

DEFINABILITY OF SECOND ORDER GENERALIZED QUANTIFIERS

JUHA KONTINEN

Academic dissertation

*To be presented, with the permission of the Faculty of Science
of the University of Helsinki, for public criticism in Auditorium 5,
the Main Building of the University, on December 18th, 2004, at 10 a.m.*

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE
DEPARTMENT OF MATHEMATICS AND STATISTICS

HELSINKI 2004

ISBN 952-91-8109-4 (paperback)
ISBN 952-10-2237-X (PDF)

HELSINKI UNIVERSITY PRINTING HOUSE
HELSINKI 2004

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my supervisor Professor Jouko Väänänen for introducing me to this subject and for his inspiring guidance and support during my doctoral studies. I am also grateful to Professor Lauri Hella and Docent Kerkko Luosto for carefully reading the manuscript and making valuable comments on it.

I am indebted to the Graduate School of Mathematical Analysis and Logic and to the Finnish Academy of Science and Letters (Vilho, Yrjö and Kalle Väisälä Foundation) for financial support.

I want to thank my friends and colleagues at the Department of Mathematics and Statistics and the participants of the Finite Model Theory Seminar for their support and interest. Especially, I want to thank Ph.Lic. Pekka Pankka and Ph.Lic. Jussi Laitila.

I express my best thanks to my parents Raili and Esko for their continuous encouragement and support. My thanks are also due to my brother Jarmo. Finally, I am grateful to Ilona for her love and her important support during my doctoral studies.

Helsinki, December 2004

Juha Kontinen

CONTENTS

1. Introduction	5
I. Preliminaries	8
2. Notation	8
3. Lindström quantifiers	8
4. Definability of Lindström quantifiers	9
5. Second order generalized quantifiers	10
II. A notion of definability for second order quantifiers	12
6. Extending first order logic by second order predicates	12
III. Definability in terms of Lindström quantifiers	15
7. Preliminaries	15
8. Some definable quantifiers	15
9. Proving non-definability	17
10. Looking for a complete characterization of definability	20
11. Quantifiers of arbitrary types	21
12. Further non-definability results	22
IV. Definability in monadic second order logic	26
13. Some non-definability results	26
V. Definability in extensions of second order logic	28
14. Definability and types	28
15. Fagin's lemma for second order structures	29
16. An ordering of second order types	33
17. A hierarchy of definability	35
18. A hierarchical quantifier	39
19. Conclusion and further directions	41
References	42

1. INTRODUCTION

Many interesting properties of finite structures cannot be expressed in first order logic (FO). Therefore, various extensions of first order logic have been studied in the context of finite model theory. First order logic can be strengthened for example by adding fixed-point operators, infinitary connectives, higher order quantification or generalized quantifiers. Lindström quantifiers [11], i.e., first order generalized quantifiers, give direct means to extend logics in a controlled and minimal way. The notion of second order generalized quantifier [1] combines the idea of Lindström quantifiers with that of higher order quantification. In this thesis we study questions concerning definability of second order generalized quantifiers. We show that some of the results on Lindström quantifiers carry over to the second order setting. On the other hand, some striking differences are discovered between first order and second order theory of definability.

First order logic cannot express, e.g., that a formula holds for an even number of elements. To acquire an extension of FO with this feature, we can extend it by a new quantifier Q_{even} with interpretation given by

$$\mathbb{M} \models Q_{\text{even}} x \varphi(x) \Leftrightarrow |\varphi^{\mathbb{M}}| \text{ is even,}$$

where $\varphi^{\mathbb{M}} = \{a \in M \mid \mathbb{M} \models \varphi(a)\}$. Let P be a unary predicate symbol. In general, the interpretation of a Lindström quantifier Q of type (1) is given by

$$\mathbb{M} \models Qx \varphi(x) \Leftrightarrow (M, \varphi^{\mathbb{M}}) \in K,$$

where M denotes the universe of \mathbb{M} and K is a class of $\{P\}$ -structures which is closed under isomorphisms. In particular, the class K corresponding to Q_{even} is

$$\{(M, A) \mid A \subseteq M \text{ and } |A| \text{ even}\}.$$

From now on, we identify a quantifier with the class of structures which interprets it. In other words, Q denotes a class of structures or a quantifier symbol depending on the context.

Intuitively, a quantifier Q is definable in terms of the quantifiers Q_1, \dots, Q_n if there is a uniform way to express what Q "says" of a formula using the quantifiers

Q_1, \dots, Q_n . The question of expressibility of Q in terms of Q_1, \dots, Q_n actually reduces to the question of axiomatizability of the class Q in the logic $\text{FO}(Q_1, \dots, Q_n)$. Indeed, if $Q = \text{Mod}(\psi)$ then

$$\models Qx \varphi(x) \leftrightarrow \psi(P/\varphi),$$

where $\psi(P/\varphi)$ is obtained by substituting every occurrence of $P(x)$ in ψ by $\varphi(x)$. The same argument shows that if the logic \mathcal{L} is closed under first order operations and has the substitution property then axiomatizability of Q in \mathcal{L} implies that $\mathcal{L} \geq \text{FO}(Q)$. Therefore, the logic $\text{FO}(Q)$ is the minimal such logic which can express Q .

Definability questions of Lindström quantifiers have been studied extensively in finite model theory. In [9] Kolaitis and Väänänen combined Lindström quantifiers with the infinitary logic $\mathcal{L}_{\infty\omega}^\omega$ in which every formula has only a finite number of variables. They showed that equivalence of structures relative to $\mathcal{L}_{\infty\omega}^\omega(Q_i)_{i \in I}$ can be characterized in terms of an Ehrenfeucht-Fraïssé type pebble game. This game is then used to prove, e.g., that the Härtig quantifier (cf. Example 3.2) is not definable in $\mathcal{L}_{\infty\omega}^\omega(Q_i)_{i \in I}$ for any finite collection $\{Q_i\}_{i \in I}$ of quantifiers of type (1). In [7] Hella studied Lindström quantifiers in connection with descriptive complexity theory. He showed that no extension of fixed point logic by finitely many Lindström quantifiers captures PTIME on unordered structures. On the other hand, in [3] Dawar proved that if PTIME can be represented as a logic at all, then there is a Lindström quantifier Q such that $\text{FO}(\{Q^{(n)}\}_{n \in \mathbb{N}})$ captures PTIME, where $Q^{(n)}$ is the so-called n -vectorization of Q .

Second order generalized quantifiers alter the definition of Lindström quantifiers in a straightforward way. Let us first restrict attention to second order generalized quantifiers of type ((1)), i.e., to quantifiers which apply to a single formula and bind 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 [1] Andersson studies the expressive power of second order generalized quantifiers on finite structures. He shows that on finite structures with at most binary relations almost any countable logic is equivalent to a uniformly obtained sublogic of $\text{FO}(\mathcal{Q})$, where \mathcal{Q} is some second order generalized quantifier of type ((1)), and that the result extends to all finite structures if \mathcal{Q} is allowed to be of type ((2)). He also shows that there is a quantifier \mathcal{Q} of type ((1)) and a Lindström quantifier Q such that $\text{FO} \equiv \text{FO}(\mathcal{Q})$ but $\text{FO}(Q) < \text{FO}(\mathcal{Q}, Q)$.

In this thesis we study definability of second order generalized quantifiers. The questions we study are very similar to questions that have been studied extensively in the case of Lindström quantifiers. Indeed, the notion of definability we shall

formulate is analogous to the notion of definability of Lindström quantifiers. Although, already monadic second order generalized quantifiers can be used to define any Lindström quantifier [1], monadic second order generalized quantifiers turn out to be relatively weak in defining second order quantifiers of non-monadic types.

Let us look at an example of definability in the second order setting.

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

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

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 $\text{SO}(\mathcal{Q})$ in a uniform way.

Definability questions of Lindström quantifiers can be reduced to questions concerning axiomatizability of classes of first order structures. Second order generalized quantifiers are classes of second order structures so 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) the interpretations of which are given locally and arbitrarily in every model. 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 second order predicates. It turns out that a defining sentence $\psi \in \mathcal{L}'$ can be converted into a uniform definition of \mathcal{Q} in the logic \mathcal{L} . So definability of \mathcal{Q} in \mathcal{L} implies that $\mathcal{L}(\mathcal{Q}) \equiv \mathcal{L}$.

We study definability of second order quantifiers in different contexts. In Chapter III we study which second order generalized quantifiers can be defined in a “first order way”. In fact, we concentrate on studying definability of second order generalized quantifiers in extensions of FO by Lindström quantifiers. It turns out that there are simple structural properties of second order predicates which can be used to prove non-definability results. We also show that $\mathcal{L}(\mathcal{Q}) \equiv \mathcal{L}$ does not imply definability of \mathcal{Q} in the logic \mathcal{L} contrary to the case with Lindström quantifiers.

In Chapter IV we study definability in monadic second order logic, MSO. We prove some non-definability results using the fact that a quantifier \mathcal{Q} for which $\text{MSO}(\mathcal{Q}) > \text{MSO}$ cannot be definable in MSO. We also show that finding a simple characterization of definable quantifiers in MSO is hard since it would yield a characterization of classes of graphs definable in MSO.

In Chapter V we take a broader perspective and study definability of second order quantifiers in terms of their types. The main result of Chapter V is the analogue of the Hierarchy Theorem in [8]. It says that for any second order type t there is a quantifier of type t which cannot be defined using any second order quantifiers of types lower than t . We also produce a concrete example of a quantifier of type ((2)) which is not definable in terms of quantifiers of lower types, namely the binary second order existential quantifier.

I. PRELIMINARIES

In this chapter we explain the notation used in this thesis. We also recall the basic concepts and some known results about Lindström quantifiers.

2. NOTATION

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 by their universes.

The cardinality of a set X is denoted by $|X|$. The image of a relation R under a function f is $f[R] = \{(f(a_1), \dots, f(a_n)) \mid (a_1, \dots, a_n) \in R\}$. With the cardinality of a structure \mathbb{M} we mean the cardinality of the domain M .

The class of all τ -structures is denoted by $\text{Str}(\tau)$, and $\text{Str}(\tau, n)$ is the restriction of $\text{Str}(\tau)$ to structures having cardinality n . For a logic \mathcal{L} , the set of τ -formulas of \mathcal{L} is denoted by $\mathcal{L}[\tau]$. If φ is a τ -sentence then the class of τ -models of φ is denoted by $\text{Mod}(\varphi)$. For a class K of τ -structures, the number of isomorphism types of structures in K with universe $\{1, \dots, n\}$ is denoted by $I_n(K)$.

