The Hierarchy Theorem for Second Order Generalized Quantifiers

Juha Kontinen

Abstract

We study definability of second order generalized quantifiers on finite structures. Our main result says that for every second order type t there exists a second order generalized quantifier of type t which is not definable in the extension of second order logic by all second order generalized quantifiers of types lower than t.

1 Introduction

The notion of *second order generalized quantifier* was introduced in [1]. Syntactically, a second order generalized quantifier $\mathcal{Q}$ of type $((1))$ is a quantifier which applies to a single formula and binds one unary second order variable. The interpretation of such a quantifier is determined by a class $\mathcal{Q}$ of second order structures (M, G), where M is a set and $G \subseteq \mathcal{P}(M)$, such that $\mathcal{Q}$ is closed under isomorphisms. In other words, the semantics of $\mathcal{Q}$ is given by

$$\mathbb{M} \models \mathcal{Q}X\,\varphi(X) \Leftrightarrow (M, \varphi^{\mathbb{M}}) \in \mathcal{Q},$$

where $\varphi^{\mathbb{M}} = \{A \subseteq M \mid \mathbb{M} \models \varphi(A)\}$. In this paper we study definability of second order generalized quantifiers. The notion of definability we study is analogous to the notion of definability of Lindström quantifiers [5]. Let us take an example.

Example 1.1. Let $\mathcal{Q}$ be the quantifier $\{(M, P) \mid P \subseteq \mathcal{P}(M)$ and $|P| \geq 2\}$ and $\psi(X)$ a formula. Then

$$\models \mathcal{Q}X\,\psi(X) \leftrightarrow \exists X \exists Y (X \neq Y \wedge \psi(X) \wedge \psi(Y)).$$

This research was financially supported by the Vilho, Yrjö and Kalle Väisälä Foundation and the Graduate School of Mathematical Analysis and Logic.

2000 *Mathematics Subject Classification.* Primary 03C80; Secondary 03C13, 03C85.

Example 1.1 shows that extending second order logic, SO, by the quantifier $\mathcal{Q}$ does not increase its expressive power. In addition, it shows that all occurrences of the quantifier $\mathcal{Q}$ can be eliminated from the formulas of $SO(\mathcal{Q})$ in a uniform way.

In the case of Lindström quantifiers, i.e., first order generalized quantifiers, definability of a quantifier Q in a logic $\mathcal{L}$ just means that the class Q is axiomatizable in $\mathcal{L}$. Since second order generalized quantifiers are classes of second order structures, there is no connection between quantifiers and classes of structures determined by sentences. Therefore, we introduce logics which are interpreted in second order structures consisting of a first order part and some second order predicates. Syntactically, these second order predicates act as first order quantifiers (so-called quantifier variables) interpretations of which are given locally and arbitrarily in every structure. This approach enables us to transform questions like "Can a logic $\mathcal{L}$ define $\mathcal{Q}$" to "Is $\mathcal{Q}$ the class of models of some sentence $\psi \in \mathcal{L}'$", where $\mathcal{L}'$ is an extension of $\mathcal{L}$ by some first order quantifier variables. It turns out that a defining sentence $\psi \in \mathcal{L}'$ can be then converted into a uniform definition of $\mathcal{Q}$ in the logic $\mathcal{L}$. In particular, definability of $\mathcal{Q}$ in $\mathcal{L}$ implies that $\mathcal{L}(\mathcal{Q}) \equiv \mathcal{L}$.

Our main result is the analogue of the Hierarchy Theorem in [4]:

Hierarchy Theorem. *Let s be a first order type and $\mathcal{B}$ the collection of all Lindström quantifiers of types lower than s. Then there is a Lindström quantifier Q of type s which is not definable in the logic $FO(\mathcal{B})$.*

We shall prove the following.

Theorem 2.21. *Let t be a second order type and $\mathbb{Q}$ the collection of all second order generalized quantifiers of types lower than t. Then there is a quantifier $\mathcal{Q}$ of type t which is not definable in the logic $SO(\mathbb{Q})$.*

Before proving Theorem 2.21, we prove a version of Fagin's lemma for second order structures which is the main result used in the proof of the Hierarchy Theorem. Analogously to the first order case, our version of Fagin's lemma induces a natural well-ordering of second order types. This well-ordering is then used to state and prove Theorem 2.21.

1.1 Notations

Vocabularies τ are finite sets consisting of relation symbols and constant symbols. All structures are assumed to be finite. The universe of a structure $\mathbb{M}$ is denoted by M. For $\tau = \emptyset$, we denote τ-structures $\mathbb{M}$ by their universes M. The cardinality of a set X is denoted by

$|X|$. With the cardinality of a structure $\mathbb{M}$ we mean the cardinality of the universe M. The class of all τ-structures is denoted by $Str(\tau)$ and $Str(\tau, n)$ is the restriction of $Str(\tau)$ to structures with cardinality n. If φ is a τ-sentence then the class of τ-models of φ is denoted by $Mod(\varphi)$. For a class K of τ-structures, the number of isomorphism types of structures in K with universe $\{1, \ldots, n\}$ is denoted by $I_n(K)$. Finally, for functions $f, g\colon \omega \to \omega$ on natural numbers, we write $f(n) \sim g(n)$ if $\lim_{n\to\infty} f(n)/g(n) = 1$.

1.2 Fagin's lemma

The Hierarchy Theorem in [4] is based on a result called Fagin's lemma. Fagin's lemma is concerned with the number of isomorphism types of finite relational structures.

Let $s = (l_1, \ldots, l_r)$ be a tuple of positive integers (a first order type). A structure of type s is a structure of the form $(M, P_1, \ldots, P_r)$, where $P_i \subseteq M^{l_i}$ for $1 \leq i \leq r$. In other words, a s-structure is a structure of vocabulary $\{P_1, \ldots, P_r\}$, where P_i is l_i-ary for $1 \leq i \leq r$. Denote the number of s-structures with universe $\{1, \ldots, n\}$ by $a_n(s)$ and the number of isomorphism types of s-structures with universe $\{1, \ldots, n\}$ by $b_n(s)$. It is easy to see that $a_n(s) = 2^{n^{l_1}+\cdots+n^{l_r}}$. The real task is to determine the value $b_n(s)$. Fagin's lemma [3] answers this question.

