τ, unary relations $\{s_{i,j} : i < i^*, j < 2^{i^*}\}$ and define $s_{i,j}^{\mathfrak{M}[U]} := B_j^i$. This completes the definition of τ and $\mathfrak{M}$.

We now prove that $\mathfrak{M}$ is as desired. Let $a, a', b, b' \in U$ and assume

$$tp_\Phi((a, b), \emptyset, \mathfrak{M}[U]) = tp_\Phi((a', b'), \emptyset, \mathfrak{M}[U]).$$

If $a \in A^*$ then $a = a'$ (due to the formula $x = c_x$), and the truth value of $R(a, b)$ is determined by $b/E_{A^*}^\Delta$. Moreover $b/E_{A^*}^\Delta = b'/E_{A^*}^\Delta$ (due to the formula $s_{b/E_{A^*}^\Delta}(y)$), so we get $R(a, b) = R(a', b')$ as desired. Symmetrically we deal with the cases $b, b'a' \in A^*$. So we can assume $a, a', b, b' \notin A^*$. By the definition of the functions S_m we have:

$$aS^\Delta b \iff (\exists m < m^*) a S_m b,$$
$$a' S^\Delta b' \iff (\exists m < m^*) a' S_m b'.$$

But due to the formulas of the form $f_m(x) = y$, the right hand side of both equations is equivalent, so we have $aS^\Delta b \Leftrightarrow a'S^\Delta b'$. Assume $a \in B_{j_1}^{i_1}$, $b \in B_{j_2}^{i_2}$. Due to the formula $s_{i,j}(x)$ we get $a' \in B_{j_1}^{i_1}$, $b' \in B_{j_2}^{i_2}$. By the construction of the B_i^j we get:

$$R(a, b) = t_{i_1}^b \iff \neg a S^\Delta b,$$
$$R(a', b') = t_{i_1}^{b'} \iff \neg a' S^\Delta b'.$$

But $b, b' \in B_{j_2}^{i_2}$ so $t_{i_1}^b = t_{i_1}^{b'}$, and as we have seen $aS\Delta b \Leftrightarrow a'S^\Delta b'$. Hence $R(a, b) = R(a', b')$ as desired. ⊣

COROLLARY 5.7. *Theorem 5.2 is true for the case* $n(\mathfrak{R}) = 2$.

PROOF. Let τ, Φ and $\mathfrak{M}$ be as in the previous lemma. For every set of types $D \subseteq S_\Phi^2(\emptyset, \mathfrak{M}[U]))$ we can easily write a formula $\chi_D(x, y)$ in τ such that $\mathfrak{M}[U] \models \chi_D(x, y)$ iff (x, y) satisfies one of the types in D. For all $U \in \mathfrak{U}$ let D_U be the collection of types $tp_\Phi((x, y), \emptyset, \mathfrak{M}[U])$ such that $(U, \mathfrak{R}[U]) \models r(x, y)$. Using the previous lemma it is easy to verify that for all $U \in \mathfrak{U}$ and $x, y \in U$ we have:

$$(U, \mathfrak{R}[U]) \models r(x, y) \iff (U, \mathfrak{M}[U]) \models \chi_{D_U}(x, y).$$

We now add to τ constants: $\{c_{true}\} \cup \{c_D : D \subseteq S_\Phi^2(\emptyset, \mathfrak{M}[U]))$. For each $U \in \mathfrak{U}$, c_{true} is interpreted in $\mathfrak{M}[U]$ by some element of U. The rest of the constants are interpreted so that for all $D \subseteq C$: $(c_D^{\mathfrak{M}[U]} = c_{true}^{\mathfrak{M}[U]}) \Leftrightarrow (D = D_U)$ holds. (assuming U has more than one element there is no problem to do that). Now the desired formula in theorem 5.2 is:

$$\varphi(x,y) := \bigwedge_{D\subseteq C} [(c_D = c_{true}) \to \chi_D(x,y)].$$

⊣

§6. Proof of the main theorem in the general case. We prove theorem 5.2 when $n(\mathfrak{R}) > 2$. From here on we assume:

> τ is a simple vocabulary. Δ is a finite set of formulas in $\tau \cup \{r\}$, such that $\varphi(\bar{x}) \in \Delta \longrightarrow lg(\bar{x}) \leq n(\mathfrak{R})$.

First we generalize definition 5.4.

DEFINITION 6.1. Let τ, Δ be as above. Let $U \in \mathfrak{U}$ and M be a simple expansion of $\mathfrak{R}[U]$ on U for $\tau \cup \{r\}$. Let $n < n(\mathfrak{R})$. We denote $k^* = k_1^*(\Delta)$ and $A = A_U^{\Delta,M}$ the existence of which follows from 5.1.1 and define:

1. For all $\varphi(x,\bar{y}) \in \Delta$ with $lg(\bar{y}) = n$ and $\bar{b} \in {}^nU$:

$$Minority_{\Delta,M}(\bar{b},\varphi) := \{x \in U : |\{x' \in U : xE_A^{\Delta,M}x' \wedge \varphi(x,\bar{b}) \equiv \varphi(x',\bar{b})\}| \leq k^*\}$$

2. Define a relation $S^n_{\Delta,M} \subseteq U \times {}^nU$:

$$aS^n_{\Delta,M}\bar{b} \Leftrightarrow a \in \bigcup\{Minority_{\Delta,M}(\bar{b},\varphi) : \varphi(x,\bar{y}) \in \Delta, lg(\bar{y}) = n\}$$

Remark 6.2. For $i \in \{1,2\}$ assume τ_i, Δ_i satisfy the assumption above, and $\mathfrak{M}_i$ is a simple expansion of $\mathfrak{R}$ on $\mathfrak{U}$ for $\tau_i \cup \{r\}$. Furthermore assume $\tau_1 \subseteq \tau_2$, $\Delta_1 \subseteq \Delta_2$ and $\mathfrak{M}_1 = \mathfrak{M}_2|\tau_1$. By 5.1.2 we may assume $k_1^*(\Delta_2) \geq k_1^*(\Delta_1)$, hence for all $U \in \mathfrak{U}$ we can assume without loss of generality (we can add elements to $A_U^{\Delta_2,\mathfrak{M}_2[U]}$ if needed) that $aS^n_{\Delta_2,\mathfrak{M}_2[U]}\bar{b} \Longrightarrow aS^n_{\Delta_1,\mathfrak{M}_1[U]}\bar{b}$.

LEMMA 6.3. *Using the notations of the previous definition*:

1. $\left|\left\{x : \left|x/E_A^{\Delta,M}\right| \leq 2\cdot k^*\right\}\right| \leq k^* \cdot 2^{|\Delta|\binom{k^*}{n(\mathfrak{R})}+1}$.
2. *For all* $\bar{b} \in {}^nU$: $\left|\left\{x \in U : xS^n_{\Delta,M}\bar{b}\right\}\right| \leq |\Delta|\cdot(k^*)^2 + k^*\cdot 2^{|\Delta|\binom{k^*}{n(\mathfrak{R})}+1}$.

PROOF. Similar to the proof of 5.5, only in (1) we have at most $\binom{k^*}{n(\mathfrak{R})}$ different choices of parameters for each formula. ⊣

Notation 6.4. Using the notations above we put: $l^* = l^*(\Delta) := |\Delta|\cdot(k^*)^2 + k^*\cdot 2^{|\Delta|\binom{k^*}{n(\mathfrak{R})}+1}$.

LEMMA 6.5 (Symmetry Lemma (with Parameters)). *Assume* τ, Δ *satisfy* 6, *and Let* $\mathfrak{M}$ *be a simple expansion of* $\mathfrak{R}$ *for* $\tau \cup \{r\}$. *Let* $n < n(\mathfrak{R})$. *Then there exists a simple vocabulary* $\tau' \supseteq \tau$, *and* $\mathfrak{M}'$ *a simple expansion of* $\mathfrak{M}$ *for* $\tau' \cup \{r\}$, *and for* $i \in \{1,2\}$ *there exists* $\Delta_i = \Delta_i(\Delta)$ *such that* τ', Δ_i *also satisfy the assumption above, and for all* $U \in \mathfrak{U}$, $a, b \in U$ *and* $\bar{c} \in {}^{n-1}U$:

$$aS^n_{\Delta,\mathfrak{M}[U]}b\bar{c} \Longrightarrow (aS^{n-1}_{\Delta_1,\mathfrak{M}'[U]}\bar{c}) \vee (bS^n_{\Delta_2,\mathfrak{M}'[U]}a\bar{c}).$$

Sh:801

PROOF. First we define a few constants we will use later: $m^* := k_2^*(\phi)$ (see assumption 5.1), where ϕ is the following formula in $\tau \cup \{s, c_1, \dots, c_{n-1}\} \cup \{r\}$:

$$\phi = \phi(y, y') := (\forall x) \bigwedge_{\psi(x,y,z_1,\dots,z_{n-1})\in\Delta} \{s(x) \to [\psi(x, y, c_1, \dots, c_{n-1}) \equiv \psi(x, y', c_1, \dots, c_{n-1})]\}$$

(s is an unary relation symbol and $c_1, \dots, c_{n-1}$ are constants not in τ). We also define: $m_1 = m_1(\Delta) := (m^*)^2 \cdot 2^{|\Delta|m^*} \cdot l^*(\Delta)$ and $m_2 = m_2(\Delta) := m^* \cdot 2^{|\Delta|m^*}$.

Let $\mathfrak{M}', \tau'$ and $\psi(x, x')$ be the vocabulary, model and formula which interpret $E_A^{\Delta,M}$ (see the proof of 3.4). We define in τ' a formula that will interpret $xS^n_{\Delta,\mathfrak{M}[U]}y\bar{z}$ in $\mathfrak{M}'[U]$ (where $lg(\bar{z}) = n-1$):

$$\chi(x, y, \bar{z}) := \bigvee_{\varphi(u,v,\bar{w})\in\Delta, lg(\bar{w})=n-1} (\exists^{\le k_1^*(\Delta)} x')[\psi(x, x') \wedge (\varphi(x, y, \bar{z}) \equiv \varphi(x', y, \bar{z}))].$$

We therefore get:

(*) for all $U \in \mathfrak{U}$, $a, b \in U$ and $\bar{c} \in {}^{n-1}U$: $\mathfrak{M}'[U] \models \chi(a, b, \bar{c}) \iff aS^n_{\Delta,\mathfrak{M}[U]}b\bar{c}$.