For functions $f, g: \omega \rightarrow \omega$ on natural numbers, we write $f(n) \sim g(n)$ if

$$\lim_{n \rightarrow \infty} f(n)/g(n) = 1.$$

Throughout the text, second order generalized quantifiers are denoted by \mathcal{Q} and Lindström quantifiers are denoted by Q . The set of natural numbers is denoted by \mathbb{N} or ω .

3. LINDSTRÖM QUANTIFIERS

Lindström defined the general notion of first order generalized quantifier in [11].

3.1. Definition. Let $s = (l_1, \dots, l_r)$ be a tuple of positive integers. A *Lindström quantifier* of type s is a class Q of structures of vocabulary $\tau_s = \{P_1, \dots, P_r\}$ such that P_i is l_i -ary for $1 \leq i \leq r$, and Q is closed under isomorphisms.

The *arity* of a quantifier Q is $\text{ar}(Q) = \max\{\text{ar}(P_1), \dots, \text{ar}(P_r)\}$. We say that Q is n -ary if $\text{ar}(Q) \leq n$. The type of the quantifier Q is called *simple* if $r = 1$, i.e., if Q applies to one formula.

3.2. Example. Let us look at some examples of Lindström quantifiers.

$$\begin{aligned} \forall &= \{(M, P) \mid P = M\} \\ \exists &= \{(M, P) \mid P \subseteq M \text{ and } P \neq \emptyset\} \\ Q_{\text{even}} &= \{(M, P) \mid P \subseteq M \text{ and } |P| \text{ is even}\} \\ \text{Most} &= \{(M, P, S) \mid P, S \subseteq M \text{ and } |P \cap S| > |P \setminus S|\} \\ \text{Some} &= \{(M, P, S) \mid P, S \subseteq M \text{ and } P \cap S \neq \emptyset\} \\ \text{I} &= \{(M, P, S) \mid P, S \subseteq M \text{ and } |P| = |S|\} \end{aligned}$$

The first example is the first order universal quantifier. The quantifiers Most and Some are examples of quantifiers used in natural language semantics. They aim to capture the truth conditions of sentences of the form "Most A's are B's" and "Some A is B". The quantifier I is the so-called Härtig quantifier.

Quantifiers can be also defined over a fixed set M .

3.3. Definition. A quantifier Q of type s over M is a subset of $\mathcal{P}(M^{l_1}) \times \dots \times \mathcal{P}(M^{l_r})$ such that Q is *permutation invariant*:

$$(A_1, \dots, A_r) \in Q \Leftrightarrow (f[A_1], \dots, f[A_r]) \in Q$$

for all $(A_1, \dots, A_r) \in \mathcal{P}(M^{l_1}) \times \dots \times \mathcal{P}(M^{l_r})$ and $f \in S_M$, where S_M is the set of permutations of M .

Note that if Q is a Lindström quantifier of type s and M is a set, then

$$\{(A_1, \dots, A_r) \mid (M, A_1, \dots, A_r) \in Q\}$$

is a quantifier of type s over M .

3.4. Definition. The extension $\text{FO}(Q)$ of first order logic by a quantifier Q is defined as follows:

- The formula formation rules of FO are extended by the rule:
if for $1 \leq i \leq r$, $\varphi_i(\bar{x}_i)$ is a formula and \bar{x}_i is an l_i -tuple of pairwise distinct variables then $Q\bar{x}_1, \dots, \bar{x}_r (\varphi_1(\bar{x}_1), \dots, \varphi_r(\bar{x}_r))$ is a formula.
- The satisfaction relation of FO is extended by the rule:

$$\mathbb{M} \models Q\bar{x}_1, \dots, \bar{x}_r (\varphi_1(\bar{x}_1), \dots, \varphi_r(\bar{x}_r)) \text{ iff } (M, \varphi_1^{\mathbb{M}}, \dots, \varphi_r^{\mathbb{M}}) \in Q,$$

$$\text{where } \varphi_i^{\mathbb{M}} = \{\bar{a} \in M^{l_i} \mid \mathbb{M} \models \varphi_i(\bar{a})\}.$$

First order logic can be extended in the same way by a collection \mathcal{B} of Lindström quantifiers. The extension of FO by the quantifiers in \mathcal{B} is denoted by $\text{FO}(\mathcal{B})$.

4. DEFINABILITY OF LINDSTRÖM QUANTIFIERS

4.1. Definition. Let \mathcal{L} and \mathcal{L}' be logics. The logic \mathcal{L}' is *at least as strong as* the logic \mathcal{L} ($\mathcal{L} \leq \mathcal{L}'$) if for every sentence $\varphi \in \mathcal{L}$ over any vocabulary there exists a sentence $\psi \in \mathcal{L}'$ over the same vocabulary such that

$$\models \varphi \leftrightarrow \psi.$$

The logics \mathcal{L} and \mathcal{L}' are *equivalent* ($\mathcal{L} \equiv \mathcal{L}'$) if $\mathcal{L} \leq \mathcal{L}'$ and $\mathcal{L}' \leq \mathcal{L}$.

4.2. Definition. Let Q be a Lindström quantifier of type s and \mathcal{B} a collection of Lindström quantifiers. We say that the quantifier Q is *definable* in terms of the quantifiers in \mathcal{B} if there is a sentence $\varphi \in \text{FO}(\mathcal{B})$ of vocabulary τ_s such that for any τ_s -structure \mathbb{M} :

$$\mathbb{M} \models \varphi \Leftrightarrow \mathbb{M} \in Q.$$

4.3. Proposition. *Let Q be a Lindström quantifier and \mathcal{B} a collection of Lindström quantifiers. If the quantifier Q is definable in terms of the quantifiers in \mathcal{B} , then*

$$\text{FO}(Q, \mathcal{B}) \equiv \text{FO}(\mathcal{B}).$$

The proof of Proposition 4.3 uses recursively the fact that if φ is the formula which defines Q and $\psi_1(\bar{x}_1), \dots, \psi_r(\bar{x}_r)$ are formulas of $\text{FO}(\mathcal{B})$, then

$$\models Q\bar{x}_1, \dots, \bar{x}_r (\psi_1(\bar{x}_1), \dots, \psi_r(\bar{x}_r)) \leftrightarrow \varphi(P_1/\psi_1, \dots, P_r/\psi_r),$$

where the formula on the right is obtained by substituting every occurrence of $P_i(\bar{x}_i)$ in φ by $\psi_i(\bar{x}_i)$. Since $Q = \text{Mod}(\varphi)$, where $\varphi = Q\bar{x}_1, \dots, \bar{x}_r (P_1(\bar{x}_1), \dots, P_r(\bar{x}_r))$, the converse of Proposition 4.3 is also true. We shall prove that the version of Proposition 4.3, where Q is replaced with a second order quantifier \mathcal{Q} , holds. On the other hand, the converse of Proposition 4.3 turns out to be false in the second order case.

The quantifiers \exists, \forall , and *Some* in Example 3.2 are clearly definable in FO. Fact 4.4 lists some well-known examples of Lindström quantifiers which are not definable in FO.

4.4. Fact. *Let $S \subseteq \mathbb{N}$ be infinite and co-infinite. The following Lindström quantifiers are not definable in FO:*

- I
- Most
- $\text{Log} = \{(M, P) \mid P \subseteq M \text{ and } |P| \geq \log(|M|)\}$
- $Q_S = \{(M, P) \mid P \subseteq M \text{ and } |P| \in S\}$

5. SECOND ORDER GENERALIZED QUANTIFIERS

The notion of second order generalized quantifier was introduced in [1]. Let $t = (s_1, \dots, s_w)$, where $s_i = (l_1^i, \dots, l_{r_i}^i)$ 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, \dots, P_w) , where $P_i \subseteq \mathcal{P}(M^{l_1^i}) \times \dots \times \mathcal{P}(M^{l_{r_i}^i})$. In the following the notations $\text{Str}(t)$ and $\text{Str}(t, n)$ are used analogously to the first order case.

5.1. Definition. 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: if $(M, P_1, \dots, P_w) \in \mathcal{Q}$ and $f: M \rightarrow N$ is a bijection such that

$$S_i = \{(f[A_1], \dots, f[A_{r_i}]) \mid (A_1, \dots, A_{r_i}) \in P_i\}$$

(abbreviated $f[[P_i]] = S_i$) for $1 \leq i \leq w$ then $(N, S_1, \dots, S_w) \in \mathcal{Q}$.

The *arity* of a second order generalized quantifier \mathcal{Q} is $\text{ar}(\mathcal{Q}) = \max\{r_1, \dots, r_w\}$ and the *variable arity* of \mathcal{Q} is $\max\{l_j^i \mid 1 \leq j \leq r_i \text{ and } 1 \leq i \leq w\}$. The type of a quantifier \mathcal{Q} is called *simple* if $w = 1$ and \mathcal{Q} is called *monadic* if its variable arity is 1.

5.2. Example. Let us look at some examples of second order generalized quantifiers. Suppose $S \subseteq \mathbb{N}$ and $k \in \mathbb{N}$.

$$\begin{aligned}
\exists_k^2 &= \{(M, P) \mid P \subseteq \mathcal{P}(M^k) \text{ and } P \neq \emptyset\} \\
\text{Even} &= \{(M, P) \mid P \subseteq \mathcal{P}(M) \text{ and } |P| \text{ even}\} \\
\text{Even}' &= \{(M, P) \mid P \subseteq \mathcal{P}(M) \text{ and } \forall X \in P (|X| \text{ even})\} \\
\text{Three} &= \{(M, P) \mid P \subseteq \mathcal{P}(M) \text{ and } |P| = 3\} \\
\text{Most}^2 &= \{(M, P) \mid P \subseteq \mathcal{P}(M) \text{ and } |P| \geq 2^{|M|-1}\} \\
\mathcal{Q}_S &= \{(M, P) \mid P \subseteq \mathcal{P}(M) \text{ and } |P| \in S\} \\
\text{TC}_k^2 &= \{(M, R_1, R_2) \mid R_i \subseteq \mathcal{P}(M^k) \times \mathcal{P}(M^k) \text{ and } R_2 \subseteq \text{TC}(R_1)\}
\end{aligned}$$

The first example is the familiar k -ary second order existential quantifier. The type of \exists_k^2 is $((k))$. The quantifier Even says that a formula holds for an even number of subsets of the universe. On the other hand, the quantifier Even' says that all the subsets satisfying a formula have an even number of elements. The quantifier TC_k^2 is the second order version of the transitive closure operator. It consists of the structures (M, R_1, R_2) , where R_2 is contained in the transitive closure $\text{TC}(R_1)$ of R_1 . The type of TC_k^2 is $((k, k), (k, k))$. The type of the other examples is $((1))$.