Fagin's lemma. *Let $s = (l_1, \ldots, l_r)$ be a first order type such that $max\{l_i : 1 \leq i \leq r\} > 1$. Then $b_n(s) \sim \frac{1}{n!} a_n(s)$.*

Proof. See [3]. $\square$

We shall prove a version of Fagin's lemma for second order structures in Section 2.1.

1.3 Second order generalized quantifiers

The notion of second order generalized quantifier was introduced in [1]. Let $t = (s_1, \ldots, s_w)$, where $s_i = (l^i_1, \ldots, l^i_{r_i})$ is a tuple of positive integers for $1 \leq i \leq w$. A *second order structure* of type t is a structure of the form $(M, P_1, \ldots, P_w)$, where

$$P_i \subseteq \mathcal{P}(M^{l^i_1}) \times \cdots \times \mathcal{P}(M^{l^i_{r_i}}).$$

In the following the class of all t-structures is denoted by $Str(t)$. Also, the notations $Str(t, n)$, $a_n(t)$, and $b_n(t)$ are used as in the case of first order structures.

Definition 1.2. A *second order generalized quantifier* $\mathcal{Q}$ of type t is a class of structures of type t such that $\mathcal{Q}$ is closed under isomorphisms.

Example 1.3. Let us look at some examples of second order generalized quantifiers. Let $S \subseteq \omega$ and $n \in \omega$.

$$\begin{array}{rcl} \exists^2_n & = & \{(M,P) \mid P \subseteq \mathcal{P}(M^n) \text{ and } P \neq \emptyset\} \\ Even & = & \{(M,P) \mid P \subseteq \mathcal{P}(M) \text{ and } |P| \text{ even}\} \\ Even' & = & \{(M,P) \mid P \subseteq \mathcal{P}(M) \text{ and } \forall X \in P(|X| \text{ even})\} \\ Three & = & \{(M,P) \mid P \subseteq \mathcal{P}(M) \text{ and } |P| = 3\} \\ \mathcal{Q}_S & = & \{(M,P) \mid P \subseteq \mathcal{P}(M) \text{ and } |P| \in S\} \end{array}$$

The first example is the correspondent of the familiar n-ary second order existential quantifier. The quantifier $Even$ says that a formula holds for an even number of subsets of the domain. On the other hand, the quantifier $Even'$ says that all the sets satisfying a formula have an even number of elements.

Definition 1.4. The extension $SO(\mathcal{Q})$ of SO by a quantifier $\mathcal{Q}$ is defined as follows:

(i) The formula formation rules of SO are extended by the rule: if for $1 \leq i \leq w$, $\varphi_i(\bar{X}_i)$ is a formula and $\bar{X}_i = (X_{1,i}, \ldots, X_{r_i,i})$ is a tuple of pairwise distinct predicate variables such that $ar(X_{j,i}) = l^i_j$ for $1 \leq j \leq r_i$, then

$$\mathcal{Q}\bar{X}_1, \ldots, \bar{X}_w\, (\varphi_1(\bar{X}_1), \ldots, \varphi_w(\bar{X}_w))$$

is a formula.

(ii) Satisfaction relation of SO is extended by the rule:

$$\mathbb{M} \models \mathcal{Q}\bar{X}_1, \ldots, \bar{X}_w\, (\varphi_1, \ldots, \varphi_w) \Leftrightarrow (M, \varphi_1^{\mathbb{M}}, \ldots, \varphi_w^{\mathbb{M}}) \in \mathcal{Q},$$

where $\varphi_i^{\mathbb{M}} = \{\bar{Y} \in \mathcal{P}(M^{l^i_1}) \times \cdots \times \mathcal{P}(M^{l^i_{r_i}}) \mid \mathbb{M} \models \varphi_i(\bar{Y})\}$.

2 Definability in extensions of second order logic

In the following an extension of SO by a collection $\mathbb{Q}$ of second order quantifiers is denoted by $SO(\mathbb{Q})$. Before formulating the notion of definability studied in this paper, we need to introduce some auxiliary logics.

Definition 2.1. Let $t = (s_1, \ldots, s_w)$ be a type, where $s_i = (l_1^i, \ldots, l_{r_i}^i)$ for $1 \le i \le w$, and let $\mathcal{G}_1, \ldots, \mathcal{G}_w$ be first order quantifier symbols of types $s_1, \ldots, s_w$.

(i) The syntax of the logic $SO(\mathcal{G}_1, \ldots, \mathcal{G}_w)$ is obtained by extending the syntax of SO by quantifiers $\mathcal{G}_1, \ldots, \mathcal{G}_w$.

(ii) The models $\mathcal{M}$ of the logic $SO(\mathcal{G}_1, \ldots, \mathcal{G}_w)$ are of the form

$$(\mathbb{M}, G_1, \ldots, G_w),$$

where $\mathbb{M}$ is a first order structure and $G_i \subseteq \mathcal{P}(M^{l_1^i}) \times \cdots \times \mathcal{P}(M^{l_{r_i}^i})$.

(iii) The satisfaction relation of SO is extended by the following rule for $1 \le i \le w$:

$$\mathcal{M} \models \mathcal{G}_i \bar{x}_1, \ldots, \bar{x}_{r_i}(\varphi_1, \ldots, \varphi_{r_i}) \text{ iff } (\varphi_1^{\mathcal{M}}, \ldots, \varphi_{r_i}^{\mathcal{M}}) \in G_i.$$

Now, for $\tau = \emptyset$, all τ-models of $SO(\mathcal{G}_1, \ldots, \mathcal{G}_w)$ are second order structures of type t. By Proposition 2.2, isomorphic structures are elementarily equivalent. Thus, for any τ-sentence φ the class $Mod(\varphi)$ is a second order quantifier of type t.