Define:

$$\chi'(x, \bar{z}) := (\exists^{>m_2} y)\chi(x, y, \bar{z}),$$
$$\Delta_1 := \Delta \cup \{\chi'(x, \bar{z})\},$$
$$\Delta_2 := \Delta \cup \{\chi(x, y, \bar{z})\}.$$

Note that by 5.1.1 we may assume that $k_1^*(\Delta) \ge \max\{m_1(\Delta), m_2(\Delta)\}$, and by 5.1.2 we may assume $k_1^*(\Delta_i) \ge k_1^*(\Delta) \ge m_i(\Delta)$ for $i \in \{1, 2\}$. We now assume toward contradiction that there exists $U \in \mathfrak{U}$, $a, b \in U$ and $\bar{c} \in {}^{n-1}U$ such that:

1. $aS^n_{\Delta,\mathfrak{M}[U]}b\bar{c}$.
2. $\neg(aS^{n-1}_{\Delta_1,\mathfrak{M}'[U]}\bar{c})$.
3. $\neg(bS^n_{\Delta_2,\mathfrak{M}'[U]}a\bar{c})$.

From (3) and $k_1^*(\Delta_2) \ge m_2$ we can find $\{b_0, \dots, b_{m_2}\}$ without repetitions such that for all $i \le m_2$: $\mathfrak{M}'[U] \models \chi(a, b, \bar{c}) \equiv \chi(a, b_i, \bar{c})$. from (1) and (*) we get that for all $i \le m_2$: $\mathfrak{M}'[U] \models \chi(a, b_i, \bar{c})$. Hence $\mathfrak{M}'[U] \models \chi'(a, \bar{c})$. from (2) and $k_1^*(\Delta_1) \ge m_1$ we can find $\{a_0, \dots, a_{m_1}\}$ without repetitions such that for all $i \le m_1$: $\mathfrak{M}[U] \models \chi'(a, \bar{c}) \equiv \chi'(a_i, \bar{c})$. We have seen that $\mathfrak{M}'[U] \models \chi'(a, \bar{c})$ so by the definition of $\chi'(x, \bar{z})$ we have for all $i \le m_1$, there exists $\{b_0^i, \dots, b_{m_2}^i\}$ without repetitions such that $i \le m_1 \wedge j \le m_2 \Rightarrow a_i S^n_{\Delta,\mathfrak{M}[U]} b_j^i \bar{c}$. By the definition of $l^*(\Delta)$ and a repeated use of the pigeon hole principle we can find a subset of $\{a_0, \dots, a_{m_1}\}$, $\{a_{i_0}, \dots, a_{i_{m^*}}\}$ such that the sets $\{\{b_0^{i_l}, \dots, b_{m_2}^{i_l}\} : l \le m^*\}$ are pairwise disjoint. without loss of generality we assume $i_l = l$ for all $l \le m^*$. Using the pigeon hole principle again we can find for all $i \le m^*$ subset of $\{b_0^i, \dots, b_{m_2}^i\}$ with $m^* + 1$ elements (and again we assume this subset is $\{b_0^i, \dots, b_{m^*}^i\}$) such that for all $\varphi(x, y, \bar{z}) \in \Delta$ and $j_1, j_2 \le m^*$ we have $\varphi(a_i, b_{j_1}^i, \bar{c}) \Leftrightarrow \varphi(a_i, b_{j_2}^i, \bar{c})$. In conclusion we got: $\{a_0, \dots, a_{m^*}\}$ without repetitions and for each $i \le m^*$: $\{b_0^i, \dots, b_{m^*}^i\}$ without repetitions such that $a_{i_1} S^n_\Delta b_j^{i_2} \bar{c} \Leftrightarrow i_1 = i_2$. Moreover the elements of $\{b_0^i, \dots, b_{m^*}^i\}$ satisfy the same formulas of the form $\varphi(a_i, y, \bar{c}) \in \Delta$ ($\bar{c}$ and a_i are parameters). Now the formula $\phi(y, y')$ (where s is taken to mean

$\{a_0, \dots, a_{m^*}\}$ and the constants c_i are taken to mean the elements c_i) interprets a $m^* + 1$-big equivalence relation on $\{y^i_j : i, j \leq m^*\}$. This is a contradiction to the definition of m^*. ⊣

We now prove a number of lemmas we need for the proof of the main theorem. First we show that we can code a delta system of n-tuples by singletons.

Definition 6.6. A set of n-tuples, $\{a^i \in {}^n A : i < i^*\}$ (A some set and i^* some natural number), is called a delta system if, there exists some $w \subseteq \{0, \dots, n-1\}$ such that: $|\{a^i_t : i < i^*\}| = 1$ for all $t \in w$, and $|\{a^i_t : i < i^*\}| = i^*$ for all $t \notin w$.

Lemma 6.7. *Let n be a natural number. Then there exists a simple vocabulary τ, and a formula $\theta(x, \bar{y})$ in τ with $lg(\bar{y}) = n$ such that: for all $U \in \mathfrak{U}$ and delta system $\langle \overline{a^i} \in {}^n U : i < i^* \rangle$, we have a simple model M for τ on U and a sequence $\langle b_i \in U : i < i^* \rangle$ such that:*

$$(\forall \bar{a} \in {}^n U)(\forall b \in U)\big[M \models \theta(b, \bar{a})\big] \textit{ iff } (\exists i < i^*)\big(b = b_i \wedge \bar{a} = \overline{a^i}\big).$$

Proof. Define $\tau = \{c^*_0, \dots, c^*_n, c_1, \dots, c_n, s_0, s_1, f_1, \dots, f_n\}$. For each $n \geq t \geq 0$ define the formulas:

$$\theta_t(x, \bar{y}) := y_0 = c_0 \wedge \dots \wedge y_t = c_t \wedge y_{t+1} = x \wedge$$
$$y_{t+2} = f_{t+2}(x) \wedge \dots \wedge y_n = f_n(x).$$
$$\theta(x, \bar{y}) := s_1(x) \wedge \bigwedge_{n \geq t \geq 0} \big[s_0(c^*_t) \rightarrow \theta_t(x, \bar{y})\big].$$

Now let $U \in \mathfrak{U}$ and assume $\langle \overline{a^i} \in {}^n U : i < i^* \rangle$ is a delta system. For simplicity assume we have some $n > t^* \geq 0$, such that: $|\{a^i_t : i < i^*\}| = 1$ for all $0 \leq t \leq t^*$, and $|\{a^i_t : i < i^*\}| = i^*$ for all $n > t > t^*$. (We can prove the lemma in the general case of a delta system but this makes the definition of θ more complicated). We can now define M:

$c^{*M}_0 \dots c^{*M}_n$ are some distinct elements of U (we assume $|U| > n$).

$c^M_t = a^1_t$ (for $1 \leq t \leq t^*$ and assuming $t^* > 0$ otherwise the definition of c^M_t is insignificant).

$s^M_0 := \{c^{*M}_{t^*}\}$.

$s^M_1 := \{a^i_{t^*+1} : i < i^*\}$ (assuming $t^* < n$ otherwise define s^M_1 to be some singleton).

$f^M_t := \{(a^i_{t^*+1}, a^i_t) : i < i^*\}$ (for $t^* + 1 < t \leq n$ and assuming $t^* + 1 < n$ otherwise the definition of f^M_t is insignificant).

Note that f^M_t are one to one functions in the relevant cases. In conclusion we define $\langle b_i = a^i_{t^*+1} : i < i^* \rangle$ (again we assume $t^* < n$ otherwise we define $\langle b_i \in U : i < i^* \rangle$ to be some constant sequence). So by our definitions we get $M \models \theta_{t^*}(b_i, \overline{a^i})$ for all $i < i^*$. Moreover if $M \models \theta_{t^*}(b, \bar{a})$ then there exists $i < i^*$ such that $b = b_i$ and $\bar{a} = \overline{a^i}$. Hence θ, M and $\langle b_i : i < i^* \rangle$ are as needed. ⊣

We now show that it is impossible to interpret a large order relation on $\mathfrak{U}$.

Lemma 6.8. *Let τ_0 be a simple vocabulary, and $\varphi(\bar{x}, \bar{y})$ a formula in $\tau_0 \cup \{r\}$ (not assuming $lg(\bar{x}) = lg(\bar{y})$). Then there exists a natural number $k^* = k^*_3 = k^*_3(\varphi)$ such that for every $\mathfrak{M}$ a simple expansion of $\mathfrak{N}$ for $\tau_0 \cup \{r\}$, and for all $U \in \mathfrak{U}$, it is* *impossible to find sequences* $\langle \overline{a_i} \in {}^{lg(\bar{x})}U : i < k^* \rangle$ *and* $\langle \overline{b_j} \in {}^{lg(\bar{y})}U : j < k^* \rangle$ *such that*:

$$(\forall i, j < k^*)\big[\mathfrak{M}[U] \models \varphi(\overline{a_i}, \overline{b_j}) \iff i \leq j\big].$$

PROOF. Let $\varphi(\bar{x}, \bar{y})$ and τ_0 be as described. For $i \in \{1,2\}$ let τ_i, θ_i be the vocabulary and formula used to code delta systems for $n = lg(\bar{x})$ and $n = lg(\bar{y})$ respectively (i.e., those from the previous lemma). Add to τ_0 new unary relation symbol and function symbol, s^*, f^*. In the vocabulary $\tau = \tau_0 \cup \tau_1 \cup \tau_2 \cup \{s^*, f^*\} \cup \{r\}$ define the formula:

$$\phi(v, v') := (\neg\exists u)\langle s^*(u) \wedge (\forall \bar{x}, \overline{x'}, \overline{y}, \overline{y'})\{[\theta_1(u, \bar{x}) \wedge (\theta_1(f^*(u), \overline{x'}) \wedge \theta_2(v, \overline{y}) \wedge \theta_2(v', \overline{y'})] \rightarrow [\varphi(\overline{x}, \overline{y}) \equiv \varphi(\overline{x}, \overline{y'}) \wedge \varphi(\overline{x'}, \overline{y}) \equiv \neg\varphi(\overline{x'}, \overline{y'})]\}\rangle$$

which will interpret a large equivalence relation. For all $m, n \in \omega$ Let $Delta(n, m)$ denote the minimal number d such that every sequence of d n-tuples has a subsequence of length m which is a delta system. We can now define $k_3^*(\varphi)$:

$$k^* = k_3^*(\varphi) := Delta(lg(\bar{x}), Delta(lg(\bar{y}), (k_2^*(\phi))^2)).$$

Toward contradiction we assume that there exist $\mathfrak{M}_0$ a model for τ_0 on $\mathfrak{U}$, $U \in \mathfrak{U}$ and sequences as in the lemma. By the definition of k^* there exist subsequences of length $(k_2^*(\phi))^2$, which are delta systems. Put $k_2 := k_2^*(\phi)$. Without loss of generality we assume these subsequences are: $\langle \overline{a_i} \in {}^{lg(\bar{x})}U : i < (k_2)^2 \rangle$ and $\langle \overline{b_j} \in {}^{lg(\bar{y})}U : j < (k_2)^2 \rangle$. Let M_1, M_2, $\langle a_i \in U : i < (k_2)^2 \rangle$ and $\langle b_j \in U : j < (k_2)^2 \rangle$ be the models and sequences used to code $\langle \overline{a_i} : i < (k_2)^2 \rangle$ and $\langle \overline{b_j} : j < (k_2)^2 \rangle$ (see 6.7). We define M a model for τ on U:

For each $i \in \{0, 1, 2\}$: $M|\tau_i := M_i$.