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

- Second order variables are introduced to FO.
- The formula formation rules of FO are extended by the rule:
if for $1 \leq i \leq w$, $\varphi_i(\bar{X}_i)$ is a formula and $\bar{X}_i = (X_{1,i}, \dots, X_{r_i,i})$ is a tuple of pairwise distinct predicate variables such that $\text{ar}(X_{j,i}) = l_j^i$ for $1 \leq j \leq r_i$, then

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

is a formula.

- Satisfaction relation of FO is extended by the rule:

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

$$\text{where } \varphi_i^{\mathbb{M}} = \{\bar{R} \in \mathcal{P}(M^{l_1^i}) \times \dots \times \mathcal{P}(M^{l_{r_i}^i}) \mid \mathbb{M} \models \varphi_i(\bar{R})\}.$$

5.4. Example. We show that Partial Fixed-Point Logic $\text{FO}(\text{PFP})$ can be captured by the second order transitive closure quantifiers TC_k^2 . So let φ be a formula of the logic $\text{FO}(\text{PFP})$. Then φ is equivalent to a formula of the form $\forall \bar{u} [\text{PFP}_{X, \bar{x}} \psi] \bar{u}$ where ψ is first order (cf. [4]). Let π be the formula

$$\text{TC}_k^2 XY, UV (\forall \bar{x} (Y(\bar{x}) \leftrightarrow \psi(X, \bar{x})), \forall \bar{y} \neg U(\bar{y}) \wedge \forall \bar{u} V(\bar{u})),$$

where $k = |\bar{x}|$. Then $\models \forall \bar{u} [\text{PFP}_{X, \bar{x}} \psi] \bar{u} \leftrightarrow \pi$ and thus $\models \varphi \leftrightarrow \pi$.

II. A NOTION OF DEFINABILITY FOR SECOND ORDER QUANTIFIERS

In this chapter we formulate the notion of definability studied in this thesis. Before we can give the exact definition we need to define some auxiliary logics.

6. EXTENDING FIRST ORDER LOGIC BY SECOND ORDER PREDICATES

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

- The syntax of the logic $\text{FO}(\mathcal{G}_1, \dots, \mathcal{G}_w)$ is obtained by extending the syntax of FO by the quantifier rules corresponding to $\mathcal{G}_1, \dots, \mathcal{G}_w$.
- The models \mathcal{M} of the logic $\text{FO}(\mathcal{G}_1, \dots, \mathcal{G}_w)$ are of the form

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

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

- The satisfaction relation of FO is extended by the following rule for $1 \leq i \leq w$:

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

Note that if G_i happens to be permutation invariant (cf. Definition 3.3) then G_i is a quantifier of type s_i over M .

Now, for $\tau = \emptyset$, all τ -models of $\text{FO}(\mathcal{G}_1, \dots, \mathcal{G}_w)$ are second order structures of type t . By Proposition 6.2, isomorphic structures are elementarily equivalent. Thus, for any τ -sentence $\varphi \in \text{FO}(\mathcal{G}_1, \dots, \mathcal{G}_w)$ the class $\text{Mod}(\varphi)$ is a second order quantifier of type t .

6.2. Proposition. Let $(\mathbb{M}, G_1, \dots, G_w)$ and $(\mathbb{N}, G'_1, \dots, G'_w)$ be models and $f: \mathbb{M} \cong \mathbb{N}$ an isomorphism such that $f[[G_i]] = G'_i$ for $1 \leq i \leq w$. Then for all $\varphi(\bar{x}) \in \text{FO}(\mathcal{G}_1, \dots, \mathcal{G}_w)$ and $\bar{a} = a_1 \dots a_n \in M$,

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

Proof. Induction on φ . □

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

6.3. Definition. Let \mathcal{B} be a collection of first and second order quantifiers and let \mathcal{Q} be a second order quantifier of type $t = (s_1, \dots, s_w)$. The quantifier \mathcal{Q} is *definable* in $\text{FO}(\mathcal{B})$ if there is a sentence $\varphi \in \text{FO}(\mathcal{B}, \mathcal{G}_1, \dots, \mathcal{G}_w)$ of vocabulary $\tau = \emptyset$ such that for all second order structures (M, G_1, \dots, G_w) of type t ,

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

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

6.4. Example. Suppose, to simplify notation, that \mathcal{Q} is a quantifier of type $t = ((1), (1))$. Then for any t -structure (M, G_1, G_2) ,

$$(M, G_1, G_2) \models \mathcal{Q}X, Y (\mathcal{G}_1 x X(x), \mathcal{G}_2 y Y(y)) \Leftrightarrow (M, G_1, G_2) \in \mathcal{Q}.$$

Next we show that the notion of definability formulated in Definition 6.3 does have the desired property.

6.5. Theorem. *If \mathcal{Q} is definable in the logic $\text{FO}(\mathcal{B})$ then $\text{FO}(\mathcal{B}) \equiv \text{FO}(\mathcal{B}, \mathcal{Q})$.*

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

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

We may assume that $\theta_1(\overline{X}_1), \dots, \theta_w(\overline{X}_w)$ do not contain other free variables than the ones appearing in $\overline{X}_1, \dots, \overline{X}_w$. If $\text{FO}(\mathcal{B})$ does not have second order variables, we may treat the variables appearing in $\overline{X}_1, \dots, \overline{X}_w$ as extra predicate symbols. Define a mapping $f: \text{FO}(\mathcal{B}, \mathcal{G}_1, \dots, \mathcal{G}_w)[\emptyset] \rightarrow \text{FO}(\mathcal{B})[\sigma]$ as follows (where the only non-trivial clause is the last one):

$$\begin{aligned} f(x = y) &= (x = y) \\ f(X(\overline{y})) &= X(\overline{y}) \\ f(\neg\varphi) &= \neg f(\varphi) \\ f(\varphi \wedge \psi) &= f(\varphi) \wedge f(\psi) \\ f(\exists x\varphi) &= \exists x f(\varphi) \\ f(Q\overline{x}_1, \dots, \overline{x}_r(\varphi_1, \dots, \varphi_r)) &= Q\overline{x}_1, \dots, \overline{x}_r(f(\varphi_1), \dots, f(\varphi_r)), \text{ for } Q \in \mathcal{B} \\ f(Q'\overline{X}_1, \dots, \overline{X}_w(\varphi_1, \dots, \varphi_w)) &= Q'\overline{X}_1, \dots, \overline{X}_w(f(\varphi_1), \dots, f(\varphi_w)), \text{ for } Q' \in \mathcal{B} \\ f(\mathcal{G}_i\overline{x}_1, \dots, \overline{x}_{r_i}(\varphi_1, \dots, \varphi_{r_i})) &= \theta_i(X_1/f(\varphi_1), \dots, X_{r_i}/f(\varphi_{r_i})) \end{aligned}$$

Our intention is to show that

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

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

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

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

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

where $A_i = \{\overline{d} \mid (\mathbb{M}, G_1, \dots, G_w) \models \psi_i(\overline{d}, \overline{b}, \overline{R})\}$. By the definition of G_i , we have

$$(A_1, \dots, A_{r_i}) \in G_i \Leftrightarrow \mathbb{M} \models \theta_i(A_1, \dots, A_{r_i}).$$

By the induction hypothesis $A_i = \{\overline{d} \mid \mathbb{M} \models f(\psi_i)(\overline{d}, \overline{b}, \overline{R})\}$, hence

$$\mathbb{M} \models \theta_i(A_1, \dots, A_{r_i}) \Leftrightarrow \mathbb{M} \models \theta_i(X_1/f(\psi_1), \dots, X_{r_i}/f(\psi_{r_i}))(\overline{b}, \overline{R}).$$

The claim now follows by the chain of equivalences. □

We claim now that

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

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

$$\begin{aligned} \mathbb{M} \models \mathcal{Q}\overline{X}_1, \dots, \overline{X}_w(\theta_1, \dots, \theta_w) &\Leftrightarrow (\mathbb{M}, G_1, \dots, G_w) \models \varphi_{\mathcal{Q}} \\ &\Leftrightarrow \mathbb{M} \models f(\varphi_{\mathcal{Q}}). \end{aligned}$$

□

III. DEFINABILITY IN TERMS OF LINDSTRÖM QUANTIFIERS

In this chapter we study definability of second order quantifiers in the logic $\text{FO}(\mathfrak{C})$, where \mathfrak{C} is the collection of all Lindström quantifiers of all types. The results in this chapter will appear in a modified form in [10].

7. PRELIMINARIES

7.1. Definition. Let \mathfrak{C} be the collection of all Lindström quantifiers of all types. Denote the logic $\text{FO}(\mathfrak{C})$ by \mathcal{L}^* . Also, $\text{FO}(\mathfrak{C}, \mathcal{G}_1, \dots, \mathcal{G}_w)$ is denoted by $\mathcal{L}^*(\mathcal{G}_1, \dots, \mathcal{G}_w)$.

Every formula of the logic $\mathcal{L}^*(\mathcal{G}_1, \dots, \mathcal{G}_w)$ contains occurrences of finitely many Lindström quantifiers. Therefore, Theorem 6.5 implies the following.

7.2. Proposition. *Let \mathcal{Q} be definable in the logic \mathcal{L}^* with a sentence φ . Then*

$$\text{FO}(\mathcal{Q}, Q_1, \dots, Q_n) \equiv \text{FO}(Q_1, \dots, Q_n),$$

where Q_1, \dots, Q_n are those Lindström quantifiers which appear in φ .

We shall first restrict attention to definability of second order generalized quantifiers of type ((1)) in the logic \mathcal{L}^* . In Section 11 we show how to treat quantifiers of arbitrary types.

8. SOME DEFINABLE QUANTIFIERS

In this section we formulate a simple property of quantifiers of type ((1)) which implies definability in the logic \mathcal{L}^* . Let us begin with some examples.

8.1. Example.

- $\psi_1 = \forall x \mathcal{G}z (z = x)$
- $\psi_2 = \exists x_1 \dots \exists x_n (\bigwedge_{i \neq j} (x_i \neq x_j) \wedge \mathcal{G}z (\bigvee_{i=1}^n z = x_i))$
- $\psi_3 = Q_{\text{even}} x (\mathcal{G}z (z = x))$

The formula ψ_1 (ψ_3 , respectively) is true in (M, G) iff G contains all (an even number of, respectively) singletons. The formula ψ_2 expresses that G contains at least one set of size n .