Proposition 2.2. *Let $(\mathbb{M}, G_1, \ldots, G_w)$ and $(\mathbb{N}, G'_1, \ldots, G'_w)$ be models and $f\colon \mathbb{M} \cong \mathbb{N}$ an isomorphism such that $f[[G_i]] = G'_i$ for $1 \le i \le w$. Then for all $\varphi(\bar{x}, \bar{X}) \in SO(\mathcal{G}_1, \ldots, \mathcal{G}_w)$, $\bar{a} = a_1 \ldots a_n \in M$, and relations $\bar{A}$,*

$$(\mathbb{M}, G_1, \ldots, G_w) \models \varphi(\bar{a}, \bar{A}) \Leftrightarrow (\mathbb{N}, G'_1, \ldots, G'_w) \models \varphi(f(\bar{a}), f(\bar{A})).$$

Proof. Induction on φ. □

In the light of the preceding discussion, we now formulate the notion of definability studied in this paper.

Definition 2.3. Let $\mathbb{Q}$ be a collection of second order quantifiers and $\mathcal{Q}$ a second order quantifier of type $t = (s_1, \ldots, s_w)$. We say that the quantifier $\mathcal{Q}$ is *definable* in the logic $SO(\mathbb{Q})$ if there is a sentence $\varphi \in SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$ of vocabulary $\tau = \emptyset$ such that for all τ-structures $(M, G_1, \ldots, G_w)$,

$$(M, G_1, \ldots, G_w) \models \varphi \Leftrightarrow (M, G_1, \ldots, G_w) \in \mathcal{Q}.$$

Theorem 2.4. *If $\mathcal{Q}$ is definable in the logic $SO(\mathbb{Q})$ then $SO(\mathbb{Q}) \equiv SO(\mathbb{Q}, \mathcal{Q})$.*

Proof. Suppose that $\varphi_{\mathcal{Q}} \in SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$ defines $\mathcal{Q}$. Let σ be a vocabulary and let $\theta_1(\bar{X}_1), \ldots, \theta_w(\bar{X}_w) \in SO(\mathbb{Q})$ be σ-formulas. We show that there is a σ-formula $\theta^* \in SO(\mathbb{Q})$ such that

$$\models \mathcal{Q}\bar{X}_1, \ldots, \bar{X}_w\,(\theta_1, \ldots, \theta_w) \leftrightarrow \theta^*.$$

Define a mapping f as follows:

$$\begin{array}{rcl} f(x = y) & = & x = y \\ f(X(\bar{y})) & = & X(\bar{y}) \\ f(\neg\varphi) & = & \neg f(\varphi) \\ f(\varphi \wedge \psi) & = & f(\varphi) \wedge f(\psi) \\ f(\exists x\varphi) & = & \exists x f(\varphi) \\ f(\mathcal{Q}'\bar{X}_1, \ldots, \bar{X}_w\,(\varphi_1, \ldots, \varphi_w)) & = & \mathcal{Q}'\bar{X}_1, \ldots, \bar{X}_w\,(f(\varphi_1), \ldots, f(\varphi_w)) \\ f(\mathcal{G}_i\bar{x}_1, \ldots, \bar{x}_{r_i}(\varphi_1, \ldots, \varphi_{r_i})) & = & \theta_i(X_1/f(\varphi_1), \ldots, X_{r_i}/f(\varphi_{r_i})) \end{array}$$

Claim 1. *Let $\mathbb{M}$ be a σ-structure and $G_i = \{\bar{A} \mid \mathbb{M} \models \theta_i(\bar{A})\}$. Then for all $\varphi(\bar{x}, \bar{X}) \in SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)[\emptyset]$, $\bar{b} = b_1 \ldots b_n \in M$, and relations $\bar{R}$:*

$$(\mathbb{M}, G_1, \ldots, G_w) \models \varphi(\bar{b}, \bar{R}) \Leftrightarrow \mathbb{M} \models f(\varphi)(\bar{b}, \bar{R}).$$

Proof. We prove the claim using induction on φ. We consider only the case $\varphi = \mathcal{G}_i\bar{x}_1, \ldots, \bar{x}_{r_i}\,(\psi_1, \ldots, \psi_{r_i})(\bar{x}, \bar{X})$. Now

$$(\mathbb{M}, G_1, \ldots, G_w) \models \varphi(\bar{b}, \bar{R}) \quad\Leftrightarrow\quad (A_1, \ldots, A_{r_i}) \in G_i,$$

where $A_i = \{\bar{d} \mid (\mathbb{M}, G_1, \ldots, G_w) \models \psi_i(\bar{d}, \bar{b}, \bar{R})\}$. By the induction hypothesis, $A_i = \{\bar{d} \mid \mathbb{M} \models f(\psi_i)(\bar{d}, \bar{b}, \bar{R})\}$. Since $\mathbb{M} \models \theta_i(A_1, \ldots, A_{r_i})$, we have

$$\mathbb{M} \models \theta_i(X_1/f(\psi_1), \ldots, X_{r_i}/f(\psi_{r_i}))(\bar{b}, \bar{R}).$$

□

We claim now that

$$\models \mathcal{Q}\bar{X}_1, \ldots, \bar{X}_w\,(\theta_1, \ldots, \theta_w) \leftrightarrow f(\varphi_{\mathcal{Q}}).$$

Let $\mathbb{M}$ be a σ-structure and $G_i = \{\bar{A} \mid \mathbb{M} \models \theta_i(\bar{A})\}$ for $1 \leq i \leq w$. By the assumption that $\varphi_{\mathcal{Q}}$ defines $\mathcal{Q}$ and Claim 1,

$$\begin{array}{rcl} \mathbb{M} \models \mathcal{Q}\bar{X}_1, \ldots, \bar{X}_w\,(\theta_1, \ldots, \theta_w) & \Leftrightarrow & (\mathbb{M}, G_1, \ldots, G_w) \models \varphi_{\mathcal{Q}} \\ & \Leftrightarrow & \mathbb{M} \models f(\varphi_{\mathcal{Q}}). \end{array}$$

□

The following example shows that a quantifier $\mathcal{Q}$ can be always defined using itself in a uniform way.