$s^{*^M} := \{a_{j \cdot k_2} : j \in \{0, 1, 2, \ldots, k_2 - 1\}\}$.

$f^{*^M} := \{(a_{j \cdot k_2}, a_{((j+1) mod(k_2)) \cdot k_2}) : j \in \{0, 1, 2, \ldots, k_2 - 1\}\}$.

Note that if π is the permutation of $\{\overline{a_{i \cdot k}} \in {}^{lg(\bar{x})}U : j < k_2\}$ defined by $\pi(\overline{a_{j \cdot k_2}}) = a_{((j+1) mod(k_2)) \cdot k_2}$, then the formula:

$$\phi'(\bar{y}, \overline{y'}) := (\neg\exists \bar{x} \in \{\overline{a_{j \cdot k_2}} \in : j < k_2\})[\varphi(\overline{x}, \overline{y}) \equiv \varphi(\overline{x}, \overline{y'}) \wedge \varphi(\pi(\overline{x}), \overline{y}) \wedge \neg\varphi(\pi(\overline{x}), \overline{y'})]$$

interprets in M a k_2-big equivalence relation on $\{\overline{b_j} : j < (k_2)^2\}$ namely the relation $\{(\overline{b_i}, \overline{b_j}) : i, j \in (k_2)^2, \exists l \in \{0, \ldots, k_2 - 1\} \text{ s. t. } i, j \in [l, l+1, \ldots, l+k_2)\}$. Hence by the properties of θ_1 and θ_2, the formula $\phi(v, v')$ interprets a k_2-big equivalence relation on $\{b_j : j < (k_2)^2\}$ which is a contradiction. ⊣

We need one more lemma before we can prove the main theorem.

LEMMA 6.9. *Let* τ *be a simple vocabulary and* $\varphi(x, y, \bar{z})$ *a formula in* $\tau \cup \{r\}$. *Then there exist a natural number* $k^* = k_4^* = k_4^*(\varphi)$ *such that for every* $\mathfrak{M}$ *a simple expansion of* $\mathfrak{N}$ *for* $\tau \cup \{r\}$ *on* $\mathfrak{U}$, *and for all* $U \in \mathfrak{U}$, *it is impossible to find for each* $l < k^*$: $\overline{c^l} \in {}^{lg(\bar{z})}U$ *and sequences* $\langle a_i^l \in U : i < k^* \rangle$ *and* $\langle b_j^l \in U : j < k^* \rangle$ *such that*:

(α) *For all* $l, i, j < k^*$, $\mathfrak{M}[U] \models \varphi(a_i^l, b_j^l, \overline{c^l})$ *iff* $i = j$.

Sh:801

(β) *For all* $l_1 < l < k^*$, *the truth value of* $\varphi(a_i^{l_1}, b_j^{l_1}, \overline{c^l})$ *in* $\mathfrak{M}[U]$ *is independent of* $i, j < k^*$.

PROOF. Put $lg(\bar{z}) = n$. Let τ' and $\theta(x, \bar{y})$ be the vocabulary and formula we get by applying lemma 6.7 to n. Define a simple vocabulary $\tau^* := \tau \cup \tau' \cup \{s_1, s_2, f\}$, and formulas in τ^*:

$$\psi_1(x, y) := s_1(x) \wedge s_2(y) \wedge (\forall \bar{z})[\theta(y, \bar{z}) \rightarrow \varphi(x, f(x), \bar{z})],$$
$$\psi_2(x, x') := s_1(x) \wedge s_1(x') \wedge (\forall y \forall \bar{z})(s_2(y) \wedge \theta(y, \bar{z})) \rightarrow [\varphi(x, f(x), \bar{z}) \equiv \varphi(x', f(x'), \bar{z})].$$

Put $k' := \max\{k_3^*(\psi_1), k_2^*(\psi_2)\} + 1$. Let k^* be large compared to k', we will not give an exact definition of k^* but it should be clear that choosing k^* large enough will give a contradiction. Let $U \in \mathfrak{U}$ and M some simple expansion of $\mathfrak{N}[U]$ for $\tau \cup \{r\}$. Assume that for all $l < k^*$ there exist $\overline{c^l} \in {}^nU$ and sequences $\langle a_i^l \in U : i < k^* \rangle$ and $\langle b_j^l \in U : j < k^* \rangle$ satisfying (α) and (β). By choosing k^* large enough and using Ramsey theorem and condition (α) we can find $w \subseteq \{0, \ldots, k^* - 1\}$ of size $(k' + 2) \cdot k'$ such that:

- $\langle \overline{c^l} : l \in w \rangle$ is a delta system.
- The truth value of the sentences $a_i^{l_1} = a_j^{l_2}$ and $\varphi(a_i^{l_1}, b_j^{l_1}, \overline{c^l})$ (in M) for $l_1, l_2, i, j \in w$ depends only on the order type of the indexes.
- For all $l_1, l_2, i_1, i_2 \in w$, $a_{i_1}^{l_1} = a_{i_2}^{l_2} \Rightarrow (l_1 = l_2) \wedge (i_1 = i_2)$.
- For all $l_1, l_2, i_1, i_2 \in w$, $b_{i_1}^{l_1} = b_{i_2}^{l_2} \Rightarrow (l_1 = l_2) \wedge (i_1 = i_2)$.

Now using (β) exactly one of the following conditions hold: either $l_1 < l < k^* \Rightarrow M \models \varphi(a_i^{l_1}, b_j^{l_1}, \overline{c^l})$ or $l_1 < l < k^* \Rightarrow M \models \neg\varphi(a_i^{l_1}, b_j^{l_1}, \overline{c^l})$. We will deal with the first case (the second can be dealt with similarly). We have three cases:

1. There exist $w' \subseteq w$ of size k', and $i^* \neq j^* \in w'$ such that for all $l < l_1 \in w'$ we have: $M \models \neg\varphi(a_{i^*}^{l_1}, b_{j^*}^{l_1}, \overline{c^l})$.
2. There exist $w' \subseteq w$ of size k', and π a permutation of w' without fixed points such that for all $l < l_1 \in w'$ and for all $j \in w'$ we have: $M \models \varphi(a_j^{l_1}, b_{\pi(j)}^{l_1}, \overline{c^l})$.
3. Neither (1) nor (2) hold.

As stated above τ' and $\theta(x, \bar{y})$ are the vocabulary and formula we get by applying lemma 6.7 to n. Let M' and $\langle c^l \in U : l \in w \rangle$ be the model and sequence we get by applying that lemma (in U) to $\langle \overline{c^l} : l \in w \rangle$. For each of the cases (1)–(3) we define M^* a simple expansion of M for $\tau^* \cup \{r\}$ and get a contradiction. In each case $M^* | \tau' := M'$. The interpretation of s_1, s_2 and f will be given for each case separately:

Case (1). Define $s_1^{M^*} := \{c^l : l \in w'\}$, $s_2^{M^*} := \{a_{i^*}^l : l \in w'\}$ and $f^{M^*} := \{(a_{i^*}^l, b_{j^*}^l) : l \in w'\}$. Then we have $l > l_1 \iff M^* \models \varphi(a_{i^*}^{l_1}, f(a_{i^*}^{l_1}), \overline{c^l})$ and hence the formula $\psi_1(x, x')$ interprets in M^* an order relation (in the sense of 6.8) on $\{c^l : l \in w'\} \times \{a_{i^*}^l : l \in w'\}$. This is a contradiction as $|w'| = k'$ and k' is larger than $k_3^*(\psi_1)$.

Case (2). Define $s_1^{M^*} := \{c^l : l \in w'\}$, $s_2^{M^*} := \{a_i^l : i, l \in w'\}$ and $f^{M^*} := \{(a_i^l, b_{\pi(i)}^l : i, l \in w'\}$. Then we have $l \neq l_1 \iff M^* \models \varphi(a_i^{l_1}, f(a_i^{l_1}), \overline{c^l})$ and hence the formula $\psi_2(x, x')$ interprets in M^* the relation $\{(a^{l_1}_{i_1}, a^{l_2}_{i_2}) : l_1 = l_2 \wedge i_1, i_2 \in w'\}$ which is $|w'|$-big. This is a contradiction as $|w'| = k'$ and k' is larger than $k^*_2(\psi_2)$.