8.2. Example. Let $M = \{a_1, \dots, a_n\}$ be a set and $G \subseteq \mathcal{P}(M)$. Choose variables x_1, \dots, x_n and, for $X = \{a_{i_1}, \dots, a_{i_m}\} \subseteq M$, define

$$\psi_X = \mathcal{G}z (\bigvee_{j=1}^m z = x_{i_j}).$$

Now, the formula $\psi_4 = \exists x_1 \dots \exists x_n (\bigwedge_{i \neq j} (x_i \neq x_j) \wedge \forall y (\bigvee_{i=1}^n (y = x_i)) \wedge (\bigwedge_{X \in G} \psi_X) \wedge (\bigwedge_{X \notin G} \neg \psi_X))$ determines the model (M, G) up to isomorphism.

Example sentences ψ_1, ψ_2, ψ_3 , and ψ_4 suggest that we need at least $n+1$ variables to define (cf. Definition 9.1) sets of size n . This is not actually true for all models (M, G) but holds under some extra assumptions on G . This fact motivates the following definition.

8.3. Definition. Let (M, G) be a structure and $n \in \mathbb{N}$. The set $\{X \in G \mid |X| \leq n \text{ or } |M \setminus X| \leq n\}$ is denoted by $G(n)$.

8.4. Theorem. Let \mathcal{Q} be a quantifier of type $((1))$. Suppose there is $n \in \mathbb{N}$ such that for any set M and $G \subseteq \mathcal{P}(M)$ we have $(M, G) \in \mathcal{Q}$ if and only if $(M, G(n)) \in \mathcal{Q}$. Then \mathcal{Q} is definable in \mathcal{L}^* .

Proof. We shall construct formulas $\psi_1, \psi_2, \psi_3, \psi_4$ and a Lindström quantifier Q of type $(n, n, 1, 1)$ such that the sentence $\varphi \in \mathcal{L}^*(\mathcal{G})$

$$\varphi = Q\bar{x}_1, \bar{x}_2, x, y (\psi_1(\bar{x}_1), \psi_2(\bar{x}_2), \psi_3(x), \psi_4(y))$$

defines \mathcal{Q} . The quantifier Q will be defined so that

$$(M, G) \models Q\bar{x}_1, \bar{x}_2, x, y (\psi_1, \psi_2, \psi_3, \psi_4) \Leftrightarrow (M, C) \in \mathcal{Q},$$

where C is a certain collection of subsets of M constructed from the relations $\psi_i^{(M, G)}$. The idea is to choose the formulas ψ_i so that $C = G(n)$.

With this in mind, we let $\psi_1(\bar{x}_1)$ be the formula

$$\mathcal{G}z \left(\bigvee_{j=1}^n z = x_j \right)$$

and $\psi_2(\bar{x}_2)$ be the formula

$$\mathcal{G}z \left(\bigwedge_{j=1}^n z \neq x_j \right).$$

Finally, $\psi_3(x)$ is the formula

$$\neg \mathcal{G}z (z \neq z)$$

and $\psi_4(y)$ is the formula

$$\mathcal{G}z (z = z).$$

Let Q be defined as follows:

$$(M, R_1, \dots, R_4) \in Q \Leftrightarrow (M, \bigcup_{i=1}^4 R'_i) \in \mathcal{Q},$$

where R'_1 is

$$\{\{a_1, \dots, a_n\} \mid (a_1, \dots, a_n) \in R_1\}$$

and R'_2 is

$$\{M \setminus \{a_1, \dots, a_n\} \mid (a_1, \dots, a_n) \in R_2\}.$$

The sets R'_3 and R'_4 are defined as follows:

$$R'_3 = \{R_3\} \setminus \{M\} \text{ and } R'_4 = \{R_4\} \setminus \{\emptyset\}.$$

2. Claim. The class \mathcal{Q} is closed under isomorphisms.

Proof of Claim. Let $f: (M, R_1, \dots, R_4) \cong (N, S_1, \dots, S_4)$. Then $f[[R'_i]] = S'_i$ for $1 \leq i \leq 4$. Since the quantifier \mathcal{Q} is closed under isomorphisms, we have

$$(M, \bigcup_{i=1}^4 R'_i) \in \mathcal{Q} \Leftrightarrow (N, \bigcup_{i=1}^4 S'_i) \in \mathcal{Q},$$

and thus

$$(M, R_1, \dots, R_4) \in \mathcal{Q} \Leftrightarrow (N, S_1, \dots, S_4) \in \mathcal{Q}.$$

□

By Claim 2, the class \mathcal{Q} is a Lindström quantifier. Hence, the logic \mathcal{L}^* is closed under \mathcal{Q} and $\varphi \in \mathcal{L}^*(\mathcal{G})$. It is now easy to verify that for any structure (M, G)

$$(M, G) \models \varphi \Leftrightarrow (M, G(n)) \in \mathcal{Q}.$$

Since the quantifier \mathcal{Q} satisfies

$$(M, G(n)) \in \mathcal{Q} \Leftrightarrow (M, G) \in \mathcal{Q}$$

for all (M, G) , the sentence φ defines \mathcal{Q} .

□

9. PROVING NON-DEFINABILITY

In the previous section we proved that using suitable Lindström quantifiers we can define any second order quantifier \mathcal{Q} such that the membership of a structure (M, G) in \mathcal{Q} is determined by checking only sets $X \in G$ for which $|X| < n$ or $|M \setminus X| < n$ for some fixed $n \in \mathbb{N}$. However, this property does not characterize all quantifiers which are definable in the logic \mathcal{L}^* (cf. Example 9.5). Therefore, we concentrate now on developing tools for proving non-definability. One simple property of $\mathcal{L}^*(\mathcal{G})$ which is instrumental in proving non-definability is that every formula of the logic $\mathcal{L}^*(\mathcal{G})$ has only a finite number of variables.

Let \mathcal{Q} be a quantifier of type ((1)). Proving the non-definability of \mathcal{Q} in the logic \mathcal{L}^* amounts to showing that there is no $\varphi \in \mathcal{L}^*(\mathcal{G})$ over vocabulary $\tau = \emptyset$ such that for all (M, G) ,

$$(M, G) \models \varphi \Leftrightarrow (M, G) \in \mathcal{Q}.$$

For this it suffices to construct structures (M_k, G_k) and (M_k, G'_k) , for $k \in \mathbb{N}$, such that $(M_k, G_k) \in \mathcal{Q}$, $(M_k, G'_k) \notin \mathcal{Q}$, and $(M_k, G_k) \equiv^k (M_k, G'_k)$, where \equiv^k means that the structures satisfy the same sentences of $\mathcal{L}^*(\mathcal{G})[\tau]$ containing at most k variables. We shall formulate a simple structural property of collections $G, G' \subseteq \mathcal{P}(M)$ implying that $(M, G) \equiv^k (M, G')$. Let us begin with some definitions.

9.1. Definition. Let σ be a vocabulary, \mathbb{M} a σ -structure, $G \subseteq \mathcal{P}(M)$, and let $a_1, \dots, a_n \in M$. A relation $X \subseteq M^l$ is *definable with parameters* a_1, \dots, a_n if there is a formula $\varphi(\bar{y}, x_1, \dots, x_n) \in \mathcal{L}^*(\mathcal{G})[\sigma]$ such that

$$X = \{\bar{b} \in M^l \mid (\mathbb{M}, G) \models \varphi(\bar{b}, a_1, \dots, a_n)\}.$$

A closely related notion to a definable relation is the notion of invariant relation.

9.2. Definition. Let σ be a vocabulary, \mathbb{M} a σ -structure, $G \subseteq \mathcal{P}(M)$, and let $a_1, \dots, a_n \in M$. A relation $X \subseteq M^l$ is called (a_1, \dots, a_n) -invariant if $f[X] = X$ for all $f: (\mathbb{M}, G) \cong (\mathbb{M}, G)$ such that $f(a_i) = a_i$ for $1 \leq i \leq n$.

Proposition 6.2 implies that all (a_1, \dots, a_n) -definable relations are (a_1, \dots, a_n) -invariant. Note, however, that there might not be any non-trivial permutations fixing the collection G .

9.3. Example. Let M be a set and $<$ a linear-ordering of M . Define G as the set of all initial segments of $<$. Then, $f[[G]] \neq G$ for all permutations f of M such that $f \neq \text{id}_M$.

On the other hand, if the quantifier G is permutation invariant then G is fixed by all permutations of M . In this special case invariant sets over $\tau = \emptyset$ can be easily characterized.

9.4. Lemma. *Let (M, G) be a structure such that G is permutation invariant. Then a set $X \subseteq M$ is (a_1, \dots, a_n) -invariant if and only if $X \subseteq \{a_1, \dots, a_n\}$ or $M \setminus X \subseteq \{a_1, \dots, a_n\}$.*

Lemma 9.4 is very useful since it tells us what (a_1, \dots, a_n) -definable sets look like in a model (M, G) , where G is permutation invariant. Example 9.3 shows that Lemma 9.4 fails if G is not required to be permutation invariant. By Example 9.5, Lemma 9.4 also fails for the definable sets.

9.5. Example. Let $\varphi = \mathcal{G}z(\mathcal{G}x(x = z))$. The sentence φ is true in a model (M, G) if and only if $A \in G$ where $A = \{a \in M \mid \{a\} \in G\}$. In particular, the formula $\mathcal{G}x(x = z)$ defines the set A without parameters.

Next, we formulate the first structural property of collections $G, G' \subseteq \mathcal{P}(M)$ implying that $(M, G) \equiv^k (M, G')$.