Example 2.5. Let $\mathcal{Q}$ be a quantifier of type $((1,1),(2))$. Then $\mathcal{Q}$ is definable in the logic $SO(\mathcal{Q})$ with the sentence

$$\psi = \mathcal{Q}XY, Z\,(\mathcal{G}_1 x, y\,(X(x), Y(y)), \mathcal{G}_2 \bar{z}\,(Z(\bar{z}))).$$

Note that ψ is uniform in the sense that if $\mathcal{Q}'$ is another quantifier of type $((1,1),(2))$, then it is enough to replace $\mathcal{Q}$ by $\mathcal{Q}'$ in ψ to acquire a defining sentence for $\mathcal{Q}'$ in $SO(\mathcal{Q}')$. A defining sentence like ψ can be easily formulated for quantifiers of arbitrary types.

The next theorem shows that the implication in Theorem 2.4 cannot be reversed.

Theorem 2.6. *There is a quantifier $\mathcal{Q}$ of type $((1))$ which is undefinable in SO and satisfies $SO \equiv SO(\mathcal{Q})$.*

Proof. There is a quantifier $\mathcal{Q}$ of type $((1))$ such that it contains structures in every cardinality and for every quantifier $\mathcal{Q}' \subseteq \mathcal{Q}$,

$$SO \equiv SO(\mathcal{Q}').$$

For a proof, see Theorem 3.1 in [1]. Now, for $A \subseteq \omega$, define

$$\mathcal{Q}_A = \{(M, G) \in \mathcal{Q} \mid |M| \in A\}.$$

The logic $SO(\mathcal{G})$ is countable, hence $\mathcal{Q}_A$ is undefinable in SO for some $A \subseteq \omega$. □

2.1 Fagin's lemma for second order structures

In this section we study the number of second order structures with universe $\{1, \dots, n\}$. We prove a version of Fagin's lemma for second order structures. The proof follows the proof in [3] closely. Let $s = (l_1, \dots, l_r)$. We write $P_s(x)$ for $x^{l_1} + \cdots + x^{l_r}$. Note that $2^{P_s(n)} = a_n(s)$.

Lemma 2.7. *Let $t = (s_1, \dots, s_w)$ be a second order type where $s_i = (l_1^i, \dots, l_{r_i}^i)$ for $1 \leq i \leq w$. Then*

$$a_n(t) = 2^{(2^{P_{s_1}(n)} + \cdots + 2^{P_{s_w}(n)})}.$$

Proof. Let $M = \{1, \dots, n\}$. Clearly $a_n(t)$ is equal to the cardinality of the set

$$\prod_{i=1}^{w} (\mathcal{P}(\prod_{j=1}^{r_i} \mathcal{P}(M^{l_j^i}))).$$

□

Definition 2.8. Let X be a set and π a permutation of X.

- Denote by $spt(\pi)$ the *support* of π
$$spt(\pi) = \{a \in X \mid \pi(a) \neq a\}.$$
- The cardinality of the set $spt(\pi)$ is denoted by $s(\pi)$.
- The number of cycles of π is denoted by $c(\pi)$.

The next lemma is a corollary of Burnside's Lemma [2]. It will be used in the proof of Theorem 2.10. In the following S_n denotes the set of permutations of the set $\{1, \dots, n\}$.

Lemma 2.9. *Let t be a second order type and let π be a permutation of the set $M = \{1, \dots, n\}$. Let $N(\pi)$ denote the number of t-structures with universe M for which π is an automorphism. Then*

$$b_n(t) = \frac{1}{n!} \sum_{\pi \in S_n} N(\pi).$$

Proof. The claim follows directly from Burnside's Lemma [2]. □

Theorem 2.10. *Let t be a second order type. Then $b_n(t) \sim \frac{1}{n!} a_n(t)$.*

Proof. We first prove the claim for type $t = ((1))$. Let π be a permutation of the set $\{1, \dots, n\}$. Denote by $\pi_{(1)}$ the permutation of $\mathcal{P}(\{1, \dots, n\})$ defined by $\pi_{(1)}(X) = \pi[X]$.

Claim 2. *Let π be a permutation of the set $\{1, \dots, n\}$. Then $N(\pi) = 2^{c(\pi_{(1)})}$.*

Proof. The cycles of $\pi_{(1)}$ form a partition P of the set $\mathcal{P}(\{1, \dots, n\})$. It is easy to see that π is an automorphism of $(\{1, \dots, n\}, G)$ if and only if G is a union of some elements of P. □

By Lemma 2.9,
$$b_n(t) = \frac{1}{n!} \sum_{\pi \in S_n} N(\pi). \tag{1}$$

Combining (1) and Claim 2,
$$b_n(t) = \frac{1}{n!} \sum_{\pi \in S_n} 2^{c(\pi_{(1)})}.$$

The original claim is equivalent to
$$\lim_{n \to \infty} \frac{n! b_n(t)}{a_n(t)} = 1.$$

Hence, it suffices to prove

$$\lim_{n\to\infty}\sum_{\pi\in S_n} 2^{(c(\pi_{(1)})-2^n)} = 1.$$

For $\pi = id$, $c(\pi_{(1)}) = 2^n$ thus it suffices to show

$$\lim_{n\to\infty}\sum_{\pi\in S'_n} 2^{(c(\pi_{(1)})-2^n)} = 0, \tag{2}$$

where $S'_n = S_n \setminus \{id\}$.

Claim 3. *Let π be a permutation of the set $\{1,\ldots,n\}$. If $s(\pi) = m$ then $s(\pi_{(1)}) \geq m(n-2)$.*

Proof. Obviously m singletons are not fixed by $\pi_{(1)}$. The number of sets $\{a,b\}$ such that $\pi(a) = a$ and $\pi(b) \neq b$ is $m(n-m)$. Finally, the number of sets $\{a,b\}$ such that $\pi(a) \notin \{a,b\}$ and $\pi(b) \notin \{a,b\}$ is $m(m-3)$. By summing these numbers, we get $m(n-2)$. $\square$

If $s(\pi_{(1)}) = x$, then clearly

$$c(\pi_{(1)}) \leq 2^n - x/2. \tag{3}$$

By the estimate (3) and Claim 3, the terms of the sequence in (2) are dominated by

$$\sum_{\pi\in S'_n} 2^{-(s(\pi)(n-2))/2}. \tag{4}$$

The number of permutations π of $\{1,\ldots,n\}$ such that $s(\pi) = m$ is less than $\binom{n}{m}m! \leq n^m$. Therefore, the sequence in (4) is dominated by

$$\sum_{m=2}^{n} n^m 2^{-m(n-2)/2}. \tag{5}$$

Now (5) is equivalent to

$$\sum_{m=2}^{n} 2^{-\frac{1}{2}m((n-2)-2log(n))}. \tag{6}$$

Let n_0 be the least n such that $(n-2) - log(n)$ is positive. If $n \geq n_0$ then the greatest summand in (6) occurs when $m = 2$. Thus, for $n \geq n_0$, (6) is dominated by

$$(n-1)2^{(2log(n)-(n-2))} = 2^{(2log(n)+log(n-1)-(n-2))},$$

which converges to 0.