Case (3). Look at $\langle c^{(l+2)\cdot k'} : l < k' \rangle$, and the sequences $\langle a^{(l+2)\cdot k'}_{j\cdot k'} : j < k' \rangle$ and $\langle b^{(l+2)\cdot k'}_{j\cdot k'} : j < k' \rangle$ for $l < k'$. Since (1) does not hold for these sequences we get (choosing $i^* = 0$ and $j^* = 1$) that there exist $l^* < l^*_1$ such that

$$\varphi(a_0^{(l_1^*+2)\cdot k'}, b_{k'}^{(l_1^*+2)\cdot k'}, c^{(l^*+2)\cdot k'}).$$

In the same way (choosing $i^* = 1$ and $j^* = 0$) we get that there exist $l^{**} < l^{**}_1$ such that

$$\varphi(a_{k'}^{(l_1^{**}+2)\cdot k'}, b_0^{(l_1^{**}+2)\cdot k'}, c^{l^{**}}).$$

Now look at $\langle c^{2k'+l} : l < k' \rangle$ and the sequences $\langle a^{2k'+l}_j : j < k' \rangle$ and $\langle b^{2k'+l}_j : j < k' \rangle$ for $l < k'$. Let π be a permutation of $\{0, \dots, k'-1\}$ without a fixed point. We show that these sequences along with π, satisfy the demands of case (2). Let $j < k'$ and $l < l_1 < k'$. If $j < \pi(j)$ then $j < \pi(j) < 2k' + l < 2k' + l_1$ and $0 < k' < (l^* + 2) \cdot k' < (l^*_1 + 2) \cdot k'$. Since the truth value of φ depends only on the order type of the indexes we get $\varphi(a_j^{2k'+l_1}, b_{\pi(j)}^{2k'+l_1}, c^{2k'+l})$ (as $\varphi(a_0^{(l_1^*+2)\cdot k'}, b_{k'}^{(l_1^*+2)\cdot k'}, c^{(l^*+2)\cdot k'})$). If $\pi(j) < j$ we get the same result, only now we use the 4-tuple $0 < k' < (l^{**} + 2) \cdot k' < (l^{**}_1 + 2) \cdot k'$. In both cases we have $\varphi(a_j^{2k'+l_1}, b_{\pi(j)}^{2k'+l_1}, c^{2k'+l})$ as needed in (2). So case (3) can not hold. ⊣

We are now ready to prove theorem 5.2 in the general case. We prove:

THEOREM 6.10. *There exist: σ a simple vocabulary, $\varphi(\bar{x})$ a formula in σ with $lg(\bar{x}) = n(\mathfrak{R})$, and $\mathfrak{N}$ a simple model for σ on $\mathfrak{U}$. Such that for all $U \in \mathfrak{U}$ and $\bar{a} \in {}^{n(\mathfrak{R})}U$:*

$$\mathfrak{N}[U] \models \varphi(\bar{a}) \iff (U, \mathfrak{R}[U]) \models r(\bar{a}).$$

PROOF. We prove the theorem by induction on $n(\mathfrak{R})$. The cases $n(\mathfrak{R}) = 0$ and $n(\mathfrak{R}) = 1$ are trivial. the case $n(\mathfrak{R}) = 2$ was proved in 5.7.

Before we turn to the proof of the induction step we pay attention to the following fact. Let $\mathfrak{R}'$ be as in 1.4.4. We say that "$\mathfrak{R}'$ is definable from $\mathfrak{R}$ by a simple expansion" if there exist a simple vocabulary τ, a simple expansion $\mathfrak{M}$ of $\mathfrak{R}$ for $\tau \cup \{r\}$ and a formula $\varphi(x_0, \dots, x_{n(\mathfrak{R}')-1})$ in $\tau \cup \{r\}$ such that for all $U \in \mathfrak{U}$ and $\bar{a} \in {}^{n(\mathfrak{R}')}U$ we have $\mathfrak{R}'[U](\bar{a})$ iff $\mathfrak{M}[U] \models \varphi(\bar{a})$. Note that if $\mathfrak{R}'$ is definable from $\mathfrak{R}$ by a simple expansion then $\mathfrak{R}'$ also satisfies assumption 5.1 (or else $\mathfrak{R}$ does not satisfy the assumption for we can define a big equivalence relation from $\mathfrak{R}$ using φ and the model $\mathfrak{M}$). If $\mathfrak{R}'$ is definable from $\mathfrak{R}$ by a simple expansion and $n(\mathfrak{R}') < n(\mathfrak{R})$ then by the induction hypothesis there exist: σ_0 a simple vocabulary, $\varphi_0(\bar{x})$ a formula in σ_0 with $g(\bar{x}) = n(\mathfrak{R}')$, and $\mathfrak{N}_0$ a simple model for σ_0 on $\mathfrak{U}$. Such that for all $U \in \mathfrak{U}$ and $\bar{a} \in {}^{n(\mathfrak{R}')}U$:

$$\mathfrak{N}_0[U] \models \varphi_0(\bar{a}) \iff \mathfrak{R}'[U](\bar{a}).$$

In that case we will say that $\mathfrak{R}'$ satisfies the induction hypothesis and that σ_0, φ_0 and $\mathfrak{N}_0$ interpret it.

We now assume $n(\mathfrak{R}) = n + 1 > 2$. We prove this case in two stages. In the first stage we show that we can interpret the relation $xS^n_{\Delta,\mathfrak{M}}\bar{y}$, so we prove:

LEMMA 6.11. *Let Δ, τ be as in 6, and let $\mathfrak{M}$ be a simple expansion of $\mathfrak{N}$ for $\tau \cup \{r\}$ on $\mathfrak{U}$. Then there exist:*

- *A simple vocabulary σ_0 ($r \notin \sigma_0$).*
- *$\varphi_0(x, \bar{y})$ a formula in σ_0 ($lg(\bar{y}) = n$).*
- *$\mathfrak{N}_0$ a simple model for σ_0 on $\mathfrak{U}$.*

Such that for all $U \in \mathfrak{U}$, $a \in U$ and $\bar{b} \in {}^nU$ we have: $\mathfrak{N}_0[U] \models \varphi_0(a, \bar{b}) \iff aS^n_{\Delta,\mathfrak{M}[U]}\bar{b}$.

PROOF OF LEMMA 6.11. Let $\mathfrak{M}^*, \tau^*, \psi(x, x')$ and $\chi(x, y, \bar{z})$ (where $lg(\bar{z}) = n - 1$) be the vocabulary, model and formulas interpreting $E_A^{\Delta,\mathfrak{M}[U]}$ and $S^n_{\Delta,\mathfrak{M}[U]}$ that were defined in the proof of the symmetry lemma (denoted there by $\mathfrak{M}', \tau'$). We also define a formula that will interpret an order relation in $\mathfrak{M}^*$:

$$\phi = \phi(\bar{x}, \bar{y}) := [\chi(x_0, x_1, \bar{y}) \equiv \chi(x_2, x_3, \bar{y})]$$

where $lg(\bar{x}) = 4$ and $lg(\bar{y}) = n - 1$. In the vocabulary τ^* we define a set of formulas: $\Delta^* := \Delta \cup \{\chi, \phi\}$. For brevity we write $M := \mathfrak{M}[U]$, $M^* := \mathfrak{M}^*[U]$, $N_0 := \mathfrak{N}_0[U]$ and similarly for other models, where $U \in \mathfrak{U}$ is understood from the context. Next we define some constants that we will use in the proof:

1. $m_1 := m_1(\Delta) := \max\{k_3^*(\phi), k_4^*(\chi)\} + 1$ for the formulas χ, ϕ defined above.
2. $m_2 := m_2(\Delta) := (m_1)^2 + m_1$.
3. For all $U \in \mathfrak{U}$ choose by induction on $m_2 \geq l$, $A_l = A_l^U \subseteq U$ such that:
   (a) $A_0 = \emptyset$.
   (b) $A_l \subseteq A_{l+1}$ for all $l < m_2$.
   (c) For all $l < m_2$, $r \leq n + 2 \cdot m_1$ and a type $p \in S^r_{\Delta^*}(A_l, \mathfrak{M}^*[U])$: if p is realized in $\mathfrak{M}^*[U]$ then it is realized already in A_{l+1}.
   (d) For all $l < m_2$, $|A_{l+1}|$ is minimal under the properties (a)-(c).
4. We write $A^* = A^{*^U} = A^U_{m_2}$.
5. Note that under these conditions there exists a bound on $|A^*|$ depending only on $|\Delta^*|, m_1, m_2$ and n, so in fact the bound depends only on n and $|\Delta|$ and we can calculate it in the beginning of the proof. We denote this bound by m_3. We do not calculate the value of m_3 but note that it increases super-exponentially as a function of $|\Delta|$.
6. $m_4 := m_4(\Delta) := l^*(\Delta_1(\Delta)) + l^*(\Delta_2(\Delta)) \cdot m_1$ (see 5.5 and 6.5).
7. $m_5 := m_5(\Delta) = 2 \cdot m_4 + m_3 + n + 2$.

Denote by $\mathfrak{S} = \mathfrak{S}^n_{\Delta,\mathfrak{M}}$ the $n+1$-place relation on $\mathfrak{U}$ defined by $\mathfrak{S}[U] := \{(x, y, \bar{z}) \in {}^{n+1}U : xS^n_{\Delta,\mathfrak{M}[U]}y\bar{z}\}$. (We keep using the existing notation and write $x\mathfrak{S}[U]y\bar{z}$ instead of $\mathfrak{S}[U](x, y, \bar{z})$, or sometimes write $xS^n_{\Delta,\mathfrak{M}[U]}y\bar{z}$ as before). Our aim is to interpret the relation $\mathfrak{S}$ by a formula in a simple model on $\mathfrak{U}$. First note the following:

FACT. *Assume there exists a number i^* such that for all $U \in \mathfrak{U}$: ${}^{n+1}U = \bigcup_{i<i^*} B_i^U$. Assume farther that for all $i < i^*$ the relation $\mathfrak{S}_i$ defined by $\mathfrak{S}_i[U] := \mathfrak{S}[U] \cap B_i^U$ is interpreted by the formula φ_i and the simple model $\mathfrak{N}_i$ for the vocabulary σ_i. Then the formula $\bigvee_{i<i^*} \varphi_i(x, y, \bar{z})$ in the vocabulary $\bigcup_{i<i^*} \sigma_i$ and the model $\mathfrak{N}$ defined by $(\forall i < i^*)\mathfrak{N}|\sigma_i = \mathfrak{N}_i$ will interpret $\mathfrak{S}$ as needed.*