9.6. Lemma. *Let (M, G) and (M, G') be structures such that G and G' are permutation invariant and $G(k) = G'(k)$. Then*

$$(M, G) \equiv^{k+1} (M, G').$$

Proof. We prove using induction on $\varphi(x_1, \dots, x_n) \in \mathcal{L}^*(\mathcal{G})$ having at most $k + 1$ variables that

$$(M, G) \models \varphi(a_1, \dots, a_n) \Leftrightarrow (M, G') \models \varphi(a_1, \dots, a_n),$$

for all $a_1, \dots, a_n \in M$. We consider the case $\varphi = \mathcal{G}x\psi(x, \bar{y})$ only since the cases $\varphi = (x = y)$, $\varphi = \neg\psi$, $\varphi = \psi \wedge \phi$, and $\varphi = Q\bar{x}_1, \dots, \bar{x}_r(\psi_1, \dots, \psi_r)$ are trivial. Assume $(M, G) \models \varphi(a_1, \dots, a_n)$. Then $X \in G$, where $X = \{c \in M \mid (M, G) \models \psi(c, a_1, \dots, a_n)\}$. The formula φ contains at most $k + 1$ variables, hence $|\{a_1, \dots, a_n\}| \leq k$. By Lemma 9.4, we have either $X \subseteq \{a_1, \dots, a_n\}$ or $M \setminus X \subseteq \{a_1, \dots, a_n\}$. Since $G(k) = G'(k)$, we have that $X \in G'$. By the induction hypothesis, $X = \{c \in M \mid (M, G') \models \psi(c, a_1, \dots, a_n)\}$ and thus $(M, G') \models \varphi(a_1, \dots, a_n)$. The converse is similar. \square

Lemma 9.6 can be used to show many quantifiers non-definable in the logic \mathcal{L}^* .

9.7. Theorem. *The following quantifiers are not definable in \mathcal{L}^* :*

- Most^2
- $\exists^2 = \exists_1^2$

Proof. We prove the claim for the quantifier \exists^2 . The non-definability of Most^2 can be shown analogously. For $k \in \mathbb{N} \setminus \{0, 1\}$, we set $M_k = \{1, \dots, 3k\}$, $G_k = \emptyset$ and $G'_k = \{A \subseteq \{1, \dots, 3k\} \mid |A| > k \text{ and } |M \setminus A| > k\}$. Suppose that φ defines \exists^2 . Then

$$(9.8) \quad (M_k, G_k) \not\models \varphi \text{ and } (M_k, G'_k) \models \varphi$$

for $k \in \mathbb{N} \setminus \{0, 1\}$. Let $r \in \mathbb{N}$ be the number of variables in φ . By Lemma 9.6, we have $(M_{r-1}, G_{r-1}) \equiv^r (M_{r-1}, G'_{r-1})$, and thus $(M_{r-1}, G_{r-1}) \models \varphi$ iff $(M_{r-1}, G'_{r-1}) \models \varphi$ contradicting (9.8). \square

There are quantifiers which cannot be shown to be non-definable using Lemma 9.6. The reason is that we have to be able to choose structures with permutation invariant quantifiers in order to apply it. For example, for the quantifier Three such structures of cardinality $n > 3$ do not exist. Our next goal is to modify Lemma 9.6 by replacing the assumption of permutation invariance with a milder condition.

Let us first return to the formula $\psi = \mathcal{G}x(x = z)$ introduced in Example 9.5. In a structure (M, G) the formula ψ defines the set $A = \{a \in M \mid \{a\} \in G\}$. If G happens to be permutation invariant then $A = \emptyset$ or $A = M$. Actually, it is not necessary that G is closed under permutations “globally” for A to be either \emptyset or M . It is enough that G contains all singletons or no singletons.

9.9. Lemma. *Let (M, G) and (M, G') be structures such that $G(k)$ and $G'(k)$ are permutation invariant and $G(k) = G'(k)$. Then*

$$(M, G) \equiv^{k+1} (M, G').$$

Proof. We prove using induction on $\varphi(\bar{x}) \in \mathcal{L}^*(\mathcal{G})$ having at most $k + 1$ variables that for all $f \in S_M$ and $\bar{a} = a_1 \dots a_n \in M$,

$$(M, G) \models \varphi(\bar{a}) \Leftrightarrow (M, G') \models \varphi(f(\bar{a})).$$

We consider only the non-trivial case $\varphi = \mathcal{G}x\psi(x, \bar{y})$. Assume $\bar{a} \in M$, $f \in S_M$, and $(M, G) \models \varphi(\bar{a})$. Then $X = \{c \in M \mid (M, G) \models \psi(c, \bar{a})\} \in G$. The induction assumption applied to id_M implies that $X = \{c \in M \mid (M, G') \models \psi(c, \bar{a})\}$. Therefore, again by the induction assumption, $g[X] = X$ for all $g \in S_M$ such that $g(a_i) = a_i$ for $1 \leq i \leq n$. So (cf. Lemma 9.4) we have either $X \subseteq \{a_1, \dots, a_n\}$ or $(M \setminus X) \subseteq \{a_1, \dots, a_n\}$. Also, by the induction assumption $f[X] = X'$ where $X' = \{c \in M \mid (M, G') \models \psi(c, f(\bar{a}))\}$. Now, since $X \in G(k)$ and $G(k)$ is permutation invariant we have that $X' \in G(k) = G'(k)$, i.e., $(M, G') \models \varphi(f(\bar{a}))$. Assuming $(M, G') \models \varphi(f(\bar{a}))$, we can show $(M, G) \models \varphi(\bar{a})$ completely analogously. \square

Lemma 9.9 now allows us to prove the non-definability of the quantifier Three in the logic \mathcal{L}^* .

9.10. Theorem. *The following quantifiers are not definable in \mathcal{L}^* :*

- Three
- \mathcal{Q}_S , where $S \subsetneq \mathbb{N}$ and $S \neq \emptyset$.

Proof. Similar to the proof of Theorem 9.7. Let us consider the quantifier Three. For $k \geq 3$, we set $M_k = \{1, \dots, 3k\}$, $G_k = \emptyset$ and $G'_k = \{A_1, A_2, A_3\}$, where $A_i = \{1, \dots, k+i\}$ for $1 \leq i \leq 3$. By the same arguments as in the proof of Theorem 9.7 and Lemma 9.9, the quantifier Three is not definable in the logic \mathcal{L}^* .

In the general case we first choose $n_1 \in S$ and $n_2 \notin S$ and then construct M_k, G_k with $|G_k| = n_2$, and G'_k with $|G'_k| = n_1$ similarly as for Three. \square

However, we conjecture that there are non-definable quantifiers which cannot be shown to be non-definable in the logic \mathcal{L}^* using Lemma 9.9. Let k be a positive integer. We conjecture that the quantifier \mathcal{Q}_u is not definable in the logic \mathcal{L}^* . Note that Lemma 9.9 cannot be used to show this.

$$\mathcal{Q}_u = \{(M, G) \mid G \text{ contains exactly } k \text{ sets of size } n \text{ for } 1 \leq n \leq |M| - 1\}.$$

10. LOOKING FOR A COMPLETE CHARACTERIZATION OF DEFINABILITY

Let us return to our example sentence $\varphi = \mathcal{G}z(\mathcal{G}x(x = z))$ introduced in Example 9.5. The second order quantifier defined by the sentence φ is a quantifier which is definable in \mathcal{L}^* but does not satisfy the assumption of Theorem 8.4. The reason is that φ contains nesting of the quantifier \mathcal{G} . We shall next show that the characterization of Theorem 8.4 is complete with respect to sentences not containing nesting of \mathcal{G} .

10.1. Definition. A formula $\psi \in \mathcal{L}^*(\mathcal{G})$ is *flat* if it does not contain nesting of \mathcal{G} . The collection of all flat formulas of $\mathcal{L}^*(\mathcal{G})$ over $\tau = \emptyset$ is denoted by \mathcal{F} . The collection of flat formulas with at most k variables is denoted by \mathcal{F}_k .

Restricting attention to flat formulas allows us to prove a version of Lemma 9.6 for arbitrary collections $G, G' \subseteq \mathcal{P}(M)$.

10.2. Lemma. *Let M be a set and $G, G' \subseteq \mathcal{P}(M)$ such that $G(k) = G'(k)$. Then*

$$(M, G) \models \varphi(\bar{a}) \Leftrightarrow (M, G') \models \varphi(\bar{a}),$$

for all $\varphi(\bar{x}) \in \mathcal{F}_{k+1}$, and $\bar{a} = a_1 \dots a_n \in M$.

Proof. We prove the claim using induction on φ . Again, it suffices to consider the case $\varphi = \mathcal{G}z\psi(z, \bar{y})$ only. Since the quantifier \mathcal{G} does not appear in ψ , the set $X = \{c \in M \mid (M, G) \models \psi(c, \bar{a})\}$ is fixed by all permutations of M which fix a_1, \dots, a_n . As in Lemma 9.4, we have that $X \subseteq \{a_1, \dots, a_n\}$ or $M \setminus X \subseteq \{a_1, \dots, a_n\}$. Now, since $G(k) = G'(k)$, it follows that $X \in G$ iff $X \in G'$. Therefore, by the induction assumption,

$$(M, G) \models \mathcal{G}z\psi(\bar{a}) \Leftrightarrow (M, G') \models \mathcal{G}z\psi(\bar{a}).$$

\square

Lemma 10.2 implies that Theorem 8.4 characterizes definable second order quantifiers completely with respect to sentences $\varphi \in \mathcal{F}$.

10.3. Theorem. *Let \mathcal{Q} be a second order quantifier of type $((1))$. Then \mathcal{Q} is definable with a sentence $\varphi \in \mathcal{F}$ if and only if \mathcal{Q} satisfies the assumption of Theorem 8.4.*

Proof. The defining sentence of \mathcal{Q} constructed in the proof of Theorem 8.4 is flat. On the other hand, by Lemma 10.2, a quantifier \mathcal{Q} defined by a flat sentence φ satisfies the assumption of Theorem 8.4 with $r - 1$, where r is the number of variables appearing in φ . \square

Finding a complete characterization of \mathcal{L}^* -definable quantifiers is an interesting open problem.

11. QUANTIFIERS OF ARBITRARY TYPES

It is easy to see that the methods we used to prove non-definability can be used with quantifiers of arbitrary types. Let us first consider monadic types $t = (s_1, \dots, s_w)$, where s_i is of the form $(1, \dots, 1)$ for $1 \leq i \leq w$. So we are now considering classes axiomatizable in the logic $\mathcal{L}^*(\mathcal{G}_1, \dots, \mathcal{G}_w)$, where the type of \mathcal{G}_i is s_i for $1 \leq i \leq w$.

Let (M, G_1, \dots, G_w) be a t -structure and $n \in \mathbb{N}$. Denote the set $\{(X_1, \dots, X_{r_i}) \in G_i : |X_j| \leq n \text{ or } |M \setminus X_j| \leq n \text{ for } 1 \leq j \leq r_i\}$ by $G_i(n)$. Now, Lemma 9.9 can be reformulated as follows.