Next we turn to the general case. Let $t = (s_1, \ldots, s_w)$ and let π be a permutation of the set $M = \{1, \ldots, n\}$. Denote by π_{s_i} the permutation of the set

$$\mathcal{P}(M^{l^i_1}) \times \cdots \times \mathcal{P}(M^{l^i_{r_i}})$$

induced by π. Then (cf. Claim 2) π is an automorphism of $N(\pi)$ many t-structures where

$$N(\pi) = 2^{\sum_{i=1}^{w} c(\pi_{s_i})}.$$

By Lemma 2.9,

$$b_n(t) = \frac{1}{n!} \sum_{\pi \in S_n} N(\pi)$$

and thus

$$b_n(t) = \frac{1}{n!} \sum_{\pi \in S_n} 2^{\sum_{i=1}^{w} c(\pi_{s_i})}.$$

The original claim is equivalent to

$$\lim_{n \to \infty} \frac{n! b_n(t)}{a_n(t)} = 1.$$

Hence, it suffices to prove

$$\lim_{n \to \infty} \sum_{\pi \in S_n} 2^{\sum_{i=1}^{w} (c(\pi_{s_i}) - 2^{P_{s_i}(n)})} = 1.$$

Since $c(\pi_{s_i}) = 2^{P_{s_i}(n)}$ for $\pi = id$, it remains to prove

$$\lim_{n \to \infty} \sum_{\pi \in S'_n} 2^{\sum_{i=1}^{w} (c(\pi_{s_i}) - 2^{P_{s_i}(n)})} = 0. \tag{7}$$

Suppose $s(\pi) = m$. We want to find a lower bound for $s(\pi_{s_i})$. If $l^i_1 = 1$, then based on the calculations made in Claim 3 the number of sets A not fixed in the first coordinate is greater than $m(n-2)$. Thus, no tuple of the form

$$(A, B_2, \ldots, B_{r_i}),$$

where $B_j \in \mathcal{P}(M^{l^i_j})$ for $2 \leq j \leq r_i$, is fixed by π_{s_i}. Hence, we have $s(\pi_{s_i}) > m(n-2)$. If $l^i_1 = q > 1$, then $\pi_q \colon M^q \to M^q$ defined applying π point-wise moves more than mn^{q-1} many tuples. Therefore, the number of singletons which are not fixed by π_{s_i} in the first coordinate is greater than

$$mn^{q-1} > m(n-2).$$

As in (3) the terms of the sequence in (7) are dominated by

$$\sum_{\pi \in S'_n} 2^{-w(s(\pi)(n-2))/2}$$

and thus by

$$\sum_{\pi \in S'_n} 2^{-(s(\pi)(n-2))/2}. \tag{8}$$

Sequence (8) is identical to sequence (4) which converges to 0. ☐

2.2 An ordering of second order types

In this section we define an ordering for second order types arising naturally from Theorem 2.10. The following results concerning first order types and structures are taken from [4]. Let S denote the set of finite sequences $(p_1, \ldots, p_a)$ of natural numbers such that $p_a \neq 0$. We denote a by $l(p)$.

Definition 2.11. Let p and q be elements of S. We set $p < q$ if

(i) $l(p) < l(q)$ or

(ii) $l(p) = l(q)$ and $p_i < q_i$ for the largest i such that $p_i \neq q_i$.

Proposition 2.12. *The relation $<$ well-orders the set S and the order-type of $<$ is ω^ω.*

Proof. See [4]. ☐

We denote the order-isomorphism between $(S, <)$ and $(\omega^\omega, \epsilon)$ by Φ. For a first order type $s = (l_1, \ldots, l_r)$, the *pattern* of s is the tuple $p(s) = (p_1, \ldots, p_a)$, where $p_i = |\{1 \leq j \leq r : l_j = i\}|$ and $a = max\{l_j : 1 \leq j \leq r\}$.

Definition 2.13. Let s and s' be first order types. Define $s < s'$ to hold if $p(s) < p(s')$.

Lemma 2.14. *Let s and s' be first order types such that $s < s'$. Then*

$$\lim_{n \to \infty} \frac{a_n(s)}{a_n(s')} = 0.$$

Proof. See [4]. ☐

Now we turn to second order types. We define an ordering for second order types analogously to Definition 2.13.

Definition 2.15. Let B denote the set

$$\{j : \omega^\omega \to \omega \mid |\{\alpha \in \omega^\omega \mid j(\alpha) > 0\}| \in \omega \setminus \{0\}\}.$$

We define an ordering $\prec$ for the set B as follows. Let j and j' be elements of B. Define $j \prec j'$ to hold if $j'(\beta) > j(\beta)$, where β is the greatest κ such that $j(\kappa) \neq j'(\kappa)$.

It is easy to see that $\prec$ is a linear-ordering.