We return to the proof of the lemma. Let $\langle p_i : i < i^* \rangle$ be an enumeration of all the Δ^* types of two variables over a set of at most m_3 parameters. Formally this means each p_i is a subset of $\Phi := \{\varphi(x, y, u_{j_1}, \dots, u_{j_k}) \in \Delta^* : k < m_3, j_1, \dots, j_k \in \{0, \dots, m_3 - 1\}\}$. For all $U \in \mathfrak{U}$ fix $\langle a_0, \dots, a_l \rangle$ some enumeration of A^{*^U} (of course $l < m_3$) and we then write $tp_{\Delta^*}((a, b), A^*, \mathfrak{M}^*[U]) = p_i$ iff $\mathfrak{M}^*[U] \models \varphi(a, b, a_{j_1}, \dots, a_{j_k}) \Leftrightarrow \varphi(x, y, u_{j_1}, \dots, u_{j_k}) \in p_i$. Note that i^* is uniformly bounded by $2^{|\Delta^*| \cdot \binom{m_3}{n}}$. For all $i < i^*$ and $U \in \mathfrak{U}$ the binary relation on $\mathfrak{U}$ defined for all $U \in \mathfrak{U}$ by $\{(x, y) \in {}^2U : tp_{\Delta^*}((x, y), A^{*^U}, \mathfrak{M}^*[U]) = p_i\}$ satisfies the induction hypothesis. Hence there exist a simple vocabulary σ^i a formula $\varphi^i(x, y)$ and $\mathfrak{N}^i$ a simple model for σ^i on $\mathfrak{U}$ such that for all $U \in \mathfrak{U}$ and $a, b \in U$:

$$\mathfrak{N}^i \models \varphi^i(a, b) \Longleftrightarrow tp_{\Delta^*}((x, y), A^{*^U}, M^*[U]) = p_i.$$

Without loss of generality we may assume that σ^i has only function symbols. We use a theorem of *Gaifman* about models with a distance function (see [2]). We get that $\varphi^i(x, y)$ is logically equivalent to some local formula. This means for all $U \in \mathfrak{U}$ the truth value of $\varphi^i(x, y)$ in $\mathfrak{N}^i[U]$ depends only on the type of (x, y) on the set of formulas $\Phi^i := \bigcup_{j \in \{1,2,3\}} \Phi^i_j$ where:

$$\Phi^i_1 := \{f_i^{\varepsilon(1)} \circ f_2^{\varepsilon(2)} \circ \dots \circ f_t^{\varepsilon(t)}(x) = y : f_1, \dots, f_t \in \sigma^i, \varepsilon \in {}^t\{1, -1\}, t \leq s\},$$
$$\Phi^i_2 := \{f_i^{\varepsilon(1)} \circ f_2^{\varepsilon(2)} \circ \dots \circ f_t^{\varepsilon(t)}(x) = x : f_1, \dots, f_t \in \sigma^i, \varepsilon \in {}^t\{1, -1\}, t \leq s\},$$
$$\Phi^i_3 := \{f_i^{\varepsilon(1)} \circ f_2^{\varepsilon(2)} \circ \dots \circ f_t^{\varepsilon(t)}(y) = y : f_1, \dots, f_t \in \sigma^i, \varepsilon \in {}^t\{1, -1\}, t \leq s\},$$

and $s = s(i)$ is a natural number that depends only on φ^i. Define for each $j \in \{1, 2, 3\}$: $\Phi_j := \bigcup_{i<i^*} \Phi^i_j$ and $\Phi = \Phi_1 \cup \Phi_2 \cup \Phi_3$. Also define $\sigma^* := \bigcup_{i<i^*} \sigma^i$ and $\mathfrak{N}^*$ is defined by $(\forall i < i^*)\mathfrak{N}^*|\sigma^i := \mathfrak{N}^i$. Using Gaifman's theorem for all $U \in \mathfrak{U}$ and $a, b, a', b' \in U$ we have $(\bigotimes)$:

$$tp_{\Phi}((a, b), \emptyset, N^*) = tp_{\Phi}((a', b'), \emptyset, N^*) \Longrightarrow$$
$$tp_{\Delta^*}((a, b), A^*, M^*) = tp_{\Delta^*}((a', b'), A^*, M^*).$$

Note that $|\Phi|$ is uniformly bounded. Moreover the bound depends only on $|\Delta|$ and n. We consider each Φ type separately, this means: Let q be a type without parameters in Φ (that is simply $q \subseteq \Phi$). As we saw the number of such types is bounded by $2^{|\Phi|}$. By the fact above we are done if we interpret the relation $\mathfrak{S}_q$ defined by: $\mathfrak{S}[U] \cap \{(x, y, \bar{z}) \in {}^{n+1}U : tp_{\Phi}((x, y), \emptyset, \mathfrak{N}_1[U]) = q\}$. Clearly the relation $\{(x, y, \bar{z}) \in {}^{n+1}U : tp_{\Phi}((x, y), \emptyset, \mathfrak{N}_1[U]) = q\}$ is definable from $\mathfrak{N}^*$ by the formula $\varphi_q(x, y, \bar{z}) := \bigwedge_{\phi \in q} \phi \wedge \bigwedge_{\phi \in \Phi \setminus q} \neg\phi$. Now one of the following holds:

1. There exist $\theta(x, y) \in \Phi_1$ such that $\theta \in q$. Then for all $U \in \mathfrak{U}$ and $a, b \in U$ we have:

$$[tp_{\Phi}((a, b), \emptyset, \mathfrak{N}^*[U]) = q] \Longrightarrow \mathfrak{N}^* \models \theta(a, b).$$

2. For all $\theta(x, y) \in \Phi_1$, $\theta \notin q$. Then for all $U \in \mathfrak{U}$ we have:

$$\{(x, y) \in {}^2U : tp_{\Phi}((x, y), \emptyset, \mathfrak{N}^*[U]) = q\} =$$
$$\{(x, y) \in A_q \times B_q : \mathfrak{N}^*[U] \models \bigwedge_{\theta(x,y) \in \Phi_1} \neg\theta(x, y)\},$$

where we define:

$$A_q := \{x \in U : tp_{\Phi_2}(x, \emptyset, \mathfrak{N}^*[U]) = q \cap \Phi_2\},$$
$$B_q := \{y \in U : tp_{\Phi_3}(y, \emptyset, \mathfrak{N}^*[U]) = q \cap \Phi_3\}.$$

Assume condition (1) is satisfied. Note that for all $U \in \mathfrak{U}$, $\theta(x, y)$ defines in $\mathfrak{N}^*[U]$ a (graph of a) 1–1 function, denote this function by f^U. The relation defined by $\{(x, \bar{z}) \in {}^nU : xS^n_{\Delta,\mathfrak{M}[U]}f^U(x)\bar{z}\}$ is a n-place relation definable form $\mathfrak{N}$ by a simple expansion (using the formula $(\forall t)\theta(x, t) \to \chi(x, t, \bar{z})$). Hence there exist a formula $\varphi_1(x, \bar{z})$, a vocabulary σ_1 and a model $\mathfrak{N}_1$ interpreting it. Now the formula $\theta(x, y) \wedge \varphi_1(x, \bar{z}) \wedge \varphi_q(x, y, \bar{z})$ and the model for $\sigma_1 \cup \sigma^*$ which is the union of $\mathfrak{N}^*$ and $\mathfrak{N}_1$ interprets $\mathfrak{S}_q$ as desired.

We now assume that condition (2) is satisfied. Let $U \in \mathfrak{U}$ and $\bar{c} \in {}^{n-1}U$. We ask a question:

$\diamondsuit^U_{q,\bar{c}}$ Does there exist for all $B \subseteq U$ with $|B| \leq m_5$ and $B \supseteq A^*$, elements $a, b \in U \setminus B$ such that $aS^n_{\Delta,\mathfrak{M}[U[}b, \bar{c}$ and $tp_\Phi((a, b), \emptyset, \mathfrak{M}^*[U]) = q$.

Assume that there exist $U \in \mathfrak{U}$ and $\bar{c} \in {}^{n-1}U$ such that the answer to $\diamondsuit^U_{q,\bar{c}}$ is YES. Choose by induction on $j \leq m_4$ a pair $(a_j, b_j) \in {}^2U$ such that:

- $a_jS^n_{\Delta,M}b_j\bar{c}$.
- $tp_\Phi((a_j, b_j), \emptyset, N^*) = q$.
- $a_j, b_j \notin A^* \cup \{a_k : k < j\} \cup \{b_k : k < j\} \cup \{c_0, \dots, c_{n-2}\}$.

This is possible by the definition of m_5 and $\diamondsuit^U_{i,\bar{c}}$. From the sequence $\langle a_0, \dots, a_{m_4}\rangle$ we omit all the elements satisfying $a_iS^{n-1}_{\Delta_1(\Delta),\mathfrak{M}'[U]}\bar{c}$ where Δ_1 and $\mathfrak{M}'$ are taken from the symmetry lemma (see 6.5). We omitted at most $l^*(\Delta_1)$ elements. Now note that for all j_1, j_2: $a_{j_1}S^n_{\Delta,M}b_{j_2}\bar{c} \Rightarrow b_{j_2}S^n_{\Delta_2(\Delta),\mathfrak{M}'[U]}a_{j_1}\bar{c}$. Hence for all a_j (after the change) we have $|\{b_j : a_iS^n_{\Delta,M}b_j\bar{c}\}| \leq l^*(\Delta_2)$. Hence we can decrease the size of the sequences by a factor of $l^*(\Delta_2)$ and get $a_iS^{n+1}_{\Delta,M}b_j\bar{c} \Leftrightarrow i = j$. Since the bound on $|\Phi|$ depends only on n, $|\Delta|$ we may assume w.l.o.g. (by increasing m_1 and using Ramsey theorem) that the Φ-type in N^* without parameters of (a_{j_1}, b_{j_2}) depends only on the order type of (j_1, j_2). Hence we have sequences $\langle a_0, \dots, a_{m_1}\rangle$ and $\langle b_0, \dots, b_{m_1}\rangle$ such that:

(*) For all $j_1, j_2 \leq m_1$: $a_{j_1}S^n_{\Delta,M}b_{j_2}\bar{c} \iff j_1 = j_2$.