11.1. Lemma. *Let (M, G_1, \dots, G_w) and (M, G'_1, \dots, G'_w) be structures of type t such that $G_i(k)$ and $G'_i(k)$ are permutation invariant and $G_i(k) = G'_i(k)$ for $1 \leq i \leq w$. Then*

$$(M, G_1, \dots, G_w) \equiv^{k+1} (M, G'_1, \dots, G'_w).$$

It is also straightforward to prove a version of Theorem 8.4 in this context.

11.2. Theorem. *Let t be a monadic type and \mathcal{Q} a quantifier of type t . Suppose there is $n \in \mathbb{N}$ such that for any set M and G_1, \dots, G_w we have $(M, G_1, \dots, G_w) \in \mathcal{Q}$ if and only if $(M, G_1(n), \dots, G_w(n)) \in \mathcal{Q}$. Then \mathcal{Q} is definable in \mathcal{L}^* .*

Since Lemma 10.2 also holds in this setting, the characterization of Theorem 10.3 can be generalized for quantifiers of monadic types.

11.3. Theorem. *Let t be a monadic type and \mathcal{Q} a quantifier of type t . Then \mathcal{Q} is definable in \mathcal{L}^* with a flat sentence if and only if \mathcal{Q} satisfies the assumption of Theorem 11.2.*

Let us then assume that $n > 1$ and $t = ((n))$. Suppose (M, G) and (M, G') are t -structures such that G and G' are permutation invariant. As in Lemma 9.6, $(M, G) \equiv^{k+1} (M, G')$ follows if G and G' contain the same n -ary relations which can be defined with k parameters in the empty vocabulary. Since G and G' are permutation invariant, we have that (a_1, \dots, a_k) -definable relations are fixed by all

permutations of M which fix the parameters a_1, \dots, a_k . However, the picture is not so simple as with unary relations. In particular, it is not so easy to describe the collection $G(k)$ since it is supposed to include all relations which are (a_1, \dots, a_k) -invariant for some $(a_1, \dots, a_k) \in M^k$. The collection $G(k)$ can be described but we do not take up this task here.

In some cases it is possible to utilize the fact that the diagonal $\{(a, \dots, a) \mid a \in M\}$ of M^n can be identified with M . To take an example, we show that the n -ary second order existential quantifier $\exists_n^2 = \{(M, G) \mid G \subseteq \mathcal{P}(M^n) \text{ and } G \neq \emptyset\}$ is not definable in the logic \mathcal{L}^* .

11.4. Proposition. *The quantifier \exists_n^2 is not definable in the logic \mathcal{L}^* .*

Proof. For $k \in \mathbb{N} \setminus \{0\}$, we set $M_k = \{1, \dots, 3k\}$, $G_k = \emptyset$ and

$$G'_k = \{\{(c, \dots, c) \mid c \in A\} \mid A \subseteq \{1, \dots, 3k\}, |A| > k \text{ and } |M \setminus A| > k\}.$$

It is easy to see that G'_k is permutation invariant for all $k \in \mathbb{N} \setminus \{0\}$. Also, none of the relations in G'_k can be defined with parameters fewer than $k + 1$. Consequently, $(M_k, G_k) \equiv^{k+1} (M_k, G'_k)$ for all $k \in \mathbb{N} \setminus \{0\}$. \square

12. FURTHER NON-DEFINABILITY RESULTS

In this section we show that logics which are equally strong over classes of first order structures may have radically different capabilities to define second order quantifiers.

Recall that if \mathcal{Q} is definable in \mathcal{L}^* then there are some Lindström quantifiers Q_1, \dots, Q_n such that

$$\text{FO}(\mathcal{Q}, Q_1, \dots, Q_n) \equiv \text{FO}(Q_1, \dots, Q_n).$$

In this section we show that the converse of the above does not hold. To prove this, we construct a second order quantifier which does not have any expressive power when combined with $\text{FO}(Q_1, \dots, Q_n)$ but is still not definable in $\text{FO}(Q_1, \dots, Q_n)$.

The construction of the quantifier \mathcal{Q}_0 in Theorem 12.1 is a slight modification of the construction in Theorem 3.1 in [1]. The construction in [1] is modified so that the resulting quantifier \mathcal{Q}_0 contains more than half of the isomorphism types of structures in every cardinality. This change does not affect the proof in [1] in any way. In particular, Claim 3 is proved in the same way as in [1].

12.1. Theorem. *Let \mathcal{B} be a countable collection of first and second order quantifiers and let t be a second order 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 $n \in \mathbb{N} \setminus \{0\}$ and for every quantifier $\mathcal{Q} \subseteq \mathcal{Q}_0$,*

$$\text{FO}(\mathcal{B}) \equiv \text{FO}(\mathcal{B}, \mathcal{Q}).$$

Proof. We assume first that $\mathcal{B} = \emptyset$ and $t = ((1))$. Let τ be a vocabulary containing countably many constant symbols and relation symbols of all arities. Let $X \in \tau$ be a unary predicate symbol. Let $(\varphi_i)_{i \in \mathbb{N}}$ be a list of all first order sentences over the

vocabulary τ . Let $\tau(m)$ denote the union of vocabularies of formulas $\varphi_0, \dots, \varphi_{m-1}$ minus the predicate X . Suppose that M is a set and $G \subseteq \mathcal{P}(M)$. We say that G is $\varphi_{<m}$ -definable over M if there is a $\tau(m)$ -structure \mathbb{M} with universe M and $i < m$ such that

$$G = \{B \mid (\mathbb{M}, B) \models \varphi_i\},$$

where X is interpreted as B .

We define a function $h: \mathbb{N} \rightarrow \mathbb{N}$ such that if $|M| \geq h(m)$, then the number of collections $G \subseteq \mathcal{P}(M)$ which are $\varphi_{<m}$ -definable over M is less than $\frac{1}{2}b_{|M|}(t)$, where $b_n(t)$ is the number of isomorphism types of t -structures with universe $\{1, \dots, n\}$. For every $m \in \mathbb{N}$ there is a polynomial $p_{\tau(m)}$ such that the number of $\tau(m)$ -structures with fixed universe M is at most $2^{p_{\tau(m)}(|M|)}$. Over any structure \mathbb{M} , each of the sentences φ_i defines at most one collection of subsets of M . Thus, over M there are at most $m2^{p_{\tau(m)}(|M|)}$ many $\varphi_{<m}$ -definable collections G . On the other hand, the number of elements in $\mathcal{P}(\mathcal{P}(M))$ is $2^{2^{|M|}}$ and obviously $b_n(t) \geq \frac{1}{n!}2^{2^n}$. We define $h(m)$ to be the least r such that

$$\frac{1}{r!}2^{2^r} > m2^{p_{\tau(m)}(r)+1}.$$

Let \mathcal{Q}_0 denote the class

$$\{(M, G) \mid (\forall m \in \mathbb{N})(|M| \geq h(m) \Rightarrow G \text{ is not } \varphi_{<m}\text{-definable over } M)\}.$$

It is easy to see that the definition of h ensures that the quantifier \mathcal{Q}_0 contains more than half of the isomorphism types of structures in every cardinality $n \in \mathbb{N} \setminus \{0\}$.

3. Claim. *Let $\mathcal{Q} \subseteq \mathcal{Q}_0$ be a quantifier. Then $\text{FO} \equiv \text{FO}(\mathcal{Q})$.*

Proof of Claim. The claim is proved using induction on the number of occurrences of \mathcal{Q} in a formula φ . We prove that any formula of $\text{FO}(\mathcal{Q})$ is equivalent to a first order formula over the same vocabulary. Let φ be a formula with $n+1$ occurrences of \mathcal{Q} . We may assume that the minimal sub-formula of φ containing $n+1$ occurrences of \mathcal{Q} is of the form $\mathcal{Q}X \psi(\bar{x})$. Otherwise, the claim follows directly from the induction hypothesis. By the induction hypothesis, $\psi(X, \bar{x})$ viewed as a sentence over the vocabulary $\tau_\psi \cup \{X, x_1, \dots, x_n\}$, where $\bar{x} = (x_1, \dots, x_n)$, is equivalent to a first order sentence ψ' over the same vocabulary. Let $i \in \mathbb{N}$ such that ψ' is equivalent to φ_i modulo renaming. Let \mathbb{M} be a τ_ψ -structure and $\bar{a} \in M$. The second order relation $\{B \subseteq M \mid (\mathbb{M}, B, \bar{a}) \models \varphi_i\}$ is $\varphi_{<i+1}$ -definable over M . So if $|M| > h(i+1)$ then $\mathbb{M} \not\models \mathcal{Q}X \psi'(X, \bar{a})$. Therefore, the formula $\mathcal{Q}X \psi'(X, \bar{x})$ is equivalent to a first order formula. Consequently, the formula φ is also equivalent to a first order formula. \square

This completes the proof of Theorem 12.1 in the case $\mathcal{B} = \emptyset$ and $t = ((1))$. The only effect the collection \mathcal{B} has on the proof is that the formula $\psi(\bar{x})$ in Claim 3 may have some free second order variables [1]. Suppose then that $t = (s_1, \dots, s_w)$ is an arbitrary second order type. Then we define \mathcal{Q}_0 analogously using a list $(\bar{\varphi}_i)_{i \in \mathbb{N}}$ of w -tuples of sentences containing every possible definition of a t -structure modulo renaming and proceed analogously. \square

12.2. Theorem. *Let \mathcal{B} be a countable collection of first and second order quantifiers. Then there is a quantifier \mathcal{Q} of type ((1)) such that $\text{FO}(\mathcal{B}) \equiv \text{FO}(\mathcal{B}, \mathcal{Q})$ but \mathcal{Q} is not definable in the logic $\text{FO}(\mathcal{B})$.*

Proof. Let \mathcal{Q}_0 be as in Theorem 12.1. Then, for all $A \subseteq \mathbb{N}$

$$\text{FO}(\mathcal{B}) \equiv \text{FO}(\mathcal{B}, \mathcal{Q}_A),$$

where $\mathcal{Q}_A = \{(M, G) \mid (M, G) \in \mathcal{Q}_0 \text{ and } |M| \in A\}$. Since \mathcal{Q}_0 contains structures in every cardinality, we have $\mathcal{Q}_A \neq \mathcal{Q}_{A'}$ for $A \neq A' \subseteq \mathbb{N}$. On the other hand, the logic $\text{FO}(\mathcal{B}, \mathcal{G})$ is countable. Thus, \mathcal{Q}_A is not definable in $\text{FO}(\mathcal{B})$ for some $A \subseteq \mathbb{N}$. \square

12.3. Remark. Since SO can be represented as $\text{FO}(\mathcal{B})$ for a countable collection \mathcal{B} of second order quantifiers, Theorem 12.2 implies that there is a second order quantifier \mathcal{Q} of type ((1)) which is not definable in SO and satisfies $\text{SO} \equiv \text{SO}(\mathcal{Q})$.