Proposition 2.16. *The ordering $\prec$ well-orders the set B and the order-type of $\prec$ is ω^{ω^ω}.*

Proof. Let $j \in B$ and let $\{\alpha_1, \ldots, \alpha_n\}$ ($\alpha_i < \alpha_j$ iff $i < j$) be the set of those ordinals α for which $j(\alpha) \neq 0$. We divide j into component functions $j_1, \ldots, j_n$ by setting $j_i(\alpha_i) = j(\alpha_i)$ and for $\beta \neq \alpha_i$, we set $j_i(\beta) = 0$. Assume that $j \in B$ and $j(\alpha) \neq 0$ for exactly one α. If this α is equal to some $k \in \omega$ we define $F(j) = \omega^{k-1} j(k)$, and if $\omega \leq \alpha$ we define $F(j) = \omega^\alpha j(\alpha)$. Finally, we extend F to a function from the set B to ω^{ω^ω} in the following way. Let $j \in B$ and let $j_1, \ldots, j_n$ be the component functions of j. Define

$$F(j) = F(j_n) + F(j_{n-1}) + \cdots + F(j_1). \tag{9}$$

Since every ordinal $\alpha < \omega^{\omega^\omega}$ has a unique representation in the form (9), the function F is bijective. It is easy to see that $j \prec j'$ if and only if $F(j) \in F(j')$. Thus F is an order-isomorphism. $\square$

Definition 2.17. Let $t = (s_1, \ldots, s_w)$ be a second order type. Define $j_t \colon \omega^\omega \to \omega$ by

$$j_t(\alpha) = |\{1 \leq i \leq w \mid \Phi(p(s_i)) = \alpha\}|.$$

Let t and t' be second order types. We define $t \prec t'$ to hold if $j_t \prec j_{t'}$.

Lemma 2.18. *Let $t = (s_1, \ldots, s_w)$ and $t' = (d_1, \ldots, d_e)$ be second order types such that $t' \prec t$. Then*

$$\lim_{n\to\infty} \frac{b_n(t')}{b_n(t)} = 0.$$

Proof. By Theorem 2.10,

$$\frac{b_n(t')}{b_n(t)} \sim 2^{((2^{P_{d_1}(n)} + \cdots + 2^{P_{d_e}(n)}) - (2^{P_{s_1}(n)} + \cdots + 2^{P_{s_w}(n)}))}. \tag{10}$$

After subtraction and by the assumption $t' \prec t$, the right-hand side of (10) is dominated by

$$2^{(e2^{P_d(n)} - 2^{P_s(n)})},$$

where $d < s$. By Lemma 2.14, $\lim_{n\to 0} P_d(n)/P_s(n) = 0$. Thus

$$\lim_{n\to 0} 2^{(e2^{P_d(n)} - 2^{P_s(n)})} = 0.$$

$\square$

Corollary 2.19. *Let $t, t_1, \ldots, t_e$ be second order types such that $t_i \prec t$ for $1 \leq i \leq e$. Then*

$$\lim_{n\to\infty} \sum_{i=1}^{e} \frac{b_n(t_i)}{b_n(t)} = 0.$$

2.3 A Hierarchy of definability

In this section we prove that on every level of the ordering $\prec$ there is a quantifier $\mathcal{Q}$, say, of type t which is not definable in the extension of second order logic by all second order quantifiers of types lower than t. The proof of Theorem 2.21 is analogous to the proof of Theorem 5 in [4]. We begin with an auxiliary definition [4].

Definition 2.20. Let φ be a formula and let $\mathcal{Q}_1, \ldots, \mathcal{Q}_y$ be the second order quantifiers appearing in φ. Then we write $\varphi(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)$. If $\mathcal{Q}'_1, \ldots, \mathcal{Q}'_y$ are quantifiers such that the type of $\mathcal{Q}'_i$ is the same as the type of $\mathcal{Q}_i$ for $1 \leq i \leq y$, then $\varphi(\mathcal{Q}'_1, \ldots, \mathcal{Q}'_y)$ denotes the formula obtained from φ by changing $\mathcal{Q}_i$ everywhere to $\mathcal{Q}'_i$. Formulas obtained from each others by a change of quantifiers are called *similar*.

Theorem 2.21. *Let t be a second order type and $\mathbb{Q}$ the collection of all second order generalized quantifiers of types lower than t. Then there is a quantifier $\mathcal{Q}$ of type t which is not definable in the logic $SO(\mathbb{Q})$.*

Proof. Let $\tau = \emptyset$ and let $(\psi_i)_{i\in\omega}$ be a list of τ-sentences of the logic $SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$ such that every τ-sentence of $SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$ is similar to some ψ_i. Suppose $n_0, \ldots, n_{j-1}$ and classes $K_0, \ldots, K_{j-1}$ have been defined already. Let $\mathcal{Q}_1, \ldots, \mathcal{Q}_y$ be the quantifiers appearing in ψ_j. The type of $\mathcal{Q}_i$ is denoted by t_i. Let $\mathcal{Q}'_1, \ldots, \mathcal{Q}'_y$ be quantifiers such that the type of $\mathcal{Q}'_i$ is t_i for $1 \leq i \leq y$. If

$$\mathcal{Q}_i \cap Str(t_i, n) = \mathcal{Q}'_i \cap Str(t_i, n)$$

for $1 \leq i \leq y$, then the sequences $(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)$ and $(\mathcal{Q}'_1, \ldots, \mathcal{Q}'_y)$ are called *n-equivalent*. It is easy to see that for n-equivalent sequences,

$$Mod(\psi_j(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)) \cap Str(t, n) = Mod(\psi_j(\mathcal{Q}'_1, \ldots, \mathcal{Q}'_y)) \cap Str(t, n).$$

The number of n-equivalence classes of sequences is at most $2^{p(n)}$, where $p(n) = \sum_{i=1}^{y} b_n(t_i)$. By Corollary 2.19, there is $n > n_{j-1}$ such that $b_n(t) > p(n)$. Let n_j be the least such n. Then $2^{b_{n_j}(t)} > 2^{p(n_j)}$ and hence there is a class K_j of t-structures of cardinality n_j such that K_j cannot be represented in the form

$$Mod(\psi_j(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)) \cap Str(t, n_j),$$

for any sequence $(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)$ of quantifiers. This concludes the construction. We define $\mathcal{Q} = \bigcup_{j\in\omega} K_j$.

Claim 4. *The quantifier $\mathcal{Q}$ is not definable in the logic $SO(\mathbb{Q})$.*

Proof. Suppose that $\mathcal{Q} = Mod(\varphi)$ for some $\varphi \in SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$. Then $\varphi = \psi_i(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)$ for some $i \in \omega$ and quantifiers $\mathcal{Q}_1, \ldots, \mathcal{Q}_y$. Thus

$$K_i = Mod(\psi_i(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)) \cap Str(t, n_i),$$

which contradicts the definition of $\mathcal{Q}$. □

This completes the proof of Theorem 2.21 . □

Next we show that the diagonalization argument used in the proof of Theorem 2.21 can be modified so that the resulting undefinable quantifier can be extended to a undefinable monotone quantifier. The proof is adapted from Theorem 7 in [4] where an analogous result is proved for Lindström quantifiers.