(**) For all $j_1, j_2, j_3, j_4 \leq m_1$:

$$tp_{\{x<y,x=y\}}((j_1, j_2), \emptyset, (\mathbb{N}, <)) = tp_{\{x<y,x=y\}}((j_3, j_4), \emptyset, (\mathbb{N}, <)) \Longrightarrow$$
$$tp_\Phi((a_{j_1}, b_{j_2}), \emptyset, N^*) = tp_\Phi((a_{j_3}, b_{j_4}), \emptyset, N^*).$$

Now w.l.o.g. we may assume that $m_1 \geq |\Phi_1|$ (otherwise replace m_1 by $\max\{m_1, |\Phi_1|\}$ in the definition of m_4). Hence there exists $0 < j^* \leq m_1$ such that: $N^* \models \neg \bigwedge_{\theta(x,y)\in\Phi_1} \theta(a_0, b_{j^*})$ (remember each $\theta(x, y)$ is a function). In addition by our definition $tp_\Phi((a_0, b_0), \emptyset, N^*) = q$, hence by condition (2) $a_0 \in A_q$. In addition we have $b_{j^*} \in B_q$ as $tp_\Phi((a_{j^*}, b_{j^*}), \emptyset, N^*) = q$, and hence $tp_\Phi((a_0, b_{j^*}), \emptyset, N^*) = q$ (see condition (2)). In the same way we get that there exists $0 \leq j^{**} < m_1$ such that $tp_\Phi((a_{j^{**}}, b_{m_2}), \emptyset, N^*) = q$. So by (**) we have $i, j < m_1 \Rightarrow tp_\Phi((a_i, b_j).\emptyset, N^*) = q$ and by $(\bigotimes)$ we get:

Sh:801

(∗∗∗) For all $j_1, j_2, j_3, j_4 \le m_1$:

$$tp_{\Delta^*}((a_{j_1}, b_{j_2}), A^*, M^*) = tp_{\Delta^*}((a_{j_3}, b_{j_4}), A^*, M^*).$$

We now prove:

Claim. There exists $m^* < m_2 - m_1 = (m_1)^2$ such that if (a', b') and (a'', b'') are pairs from $A_{m^*+m_1}$ that satisfy the same Δ^*-type over A_{m^*} in M^*, then $a' S^n_{\Delta,M} b' \bar{c} \equiv a'' S^n_{\Delta,M} b'' \bar{c}$.

PROOF. Assume the claim does not hold. Then for all $m < m_2 - m_1$ let (a'_m, b'_m) and (a''_m, b''_m) be pairs from A_{m+m_1} realizing the same Δ^*-type over A_{m^*} in M^*, and $\neg(a' S^n_{\Delta,M} b' \bar{c} \equiv a'' S^n_{\Delta,M} b'' \bar{c})$. Choose $\overline{c^m} \in {}^n A_{m+1}$ realizing $tp_{\Delta^*}(\overline{c^m}, A_m, M^*)$ (this is possible, see the definition of A_{m+1}). Now look at the formula $\phi(\bar{x}, \bar{y}) \in \Delta^*$. If $l_1 < l_2 < m_2 - m_1$ then (a'_{l_2}, b'_{l_2}) and (a''_{l_2}, b''_{l_2}) realizes the same Δ^*-type over A_{l_2} in M^*. Since $\overline{c^{l_1}} \subseteq A_{l_2}$ (as $l_1 < l_2$) and since $\chi(x, y, \bar{z}) \in \Delta^*$ interprets the relation $S^n_{\Delta,M}$ in M^*, we get that $a'_{l_2} S^{n+1}_{\Delta,M} b'_{l_2} \overline{c^{l_1}} \equiv a''_{l_2} S^{n+1}_{\Delta,M} b''_{l_2} \overline{c^{l_1}}$ hence $M^* \models \phi((a'_{l_2}, b'_{l_2}, a''_{l_2}, b''_{l_2}), \overline{c^{l_1}})$. On the other hand if $m_1 + l_2 \le l_1 < m_2$ then by the choice of (a'_{l_2}, b'_{l_2}) and (a''_{l_2}, b''_{l_2}) as a counter example we have $\neg(a'_{l_2} S^n_{\Delta,M} b'_{l_2} \bar{c} \equiv a''_{l_2} S^n_{\Delta,M} b''_{l_2} \bar{c})$. But $a'_{l_2}, b'_{l_2}, a''_{l_2}, b''_{l_2} \in A_{l_2+m_1} \subseteq A_{l_1}$ and $\overline{c^{l_1}}$ realizes the same Δ^*-type over A_{l_1} as $\bar{c}$, so by the definition of ϕ and M^* we have $M^* \models \neg\phi((a'_{l_2}, b'_{l_2}, a''_{l_2}, b''_{l_2}), \overline{c^{l_1}})$. Hence if we define $\langle \overline{d_l} = (a'_{l\cdot m_1}, b'_{l\cdot m_1}, a''_{l\cdot m_1}, b''_{l\cdot m_1}) : l < m_1 \rangle$ and $\langle \overline{e_l} = \overline{c^{l\cdot m_1}} : l < m_1 \rangle$, then $\phi(\bar{x}, \bar{y})$ defines an order relation in the sense of 6.8 on them, in contradiction to $m_1 > k^*_3(\phi)$. This completes the proof of the claim. ⊣

Now let m^* be the one from the claim above. For all $l < m_1$ we choose from A_{m^*+l+1} the sequence $\overline{c^l} ^\frown \langle a^l_j, b^l_j : j < m_1 \rangle$ that realizes the same Δ^*-type over A_{m^*+l} in M^* as $\bar{c} ^\frown \langle a_j, b_j : j < m_1 \rangle$.

Claim. The sequences $\langle \bar{c^l} : l < m_1 \rangle$ and $\langle a^l_i : i < m_1 \rangle$, $\langle b^l_j : j < m_1 \rangle$ satisfy the demands of lemma 6.9 for the formula $\chi(x, y, \bar{z})$.

PROOF. (α) follows directly from (∗) and the equality of types. For (β) let $l_1 < l < m_1$ and $j_1, \dots, j_4 < m_1$. Then $(a^{l_1}_{j_1}, b^{l_1}_{j_2})$ and $(a^{l_1}_{j_3}, b^{l_1}_{j_4})$ are pairs from $A_{m^*+l_1}$ and from (∗∗∗) and the equality of types we get that these pairs realize the same Δ^*-types over $A_{m^*+l_1}$ and in particular over A_{m^*}. So by the claim we have ${}^{l_1}_{j_1} S^n_{\Delta,M} b^{l_1}_{j_2} \bar{c} \equiv a^{l_1}_{j_3} S^n_{\Delta,M} b^{l_1}_{j_4} \bar{c}$. In conclusion as $\bar{c}$ and $\overline{c^{l_2}}$ realizes the same Δ^*-type over $A_{m^*+l_1}$ and by the interpretation of the formula ϕ in M^* we get $a^{l_1}_{j_1} S^n_{\Delta,M} b^{l_1}_{j_2} \overline{c^l} \equiv a^{l_1}_{j_3} S^n_{\Delta,M} b^{l_1}_{j_4} \overline{c^l}$ as in (β). ⊣

This leads to a contradiction as $m_1 > k^*_4(\chi)$.

We are left with the case where for all $U \in \mathfrak{U}$ and $\bar{c} \in {}^{n-1}U$ the answer to $\diamondsuit^U_{q,\bar{c}}$ is NO. In this case: For all $U \in \mathfrak{U}$ and $\bar{c} \in {}^{n-1}U$ there exist $B = B_{\bar{c}} \subseteq U$ with size $\le m_5$ such that there are no $u, v \notin B$ satisfying $u S^n_{\Delta,\mathfrak{M}[U]} v \bar{z}$ and $tp_\Phi((u, v), \emptyset, \mathfrak{N}^*[U]) = q$. Define $\mathfrak{T}$ a n-place relation on $\mathfrak{U}$ as follows. For all $U \in \mathfrak{U}$, $a \in U$ and $\bar{c} \in {}^{n-1}U$: $\mathfrak{T}[U](a, \bar{c})$ iff there exists $B \subseteq U$ of size $\le m_5$ such that:

- There are no $u, v \notin B$ with: $u S^n_{\Delta,\mathfrak{M}[U]} v \bar{z}$ and $tp_\Phi((u, v), \emptyset, \mathfrak{N}^*[U]) = q$.
- B is minimal under the previous demand.
- $a \in B$.

Sh:801

It is clear that $\mathfrak{T}$ is definable from $\mathfrak{N}$ in a simple expansion, and hence satisfies the induction hypothesis. Let $\sigma^{**}, \varphi^{**}(x, \bar{z})$ and $\mathfrak{N}^{**}$ be the formula, vocabulary and model that interprets $\mathfrak{T}$. We define $m_6 := m_5 \cdot Delta(m_5, 3)$ and show that for all $U \in \mathfrak{U}$ and $\bar{c} \in {}^{n-1}U$:

$$|\{x \in U : \mathfrak{N}^{**}[U] \models \varphi^{**}(x, \bar{c})\}| \leq m_6.$$

Assume toward contradiction that U and $\bar{c}$ does not satisfy that claim. Then by the definition of φ^{**} we have a sequence $\langle B_l \subseteq U : l < m_7 \rangle$ such that:

1. For all $l < m_7$, $m_5 \geq |B_l|$.
2. For all $l < m_7$, there are no $u, v \notin B_l$ s.t. $uS^n_{\Delta, \mathfrak{M}[U]} v\bar{c}$ and
$$tp_\Phi((u, v), \emptyset, \mathfrak{N}^*[U]) = q.$$
3. For all $l < m_7$, B_l is minimal under (1) and (2).
4. For all $l < m_7$, $B_l \nsubseteq \bigcup_{m<l} B_m$.
5. $\bigcup_{l<m_7} B_l = \{x : \mathfrak{N}^{**}[U] \models \varphi^{**}(x, \bar{c})\}$.

To get this sequence we start with a sequence of all the sets satisfying claims (1)–(3) in some order, and omits those that do not satisfy claim (4). claim (5) follows straight from the definition of φ^{**}. Now, by (1) and the assumption we get:

$$m_6 < \left|\{x : m^{**} \models \varphi_i^*(x, \bar{c})\}\right| = \left|\bigcup_{l<m_7} B_l\right| \leq m_7 \cdot max\{|B_l| : l < m_7\} \leq m_7 \cdot m_5$$

so we have $m_7 \geq m_6/m_5$. By the definition of m_6 and Ramsey theorem we have $B^* \subseteq U$ and $l_1 < l_2 < l_3 \leq m_7$ such that $i \neq j \Rightarrow B_{l_i} \cap B_{l_j} = B^*$. We prove that B^* satisfies (1) and (2). Since $B^* \subsetneqq B_{l_3}$ this will be a contradiction to the minimality of B_{l_3}. Obviously B^* satisfies (1), to show (2) take some $a, b \notin B^*$ then by $i \neq j \Rightarrow B_{l_i} \cap B_{l_j} = B^*$ we have $j \in (1, 2, 3\}$ such that $a, b \notin B_{l_j}$ and since B_{l_j} satisfies (2) we get $\neg aS^{n,i}_{\Delta,M} b\bar{c}$ or $tp_\Phi((u, v), \emptyset, \mathfrak{N}^*[U]) \neq q$ as needed.