Theorem 12.2 can be improved to yield a quantifier \mathcal{Q} which is not definable in \mathcal{L}^* and satisfies $\text{FO}(\mathcal{Q}) \equiv \text{FO}$. The cardinality argument used in the proof of Theorem 12.2 does not work since the logic $\mathcal{L}^*(\mathcal{G})$ is uncountable. We use diagonalization to overcome this difficulty. We begin with an auxiliary definition [8].

12.4. Definition. Let $\varphi \in \mathcal{L}^*(\mathcal{G})$ be a formula and let Q_1, \dots, Q_n be the Lindström quantifiers appearing in φ . Then we write $\varphi = \varphi(Q_1, \dots, Q_n)$. If Q'_1, \dots, Q'_n are Lindström quantifiers such that the type of Q_i is the same as the type of Q'_i for $1 \leq i \leq n$, then $\varphi(Q'_1, \dots, Q'_n)$ denotes the formula obtained from φ by changing Q_i everywhere to Q'_i . Formulas obtained from each other by a change of quantifiers are called *similar*.

12.5. Theorem. *There is a quantifier \mathcal{Q} of type ((1)) with the following properties:*

- (1) $\text{FO}(\mathcal{Q}) \equiv \text{FO}$.
- (2) *The quantifier \mathcal{Q} is not definable in the logic \mathcal{L}^* .*

Proof. Let \mathcal{Q}_0 be as in Theorem 12.1. We shall define a sub-class \mathcal{Q} of \mathcal{Q}_0 which is not definable in \mathcal{L}^* . Let $\tau = \emptyset$ and let $(\psi_i)_{i \in \mathbb{N}}$ be a list of τ -sentences of $\mathcal{L}^*(\mathcal{G})$ such that every τ -sentence φ of $\mathcal{L}^*(\mathcal{G})$ is similar to some ψ_i . We define \mathcal{Q} inductively as follows. Assume that the numbers n_0, \dots, n_{j-1} and the classes $K_0, \dots, K_{j-1} \subseteq \mathcal{Q}_0$ have been defined already. Suppose (Q_1, \dots, Q_y) are the Lindström quantifiers appearing in ψ_j . The type of Q_i is denoted by s_i . Let (Q'_1, \dots, Q'_y) be Lindström quantifiers such that the type of Q'_i is s_i for $1 \leq i \leq y$. If

$$Q_i \cap \text{Str}(\tau_{s_i}, n) = Q'_i \cap \text{Str}(\tau_{s_i}, n)$$

for $1 \leq i \leq y$, then the sequences (Q_1, \dots, Q_y) and (Q'_1, \dots, Q'_y) are called *n-equivalent*. It is easy to see that for n-equivalent sequences we have

$$\text{Mod}(\psi_j(Q_1, \dots, Q_y)) \cap \text{Str}(t, n) = \text{Mod}(\psi_j(Q'_1, \dots, Q'_y)) \cap \text{Str}(t, n),$$

where $t = ((1))$. The number of τ_{s_i} -structures with universe $\{1, \dots, n\}$ is less than $2^{P_i(n)}$ where P_i is a polynomial. Therefore, the number of n-equivalence classes of

sequences is clearly less than $2^{f(n)}$ where $f(n) = \sum_{i=1}^y 2^{P_i(n)}$. Define n_j as the least n such that $n_j > n_{j-1}$ and

$$\frac{1}{2} \frac{1}{n!} 2^{2^n} > f(n).$$

Then there is $K_j \subseteq \mathcal{Q}_0$ of structures (M, G) with $|M| = n_j$ such that K_j cannot be represented in the form

$$\text{Mod}(\psi_j(Q_1, \dots, Q_y)) \cap \text{Str}(t, n_j),$$

for any sequence of quantifiers (Q_1, \dots, Q_y) . This concludes the construction. We define $\mathcal{Q} = \bigcup_{i \in \mathbb{N}} K_i$.

4. Claim. *The quantifier \mathcal{Q} is not definable in the logic \mathcal{L}^* .*

Proof of Claim. Suppose that $\mathcal{Q} = \text{Mod}(\varphi)$ for some $\varphi \in \mathcal{L}^*(\mathcal{G})$. Then $\varphi = \psi_j(Q_1, \dots, Q_y)$ for some $j \in \mathbb{N}$ and Lindström quantifiers Q_1, \dots, Q_y . Thus $K_j = \text{Mod}(\psi_j(Q_1, \dots, Q_y)) \cap \text{Str}(t, n_j)$, which contradicts the definition of \mathcal{Q} . \square

This completes the proof of Theorem 12.5. \square

IV. DEFINABILITY IN MONADIC SECOND ORDER LOGIC

In this chapter we study definability of second order quantifiers of type ((1)) in monadic second order logic, MSO. Proving non-definability turns out to be much harder in the case of MSO. We show that finding a complete characterization of definability is hard since this characterization would also characterize MSO-definable classes of graphs.

13. SOME NON-DEFINABILITY RESULTS

13.1. Theorem. *Let $S \subseteq \mathbb{N}$ be infinite and co-infinite. The following quantifiers are not definable in MSO:*

- $\text{Log}^2 = \{(M, R) \mid R \subseteq \mathcal{P}(M) \text{ and } |R| \geq \log(|M|)\}$
- $\mathcal{Q}_S = \{(M, R) \mid R \subseteq \mathcal{P}(M) \text{ and } |R| \in S\}$
- $\text{I}^2 = \{(M, R_1, R_2) \mid R_i \subseteq \mathcal{P}(M) \text{ and } |R_1| = |R_2|\}$

Proof. Suppose that I^2 is definable in MSO. Then the formula φ

$$\varphi = \text{I}^2 X, Y (\exists x(P_1(x) \wedge \forall y(X(y) \leftrightarrow y = x)), \exists x(P_2(x) \wedge \forall y(Y(y) \leftrightarrow y = x)))$$

is equivalent to some ψ in MSO over the vocabulary $\{P_1, P_2\}$. It is easy to verify that $\text{Mod}(\varphi) = \text{I}$ and hence the H\"artig quantifier I is definable in MSO. This is a contradiction since the quantifier I is not definable in FO (Fact 4.4) and $\text{MSO} \equiv \text{FO}$ on vocabularies containing only unary predicates. The non-definability of the quantifiers Log^2 and \mathcal{Q}_S can be proved analogously. \square

Let \mathcal{Q} be any of the quantifiers in Theorem 13.1. Theorem 13.1 is based on the observation that $\text{MSO}(\mathcal{Q}) > \text{MSO}$. By Theorem 12.2, there are quantifiers \mathcal{Q}' which are not definable in MSO and $\text{MSO}(\mathcal{Q}') \equiv \text{MSO}$. For such quantifiers \mathcal{Q}' , non-definability has to be established using different means.

Next we show that finding a complete characterization of MSO-definable quantifiers seems hard since such a characterization would also characterize classes of graphs definable in MSO.

Let $V = (M, R)$ be a graph, i.e., the relation $R \subseteq M^2$ is symmetric and irreflexive. Denote by \mathbb{M}_V the following second order structure (M, G) where

$$G = \{\{a, b\} \subseteq M \mid a, b \in M \text{ and } R(a, b)\}.$$

Suppose that \mathcal{C} is a class of graphs. Denote by $\mathcal{Q}_{\mathcal{C}}$ the following second order quantifier

$$\mathcal{Q}_{\mathcal{C}} = \{\mathbb{M}_V \mid V \in \mathcal{C}\}.$$

13.2. Proposition. *Let \mathcal{C} be a class of graphs. Then \mathcal{C} is definable in MSO if and only if the second order quantifier $\mathcal{Q}_{\mathcal{C}}$ is definable in MSO.*

The other half of the claim follows from Lemma 13.3.

13.3. Lemma. *Let $\varphi \in \text{MSO}[\{R\}]$ be a sentence. Then there is $\varphi^* \in \text{MSO}(\mathcal{G})[\emptyset]$ such that for any graph V ,*

$$V \models \varphi \Leftrightarrow \mathbb{M}_V \models \varphi^*.$$

Proof. Define a function $f: \text{MSO}[\{R\}] \rightarrow \text{MSO}(\mathcal{G})[\emptyset]$ as follows:

$$\begin{aligned} f(x = y) &= x = y \\ f(X(\bar{y})) &= X(\bar{y}) \\ f(R(x, y)) &= \mathcal{G}z(z = x \vee z = y) \\ f(\neg\varphi) &= \neg f(\varphi) \\ f(\varphi \wedge \psi) &= f(\varphi) \wedge f(\psi) \\ f(\exists x\varphi) &= \exists x f(\varphi) \\ f(\exists X\varphi) &= \exists X f(\varphi) \end{aligned}$$

5. Claim. *Let V be a graph. Then for all $\varphi(\bar{x}, \bar{X}) \in \text{MSO}[\{R\}]$, $\bar{b} = b_1 \dots b_n \in M$, and sets \bar{A} :*

$$V \models \varphi(\bar{b}, \bar{A}) \Leftrightarrow \mathbb{M}_V \models f(\varphi)(\bar{b}, \bar{A}).$$

Proof of Claim. Induction on φ . □

This completes the proof of Lemma 13.3. □

Proof of Proposition 13.2. Suppose $\mathcal{C} = \text{Mod}(\varphi)$ for some $\varphi \in \text{MSO}[\{R\}]$. By Lemma 13.3 we have that

$$\mathcal{Q}_{\mathcal{C}} = \text{Mod}(f(\varphi) \wedge \psi)$$

where ψ is the formula

$$\forall X(\mathcal{G}z(X(z)) \rightarrow \exists x\exists y(x \neq y \wedge \forall z(X(z) \leftrightarrow (z = x \vee z = y)))).$$

Suppose then that \mathcal{Q} is definable in MSO and $\mathcal{Q} = \mathcal{Q}_{\mathcal{C}}$ for some class of graphs \mathcal{C} . Now $\text{Mod}(\varphi) = \mathcal{C}$ where

$$\varphi = \mathcal{Q}X(\exists x\exists y(R(x, y) \wedge \forall z(X(z) \leftrightarrow (z = x \vee z = y)))).$$

Since $\text{MSO}(\mathcal{Q}) \equiv \text{MSO}$, we get that \mathcal{C} is definable in MSO. □

Proposition 13.2 shows that finding a complete characterization of definability in the case of MSO is difficult. On the other hand, such a characterization could be possible for some interesting subclasses of quantifiers.