Definition 2.22. Let $\mathcal{Q}$ be a second order quantifier. We say that $\mathcal{Q}$ is a *monotone quantifier* if for every structure $(M, G_1, \ldots, G_w) \in \mathcal{Q}$ and $G'_1, \ldots, G'_w$, where $G'_i \subseteq \mathcal{P}(M^{l^i_1}) \times \cdots \times \mathcal{P}(M^{l^i_{r_i}})$ and $G_i \subseteq G'_i$ for $1 \leq i \leq w$, we have that $(M, G'_1, \ldots, G'_w) \in \mathcal{Q}$.

Theorem 2.23. *Let t be a type and $\mathbb{Q}$ the collection of all second order generalized quantifiers of types lower than t. Then there is a monotone quantifier $\mathcal{Q}$ of type t which is not definable in the logic $SO(\mathbb{Q})$.*

Before going to the proof of Theorem 2.23, we introduce an auxiliary concept. Let $M = \{1, \ldots, n\}$ and let $(M, G_1, \ldots, G_w)$ be a t-structure. Note that

$$0 \leq \sum_{i=1}^{w} |G_i| \leq \sum_{i=1}^{w} 2^{P_{s_i}(n)}.$$

Suppose $\sigma \colon \omega \to \omega$. We say that $(M, G_1, \ldots, G_w)$ is *σ-balanced* if $\sum_{i=1}^{w} |G_i| = \sigma(n)$. Define a function $f \colon \omega \to \omega$ as follows: $f(n)$ is the least y such that

$$b_n(t) \leq (\sum_{i=1}^{w} 2^{P_{s_i}(n)} + 1) N_y,$$

where N_y is the number of isomorphism types of t-structures with universe $\{1, \ldots, n\}$ satisfying $\sum_{i=1}^{w} |G_i| = y$. Let $Str_f(t) = \{\mathcal{M} \in Str(t) \mid \mathcal{M}$ is f-balanced$\}$.

Lemma 2.24. *Let $t = (s_1, \ldots, s_w)$ and $t' = (d_1, \ldots, d_e)$ be types such that $t' \prec t$. Then*

$$\lim_{n \to \infty} \frac{b_n(t')}{I_n(Str_f(t))} = 0.$$

Proof. By the choice of f, we have

$$I_n(Str_f(t)) \geq \frac{b_n(t)}{\sum_{i=1}^{w} 2^{P_{s_i}(n)} + 1}.$$

Therefore,

$$\frac{b_n(t')}{I_n(Str_f(t))} \leq \frac{b_n(t')(\sum_{i=1}^{w} 2^{P_{s_i}(n)} + 1)}{b_n(t)}. \tag{11}$$

By the estimate in Lemma 2.18, the right-hand side of (11) is eventually dominated by

$$c(\sum_{i=1}^{w} 2^{P_{s_i}(n)} + 1)2^{(e2^{P_d(n)} - 2^{P_s(n)})}, \tag{12}$$

where $c > 1$ and $d < s$. By Lemma 2.14, $\lim_{n\to 0} 2^{P_d(n)}/2^{P_s(n)} = 0$, hence there are $u, h, r \in \omega$ such that, for $n \geq r$, the sequence in (12) is dominated by

$$(2^{n^u} + h)2^{-2^n} \longrightarrow 0.$$

$\square$

Proof of Theorem 2.23. We modify the proof of Theorem 2.21 so that the diagonalization is realized with respect to the class $Str_f(t)$. In other words, we ensure that

$$K_j \neq Mod(\psi_j(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)) \cap Str_f(t, n_j),$$

for all sequences $(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)$ of quantifiers. This is possible by Lemma 2.24, since the number of collections of the form

$$Mod(\psi_j(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)) \cap Str_f(t, n_j)$$

is at most the number of n_j-equivalence classes of sequences of quantifiers. In particular, we choose $K_j \subseteq Str_f(t)$ for all $j \in \omega$. Let $\mathcal{Q} = \bigcup_{i\in\omega} K_i$, and let $\mathcal{Q}'$ be the following monotone quantifier

$$\{(M, G_1, \ldots, G_w) \mid \forall i(G_i' \subseteq G_i) \text{ for some } (M, G_1', \ldots, G_w') \in \mathcal{Q}\}.$$

Since $\mathcal{Q}' \cap Str_f(t) = \mathcal{Q}$ and $\mathcal{Q} \neq (Mod(\varphi) \cap Str_f(t))$ for all $\varphi \in SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$, the quantifier $\mathcal{Q}'$ is undefinable in the logic $SO(\mathbb{Q})$. $\square$

Theorem 2.21 can be also improved to yield an undefinable *recursive quantifier*, i.e., there is a *Turing machine* which accepts a coding of a structure $(M, G_1, \ldots, G_w)$ if and only if $(M, G_1, \ldots, G_w) \in \mathcal{Q}$.

Theorem 2.25. *Let t be a type and $\mathbb{Q}$ the collection of all second order generalized quantifiers of types lower than t. Then there is a recursive quantifier $\mathcal{Q}$ of type t which is not definable in the logic $SO(\mathbb{Q})$.*

The algorithm we use is the same as in the first order case. However, for reader's convenience, we explain the idea of the algorithm. The following, modulo notational changes, is taken from [4].

Proof of Theorem 2.25. Let $(\psi_i)_{i\in\omega}$ be an effective enumeration of t-sentences of $SO(\mathbb{Q}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$ up to similarity.