We use Gaifman's theorem again on the formula $\varphi^{**}(x, \bar{z})$ (w.l.o.g. σ^{**} have only function symbols). We get that for all $U \in \mathfrak{U}$ the truth value of $\varphi^{**}(x, \bar{z})$ in $\mathfrak{M}^{**}[U]$ depends only in the type without parameters of $(x, \bar{z})$ in $\mathfrak{M}^{**}[U]$ for the set of formulas $\Psi := \bigcup_{j \in \{1,2,3\}} \Psi_j$ where:

$$\begin{aligned}
\Psi_1 := \{&f_i^{\varepsilon(1)} \circ f_2^{\varepsilon(2)} \circ \cdots \circ f_t^{\varepsilon(t)}(x) = z_i : f_1, \ldots, f_t \in \sigma^{**},\\
&\varepsilon \in {}^t\{1, -1\}, t \leq s, i < n-1\},\\
\Psi_2 := \{&f_i^{\varepsilon(1)} \circ f_2^{\varepsilon(2)} \circ \cdots \circ f_t^{\varepsilon(t)}(x) = x : f_1, \ldots, f_t \in \sigma^{**},\\
&\varepsilon \in {}^t\{1, -1\}, t \leq s\},\\
\Psi_3 := \{&f_i^{\varepsilon(1)} \circ f_2^{\varepsilon(2)} \circ \cdots \circ f_t^{\varepsilon(t)}(z_i) = z_j : f_1, \ldots, f_t \in \sigma^{**},\\
&\varepsilon \in {}^t\{1, -1\}, t \leq s, i, j < n-1\},
\end{aligned}$$

and s is a natural number that depends only on φ^{**}. Note that $|\Psi|$ is uniformly bounded. Again we separate into cases according to the Ψ-types. Let q_1, q_2 be Ψ-types without parameters (formally $q_1, q_2 \subseteq \Psi$). The number of such types is bounded by $2^{|\Psi|}$ so as we saw it is enough to interpret the relation $\mathfrak{S}_{q,q_1,q_2}$ defined by:

$$\mathfrak{S}_q[U] \cap \{(x, y, \bar{z}) \in {}^{n+1}U : tp_\Psi((x, \bar{z}), \emptyset, \mathfrak{N}^{**}[U]) = q_1 \wedge \\ tp_\Psi((y, \bar{z}), \emptyset, \mathfrak{N}^{**}[U]) = q_2\}.$$

The relation $\{(x, y, \bar{z}) \in {}^{n+1}U : tp_\Psi((x, \bar{z}), \emptyset, N^{**}) = q_1 \wedge tp_\Psi((x, y), \emptyset, N^{**}) = q_2\}$ is definable in $\mathfrak{N}^{**}$ by a formula denoted $\varphi_{q_1,q_2}(x, y, \bar{z})$. Now, for $l \in \{1, 2\}$ one of the following hold:

1. There exists $\theta(x, z_i) \in \Psi_1$ such that $\theta \in q_l$. Then for all $U \in \mathfrak{U}$, $a \in U$ and $\bar{c} \in {}^{n-1}U$ we have:

$$[tp_\Psi((a, \bar{c}), \emptyset, \mathfrak{N}^{**}[U]) = q_l] \Longrightarrow \mathfrak{N}^{**}[U] \models \theta(a, c_i).$$

2. For all $\theta(x, z_i) \in \Phi_1$, $\theta \notin q_l$, and then for all $U \in \mathfrak{U}$ we have:

$$\{(x, \bar{z}) \in {}^nU : tp_\Psi((x, \bar{z}), \emptyset, \mathfrak{N}^{**}[U]) = q_l\} = \\ \{(x, \bar{z}) \in A'_{q_l} \times B'_{q_l} : \mathfrak{N}^{**}[U] \models \bigwedge_{\theta(x,z_i)\in\Psi_1} \neg\theta(x, z_i)\},$$

where we define:

$$A'_{q_l} := \{x \in U : tp_{\Psi_2}(x, \emptyset, \mathfrak{N}^{**}[U]) = q_l \cap \Psi_2\},$$
$$B'_{q_l} := \{\bar{z} \in {}^{n-1}U : tp_{\Psi_3}(\bar{z}, \emptyset, \mathfrak{N}^{**}[U]) = q_l \cap \Phi_3\}.$$

Assume that there exists $l \in \{1, 2\}$ such that (1) holds. Then, as we have seen, $\theta(x, z_i)$ defines in each $\mathfrak{N}^{**}[U]$ (a graph of) a one to one function, denote that function by f^u. the relation defined by:

$$\{(x, y, z_0, \dots, \widehat{z_i}, \dots, z_{n-2}) \in {}^nU : xS^n_{\Delta,\mathfrak{M}[U]}y, z_0, \dots, z_{i-1}, f^U(x), z_{i+1}, \dots, z_{n-2}\},$$

satisfies the induction hypothesis. Let $\varphi^1(x, y, z_0, \dots, \hat{z}_j, \dots, j_{n-2})$ be the formula in vocabulary σ^1 that interprets this relation in the simple model $\mathfrak{N}^1$. In the same way we interpret the relation defined by:

$$\{(x, y, z_0, \dots, \widehat{z_i}, \dots, z_{n-2}) \in {}^nU : xS^n_{\Delta,\mathfrak{M}[U]}y, z_0, \dots, z_{i-1}, f^U(y), z_{i+1}, \dots, z_{n-2}\}$$

using the formula $\varphi^2(x, y, z_0, \dots, \hat{z}_j, \dots, z_{n-2})$ in the vocabulary σ^2 and the simple model $\mathfrak{N}^2$. Now the formula

$$\varphi_q(x, y, \bar{z}) \wedge \varphi_{q_1,q_2}(x, y, \bar{z}) \wedge \varphi^l(x, y, z_0, \dots, \hat{z}_j, \dots, j_{n-2})$$

in the vocabulary $\sigma^* \cup \sigma^{**} \cup \sigma^l$ and the union of models: $\mathfrak{N}^{**}, \mathfrak{N}^*, \mathfrak{N}^l$, interprets the relation $\mathfrak{S}_{q,q_1,q_2}$ as needed.

Assume then that for each $l \in \{1, 2\}$ (2) holds. Seeking a contradiction we assume that there exists $U \in \mathfrak{U}$ such that:

- $\mathfrak{S}_{q,q_1,q_2}[U] \neq \emptyset$.
- For all $l \in \{1, 2\}$, $|A'_{q_l}| > m_6 + |\Psi_1|$.

So we have $a, b \in U$ and $\bar{c} \in {}^{n-1}U$ such that $\mathfrak{S}_{q,q_1,q_2}[U](a, b, \bar{c})$. Recall that we are assuming $\neg\diamondsuit^U_{q,\bar{c}}$. Hence we have $\mathfrak{N}^{**} \models \varphi^{**}(a, \bar{c}) \vee \varphi^{**}(b, \bar{c})$, because we can choose some minimal $B_{\bar{c}}$ (there is one because of $\neg\diamondsuit^U_{q,\bar{c}}$) and then $a, b \notin B_{\bar{c}}$ is contradicting the definition of $B_{\bar{c}}$. w.l.o.g we assume that $\mathfrak{N}^{**} \models \varphi^{**}(a, \bar{c})$. Now q_1 satisfies (2) so $\bar{c} \in B'_{q_1}$ as $tp_{\Psi}((a, \bar{c}), \emptyset, N^{**}) = q_1$. Note that

$$\left|\{a' \in A'_{q_1} : \mathfrak{N}^{**}[U] \not\models \bigwedge_{\theta(x,z_i) \in \Psi_1} \neg\theta(a, c_i)\}\right| \leq |\Psi_1|$$

(again each $\theta(x, z_i)$ is a function) and hence A'_{q_1} has more than m_6 (distinct) elements $\{a_0, \dots a_{m_6}\}$ satisfying $tp_{\Psi}((a_i, \bar{c}), \emptyset, \mathfrak{N}^{**}[U]) = q_1$. But we also have $tp_{\Psi}((a, \bar{c}), \emptyset, \mathfrak{N}^{**}[U]) = q_1$ and $\mathfrak{N}^{**} \models \varphi^{**}(a, \bar{c})$, so for all $0 \leq i \leq m_6$, $\mathfrak{N}^{**} \models \varphi^{**}(a_i, \bar{c})$ which is a contradiction.

Finally we assume that there is no $U \in \mathfrak{U}$ satisfying the two demands above. We then divide $\mathfrak{U}$ into three parts $\mathfrak{U}_i : i \in \{1, 2, 3\}$ such that:

- $\mathfrak{S}_{q,q_1,q_2}[U] = \emptyset \iff U \in \mathfrak{U}_1$.
- $|A'^{\,U}_{q_1}| < m_6 + |\Psi_1| \iff U \in \mathfrak{U}_1$.
- $|A'^{\,U}_{q_2}| < m_6 + |\Psi_1| \iff U \in \mathfrak{U}_2$.