`Input`: $\mathcal{M} = (M, G_1, \ldots, G_w)$ of type t and cardinality n.
`Algorithm`:
  $j \leftarrow 0$; $n_{-1} \leftarrow 0$;
  `repeat` $\{n_j \leftarrow min\{n' > n_{j-1} \mid \texttt{Cond}(n', j)\}$; $j \leftarrow j+1\}$
  `until` $n_j \geq n$;
  `if` $n_j > n$
  `then reject` $\mathcal{M}$
  `else`
    $\{$`FindQuant`(K, n, j);
    `if` $\mathcal{M} \in K$
    `then accept` $\mathcal{M}$
    `else reject` $\mathcal{M}\}$.

The condition `Cond(n,j)` is the condition $b_n(t) > \sum_{i=1}^{y} b_n(t_i)$ where t_i refer to the types of quantifiers occurring in ψ_j. In other words, the algorithm computes for each j the least $n_j > n_{j-1}$ (until $n_j \geq n$) such that there is guaranteed to exist a class of t-structures of cardinality n_j which cannot be represented in the form

$$Mod(\psi_j(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)) \cap Str(t, n_j), \tag{13}$$

for any sequence $(\mathcal{Q}_1, \ldots, \mathcal{Q}_y)$ of quantifiers. If $n = n_j$ for some j then the procedure `FindQuant`(K, n, j) returns a canonical class $K(n, j)$ of t-structures of cardinality n which cannot be represented in the form (13). The structure $\mathcal{M}$ is accepted by the algorithm if $n = n_j$ for some j and $\mathcal{M} \in K(n, j)$. So, we have $\mathcal{Q} = \bigcup_{j\in\omega} K(n_j, j)$.

The procedure `FindQuant` is computed as follows:

  $K \leftarrow \emptyset$;
  `repeat`
    $\{K \leftarrow$ `Next`(K);
    `repeat`
      $\{\bar{\mathcal{Q}} \leftarrow (\emptyset, \ldots, \emptyset)$;
      $b \leftarrow$ `Definable`$(K, \bar{\mathcal{Q}}, j)$;
      $\bar{\mathcal{Q}} \leftarrow$ `NextSeq`$(\bar{\mathcal{Q}})\}$

`until` b `or` $\bar{\mathcal{Q}} = (\emptyset, \ldots, \emptyset)\}$
`until not` b;

To ensure that the class of t-structures accepted by the algorithm is closed under isomorphisms, the algorithm first defines a canonical linear order $<$ of the structure $\mathcal{M}$. For each type t, the order $<$ induces a lexicographic order $<_t$ between the subsets of $Str(t, M)$, where $Str(t, M)$ is the set of t-structures with universe M. The procedure `Next`($\mathcal{Q}$) returns the least quantifier $\mathcal{Q}' \subseteq Str(t, M)$ which is greater than $\mathcal{Q}$ with respect to $<_t$. If such $\mathcal{Q}'$ does not exist, `Next`($\mathcal{Q}$) $= \emptyset$. `NextSeq`($\bar{\mathcal{Q}}$) is a similar function for sequences of quantifiers of types lower than t. The condition `Definable`($\mathcal{Q}, \bar{\mathcal{Q}}, j$) is satisfied just in case $\mathcal{Q} = \{\mathcal{M}' \in Str(t, M) \mid \mathcal{M}' \models \psi_j(\bar{\mathcal{Q}})\}$. □

Suppose $((1)) \prec t$ and $\mathcal{Q}$ is a quantifier of type t as in Theorem 2.21. Although $\mathcal{Q}$ is undefinable in the logic $SO(\mathbb{Q})$, where $\mathbb{Q}$ is the collection of quantifiers of types lower than t, we are bound to have

$$SO(\mathbb{Q}, \mathcal{Q}) \equiv SO(\mathbb{Q}).$$

The reason is that any class of finite first order structures is axiomatizable in $SO(M_1^1)$, where M_1^1 is the collection of quantifiers of type $((1))$ (cf. Theorem 6.4 and Corollary 6.5 in [1]). In fact, even the following version of Theorem 2.21 holds.

Theorem 2.26. *Let t be a second order type and $\mathbb{Q}$ the collection of all second order generalized quantifiers of types lower than t. Then there is a quantifier $\mathcal{Q}$ of type t which is not definable in the logic $SO(\mathbb{Q})$ and $FO \equiv FO(\mathcal{Q})$.*

The proof of Theorem 2.26 is based on the following lemma.

Lemma 2.27. *Let t be a type. Then there is a quantifier $\mathcal{Q}_0$ of type t such that $\mathcal{Q}_0$ contains more than half of the isomorphism types of structures in every cardinality $n \in \omega \setminus \{0\}$ and for every quantifier $\mathcal{Q} \subseteq \mathcal{Q}_0$,*

$$FO \equiv FO(\mathcal{Q}).$$

Proof. Straightforward generalization of the proof of Theorem 3.1 in [5]. □

Proof of Theorem 2.26. We modify the diagonalization argument in the proof of Theorem 2.21 as follows. We let n_j be the least n such that $\frac{1}{2}b_n(t) > p(n)$, where $p(n) = \sum_{i=1}^{y} b_n(t_i)$. By the choice of n_j, we can choose K_j so that $K_j \subseteq \mathcal{Q}_0$. Since $\bigcup_{j\in\omega} K_j \subseteq \mathcal{Q}_0$, the claim follows by Lemma 2.27. □

Acknowledgments

I am grateful to Professor Jouko Väänänen, who is the supervisor of my doctoral studies. I would also like to thank Professor Lauri Hella for his valuable comments.

References

[1] A. Andersson. On second-order generalized quantifiers and finite structures. *Ann. Pure Appl. Logic*, 115(1-3):1–32, 2002.

[2] W. Burnside. *Theory of groups of finite order*. Dover Publications Inc., New York, 1955. 2d ed.

[3] R. Fagin. The number of finite relational structures. *Discrete Math.*, 19(1):17–21, 1977.

[4] L. Hella, K. Luosto, and J. Väänänen. The hierarchy theorem for generalized quantifiers. *J. Symbolic Logic*, 61(3):802–817, 1996.

[5] J. Kontinen. Definability of second-order generalized quantifiers. Preprint 380, Department of Mathematics and Statistics, University of Helsinki, 2004.

Department of Mathematics and Statistics
P.O. Box 4 (Yliopistonkatu 5)
FIN-00014 University of Helsinki
Finland
E-mail: juha.kontinen@helsinki.fi