By our assumption $\bigcup_{i \in \{1,2,3\}} \mathfrak{U}_i = \mathfrak{U}$. Now for each $i \in \{1, 2, 3\}$ it is easy to interpret $\mathfrak{S}_{q,q_1,q_2}$ restricted to $\mathfrak{U}_i$ (using the formula $(\exists x)x \neq x$, or by adding a bounded number of constants to the vocabulary interpreted as the elements of $A'^{\,U}_{q_1}$ or $A'^{\,U}_{q_2}$ and using the induction hypothesis). Assume then that for each $i \in \{1, 2, 3\}$ the formula $\varphi_i^{***}(x, y, \bar{z})$ in the vocabulary σ_i^{***} interprets in the model $\mathfrak{N}_i^{***}$, the relation $\mathfrak{S}_{q,q_1,q_2}$ restricted to $\mathfrak{U}_i$. We now define $\sigma^{***} = \bigcup_{i \in \{1,2,3\}} \sigma_i^{***} \cup \{s_1, s_2, s_3\}$ (w.l.o.g the union is disjoint), and a model $\mathfrak{N}^{***}$ for σ^{***}, such that for each $i \in \{1, 2, 3\}$: $(\mathfrak{N}^{***}|\mathfrak{U}_i)|\sigma_i^{***} := \mathfrak{N}_i^{***}$, and for all $U \in \mathfrak{U}$, $s_i^{\mathfrak{N}^{***}[U]} \neq \emptyset$ iff $U \in \mathfrak{U}_i$. (if $i \neq j$ the definition of $(\mathfrak{N}^{***}|\mathfrak{U}_i)|\sigma_j^{***}$ is insignificant). Now the formula:

$$\varphi^{***}(x, y, \bar{z}) := \bigvee_{i \in \{1,2,3\}} (\exists u s_i(u)) \longrightarrow \varphi_i^{***}(x, y, \bar{z})$$

interprets $\mathfrak{S}_{q,q_1,q_2}$ in the model $\mathfrak{N}^{***}$ as required. This completes the proof of lemma 6.11. ⊣

In the second stage of proving theorem 6.10 we interpret $\mathfrak{R}$ itself. We prove the following:

Lemma 6.12. *There exist a simple vocabulary σ, and a finite set Φ of formulas in σ, and a simple model $\mathfrak{N}$ for σ on $\mathfrak{U}$. Such that for all $U \in \mathfrak{U}$ and $\bar{x}, \overline{x'} \in {}^{n(\mathfrak{R})}U$ if $tp_{\Phi}(\bar{x}, \emptyset, \mathfrak{N}[U]) = tp_{\Phi}(\bar{x}', \emptyset, \mathfrak{N}[U])$ then $(U, \mathfrak{R}[U]) \models r(\bar{x}) \equiv r(\bar{x}')$.*

Proof. Define: $\Delta := \{r(x_0, \dots, x_{n(\mathfrak{R})-1})\}$, and we denote the first variable by x and the last n variables by $\bar{y}$ (so $\Delta := \{r(x, \bar{y})\}$).

We define σ and Φ simultaneously and also we define $\mathfrak{N}[U]$ for some $U \in \mathfrak{U}$. Let $a, a' \in U$ and $\overline{b}, \overline{b'} \in {}^nU$, and assume that $tp_{\Phi}(a\overline{b}, \emptyset, \mathfrak{N}[U]) = tp_{\Phi}(a'\overline{b'}, \emptyset, \mathfrak{N}[U])$, where σ, Φ and $\mathfrak{N}$ will be defined.

Let $\sigma_0, \varphi_0(x, \bar{y})$ and $\mathfrak{M}_0$ be those who interpret $xS^n_{\Delta,\mathfrak{M}[U]}\bar{y}$, i.e., those we get from applying the previous lemma to Δ (where $\tau = \emptyset$ and $\mathfrak{M} = \mathfrak{R}$). For brevity we write $M = \mathfrak{M}[U]$, $\mathfrak{N}[U] = N$, $\mathfrak{N}_0[U] = N_0$ and $R = \mathfrak{R}[U]$. We add σ_0 to σ, φ_0 to Φ and demand $N|\sigma_0 = N_0$. Now we have:

$$aS^n_{\Delta,M}\overline{b} \equiv a'S^n_{\Delta,M}\overline{b'}.$$

We write $E = E^{\Delta,M}_{A^{\Delta,M}_U}$ and define $A^* = A^{\Delta,M}_U \cup \{x : |x/E| \leq 2 \cdot k^*_1(\Delta)\}$. for all $\alpha \in A^*$ we add to σ a constant c_α and put $c^N_\alpha := \alpha$. In addition for each equivalence class x/E we add to σ a unary relation symbol $s_{x/E}$ and put $s^N_{x/E} := x/E$. Note that both $|A^*|$ and the number of equivalence classes is uniformly bounded. We add to Φ formulas of the form $x = c$ and $y_i = c$ for each constant $c \in \sigma$, and formulas of the form $s(x)$ and $s(y_i)$ for each relation symbol $s \in \sigma$. Now for each constant $c \in \sigma$ the relation class on $\mathfrak{U}$, $\mathfrak{R}_c$ defined by $\mathfrak{R}_c[U'] := \mathfrak{R}[U'](c^{\mathfrak{N}[U']}, \bar{y})$ satisfies the induction hypothesis. That is it is a class of n-place relation not satisfying condition (1) in theorem 3.6. Hence we can add to σ and Φ the dictionaries and formulas we get from applying the induction hypothesis to each $\mathfrak{R}_c$, and expand $\mathfrak{N}$ accordingly. Assume $a \in A^*$, then (due to the formula $x = c_a$) we have $a = a'$. Because of the formulas we added to Φ for the relation $\mathfrak{R}_{c_a}$ we have:

$$\mathfrak{R}_{c_a}[U](\bar{b}) \equiv \mathfrak{R}_{c_a}[U](\overline{b'}).$$

This implies $R(c^N_a, \bar{b}) \equiv R(c^N_{a'}, \overline{b'})$. But since $c^N_a = a = a' = c^N_{a'}$ we get $R(a, \bar{b}) \equiv R(a', \overline{b'})$, as claimed. This proves the cases $a \in A^*$ and $a' \in A^*$.

Now for each x/E (where $x \notin A^*$) and $\bar{y} \in {}^nU$ we define $t^{x/E}_{\bar{y}} \in \{\mathbb{T}, \mathbb{F}\}$ to be the truth value the formula $r(-, \bar{y})$ gets for the majority of elements in x/E. This means: $t^{x/E}_{\bar{y}} = \mathbb{T}$ iff $|\{x' : xEx' \wedge R(x', \bar{y})\}| > k^*_1(\Delta)$. Note that this is true as $x \notin A^*$ and so $|x/E| > 2 \cdot k^*_1(\Delta)$. We get:

$$\neg a S^n_{\Delta,M} \bar{b} \Rightarrow [R(a, \bar{b}) \equiv (t^{a/E}_{\bar{b}} = \mathbb{T})],$$

and since Δ has only one formula we get:

$$a S^n_{\Delta,M} \bar{b} \Rightarrow [R(a, \bar{b}) \equiv (t^{a/E}_{\bar{b}} = \mathbb{F})].$$

For each x/E we have a class of relations $\mathfrak{R}_{x/E}$ on $\mathfrak{U}$ defined by $\mathfrak{R}_{x/E}[U'] := \{\bar{y} \in {}^nU' : t^{x/E}_{\bar{y}} = \mathbb{T}\}$, which satisfies the induction hypothesis. Hence we can add to σ and Φ the dictionaries and formulas we get form applying the induction hypothesis to each $\mathfrak{R}_{x/E}$ and expand $\mathfrak{N}$ accordingly. We get for all $x \notin A^*$:

$$\mathfrak{R}_{x/E}[U](\bar{b}) \equiv \mathfrak{R}_{x/E}[U](\overline{b'}).$$

Since $a/E = a'/E$ (due to the formula $s_{a/E}(x)$), we have $t^{a/E}_{\bar{b}} = t^{a/E}_{\overline{b'}}$. Assume $\neg a S^n_{\Delta',M'} \bar{b}$ (as we saw $a S^n_{\Delta',M'} \bar{b} \equiv a' S^n_{\Delta',M'} \overline{b'}$) the we have:

$$R(a, \bar{b}) \Leftrightarrow (t^{a/E}_{\bar{b}} = \mathbb{T}) \Leftrightarrow (t^{a/E}_{\overline{b'}} = \mathbb{T}) \Leftrightarrow (t^{a'/E}_{\overline{b'}} = \mathbb{T}) \Leftrightarrow R(a', \overline{b'})$$

as claimed. If $a S^n_{\Delta',M'} \bar{b}$ then again we get:

$$R(a, \bar{b}) \Leftrightarrow (t^{a/E}_{\bar{b}} = \mathbb{F}) \Leftrightarrow (t^{a/E}_{\overline{b'}} = \mathbb{F}) \Leftrightarrow (t^{a'/E}_{\overline{b'}} = \mathbb{F}) \Leftrightarrow R(a', \overline{b'}).$$

This completes the proof of lemma 6.12. ⊣

From the lemma it is easy to prove that $\mathfrak{R}$ is interpretable by a formula in a simple model. The proof is identical to the binary case (see the proof of 5.7). This completes the proof of theorem 6.10. ⊣

REFERENCES

[1] JOHN T. BALDWIN, *Definable second order quantifiers*, ***Model theoretic logics*** (J. Barwise and S. Feferman, editors), Perspectives in Mathematical Logic, Springer-Verlag, 1985, pp. 446–478.

[2] HAIM GAIFMAN, *On local and nonlocal properties*, ***Logic Colloquium '81*** (J. Stern, editor), North Holland, 1982, pp. 105–135.

[3] I. A. LAVROV, *The effective non-separability of the set of identically true formulae and the set of finitly refutable formulae for certain elementary theories*, ***Algebra i Logika***, vol. 2 (1963), no. 1, pp. 5–18, (Russian).

[4] SAHARON SHELAH, *There are just four second-order quantifiers*, ***Israel Journal of Mathematics***, vol. 15 (1973), pp. 282–300.

[5] ———, *Classifying of generalized quantifiers*, ***Around classification theory of models***, Lecture Notes in Mathematics, no. 1182, Springer-Verlag, 1986, pp. 1–46.

[6] ———, *On quantification with a finite universe*, this JOURNAL, vol. 65 (2000), pp. 1055–1075.

THE HEBREW UNIVERSITY OF JERUSALEM
EINSTEIN INSTITUTE OF MATHEMATICS
EDMOND SAFRA CAMPUS
GIVAT RAM
JERUSALEM 91904, ISRAEL
and
DEPARTMENT OF MATHEMATICS
HILL CENTER-BUSCH CAMPUS RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY
110 FRELINGHUYSEN ROAD
PISCATAWAY, NJ 08854-8019, USA
E-mail: shelah@math.huji.ac.il

THE HEBREW UNIVERSITY OF JERUSALEM
EINSTEIN INSTITUTE OF MATHEMATICS
EDMOND SAFRA CAMPUS
GIVAT RAM
JERUSALEM 91904, ISRAEL
E-mail: mord@math.huji.ac.il