V. DEFINABILITY IN EXTENSIONS OF SECOND ORDER LOGIC

In the previous chapters we have mainly concentrated on definability of quantifiers of type $((1))$. In this chapter we study definability of second order quantifiers in terms of their types. The main result of this chapter is the analogue of the Hierarchy Theorem in [8]:

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 $\text{FO}(\mathcal{B})$.*

We shall prove the following.

Theorem 17.1. *Let t be a second order type and \mathcal{Q} the collection of all second order generalized quantifiers of types lower than t . Then there is a quantifier Q of type t which is not definable in the logic $\text{SO}(\mathcal{Q})$.*

Before proving Theorem 17.1, 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 17.1.

14. DEFINABILITY AND TYPES

A simple coding argument shows that for any finite number of Lindström quantifiers Q_1, \dots, Q_n there is a Lindström quantifier Q such that $\text{FO}(Q) \equiv \text{FO}(Q_1, \dots, Q_n)$. The quantifier Q ends up having a more complex type than any of the quantifiers Q_1, \dots, Q_n . Actually, the same result holds even for some countable collections of Lindström quantifiers (Fact 14.1). The proof of Fact 14.1 can be found in [6] and in [12] in a slightly different context.

14.1. Fact. *Suppose \mathcal{B} is a countable collection of n -ary Lindström quantifiers. Then there is a Lindström quantifier Q of type $(n + 1)$ such that $\text{FO}(Q) \geq \text{FO}(\mathcal{B})$.*

In this section we prove that the analogue of Fact 14.1 holds also for second order quantifiers. The proof of Proposition 14.2 is similar to the proof of Lemma 1.6 in [6].

14.2. Proposition. *Let \mathcal{B} be a countable collection of second order quantifiers such that the variable arity of Q' is at most n for all $Q' \in \mathcal{B}$. Then there is a quantifier Q of type $((n + 1))$ such that Q' is definable in $\text{SO}(Q)$ for all $Q' \in \mathcal{B}$.*

Proof. We write $\mathcal{B} = \{Q_i\}_{i \in \mathbb{N} \setminus \{0\}}$. Without loss of generality we can assume that the type of Q_i is (s_1, \dots, s_i) where s_j is a sequence of n 's of length i . This simplification can be justified by using dummy variables analogously to the first order case (see Theorem 12.1.4. in [4]). We can also assume that none of the quantifiers Q_i contains structures of the form $(M, \emptyset, \dots, \emptyset)$.

Let $(M, G_1^i, \dots, G_i^i) \in \mathcal{Q}_i$. We use basically the same coding trick as in the first order case [6] and code each $\bar{X} \in G_j^i$ by a relation $R \subseteq M^{n+1}$. For the purposes of coding we choose a bijection

$$f: \{(i, j, k) \mid i \in \mathbb{N} \setminus \{0\}, 1 \leq j, k \leq i\} \rightarrow \{2^i \mid i \in \mathbb{N}\}$$

and define \mathcal{Q} to be the class of structures (M, G) with $G \subseteq \mathcal{P}(M^{n+1})$ such that there exists $i \in \mathbb{N} \setminus \{0\}$, G_1^i, \dots, G_i^i with $(M, G_1^i, \dots, G_i^i) \in \mathcal{Q}_i$, and distinct elements $a_1^{1,1}, \dots, a_{f(i,1,1)}^{1,1}, \dots, a_1^{i,i}, \dots, a_{f(i,i,i)}^{i,i} \in M$ for which

$$G = \{R_{\bar{X}} \mid \bar{X} \in G_j^i \text{ for some } 1 \leq j \leq i\},$$

where $R_{\bar{X}} = \bigcup_{1 \leq k \leq i} \{(\bar{b}, a_d^{j,k}) \mid \bar{b} \in X_k \text{ and } 1 \leq d \leq f(i, j, k)\}$. The quantifier \mathcal{Q} is as wanted, since for $i \in \mathbb{N} \setminus \{0\}$ we have

$$\models \mathcal{Q}_i \bar{X}_1, \dots, \bar{X}_i (\psi_1, \dots, \psi_i) \leftrightarrow \theta_i \vee \chi,$$

where

$$\begin{aligned} \chi &= \exists x_1 \dots x_l \left(\bigwedge_{p \neq q} (x_p \neq x_q) \wedge \mathcal{Q} R \bigvee_{1 \leq j \leq i} \phi_j \right) \\ \phi_j &= \exists R_1 \dots R_i (\psi_j(\bar{R}) \wedge \forall z_1 \dots z_{n+1} (R(z_1, \dots, z_{n+1}) \leftrightarrow \bigvee_{1 \leq s \leq i} v_s)) \\ v_s &= R_s(z_1, \dots, z_n) \wedge \bigvee_{l_{(j,s-1)} < h \leq l_{(j,s)}} (z_{n+1} = x_h) \end{aligned}$$

and where $l_{(j,s)} = (f(i, 1, 1) + f(i, 1, 2) + \dots + f(i, j, s))$, $l = l_{(i,i)}$, $l_{(1,0)} = 0$ and $l_{(j,0)} = l_{(j-1,i)}$ for $j > 1$. The formula θ_i takes care of the interpretations of \mathcal{Q}_i on models with less than l elements. \square

14.3. Remark. It is worth noting that to prove definability (as formulated in Definition 6.3) of \mathcal{Q}_i in the logic $\text{SO}(\mathcal{Q})$ the proof of Proposition 14.2 has to be slightly modified. We have to replace the formula ψ_j with $\mathcal{G}_j \bar{x}_1, \dots, \bar{x}_i (X_1(\bar{x}_1), \dots, X_i(\bar{x}_i))$ for $1 \leq j \leq i$ in χ and θ_i . It is now easy to see that $\text{Mod}(\theta_i \vee \chi) = \mathcal{Q}_i$ as wanted.

15. FAGIN'S LEMMA FOR SECOND ORDER STRUCTURES

The Hierarchy Theorem in [8] is based on a result called Fagin's lemma [5]. Fagin's lemma is concerned with the number of isomorphism types of finite relational structures. Let $s = (l_1, \dots, l_r)$ be a first order type. Denote the number of τ_s -structures with universe $\{1, \dots, n\}$ by $a_n(s)$ and the number of isomorphism types of τ_s -structures with universe $\{1, \dots, n\}$ by $b_n(s)$. It is easy to see that $a_n(s) = 2^{(n^{l_1} + \dots + n^{l_r})}$. The real task is to determine the value $b_n(s)$. Fagin's lemma answers this question.

Fagin's lemma. *Let $s = (l_1, \dots, l_r)$ 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 [5]. \square

In this section we prove a version of Fagin's lemma for second order structures. The proof follows the proof in [5] closely. Let $s = (l_1, \dots, l_r)$. We write $P_s(x)$ for $x^{l_1} + \dots + x^{l_r}$. Also, for a second order type t , the notations $a_n(t)$ and $b_n(t)$ are used as for first order structures.

15.1. Lemma. *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*

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

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

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

□

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

- Denote by $\text{spt}(\pi)$ the *support* of π

$$\text{spt}(\pi) = \{a \in X \mid \pi(a) \neq a\}.$$

- The cardinality of the set $\text{spt}(\pi)$ 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 15.4. Recall that S_n denotes the set of permutations of the set $\{1, \dots, n\}$.

15.3. Lemma. *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].

□

15.4. Theorem. *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 $\pi_{(1)}(X) = \pi[X]$.

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

Proof of Claim. 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 15.3,

$$(15.5) \quad b_n(t) = \frac{1}{n!} \sum_{\pi \in S_n} N(\pi).$$

Combining (15.5) and Claim 6,

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

The original claim is equivalent to

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

Hence, it suffices to prove

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

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

$$(15.6) \quad \lim_{n \rightarrow \infty} \sum_{\pi \in S'_n} 2^{(c(\pi_{(1)}) - 2^n)} = 0,$$

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

7. Claim. *Let $n > 5$ and let π be a permutation of the set $\{1, \dots, n\}$. If $s(\pi) = m$ then $s(\pi_{(1)}) \geq m(n-2)$.*

Proof of Claim. 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 $\{a, b\}$ (plus their complements) such that $\{\pi(a), \pi(b)\} \cap \{a, b\} = \emptyset$ is at least $m(m-3)$. By summing these numbers, we get $m(n-2)$. \square

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

$$(15.7) \quad c(\pi_{(1)}) \leq 2^n - x/2.$$

By the estimate (15.7) and Claim 7, the terms of the sequence in (15.6) are dominated by

$$(15.8) \quad \sum_{\pi \in S'_n} 2^{-(s(\pi)(n-2))/2}.$$

The number of permutations π of $\{1, \dots, n\}$ such that $s(\pi) = m$ is less than $\binom{n}{m} m! \leq n^m$. Therefore, the sequence in (15.8) is dominated by

$$(15.9) \quad \sum_{m=2}^n n^m 2^{-m(n-2)/2}.$$

Now (15.9) is equal to

$$(15.10) \quad \sum_{m=2}^n 2^{-\frac{1}{2}m((n-2)-2\log(n))}.$$

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 (15.10) occurs when $m = 2$. Thus, for $n \geq n_0$, (15.10) is dominated by

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

which converges to 0.

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

$$\mathcal{P}(M^{l_1^i}) \times \dots \times \mathcal{P}(M^{l_{r_i}^i})$$

induced by π . Then (cf. Claim 6) π is an automorphism of $N(\pi)$ many t -structures where

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

By Lemma 15.3,

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

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

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

Hence, it suffices to prove

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

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

$$(15.11) \quad \lim_{n \rightarrow \infty} \sum_{\pi \in S'_n} 2^{\sum_{i=1}^w (c(\pi_{s_i}) - 2^{P_{s_i}(n)})} = 0.$$

Suppose $s(\pi) = m$. We want to find a lower bound for $s(\pi_{s_i})$. If $l_1^i = 1$, then based on the calculations made in Claim 7 the number of sets A not fixed in the first coordinate is greater than $m(n-2)$. Thus, no tuple of the form

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

where $B_j \in \mathcal{P}(M^{l_j^i})$ for $2 \leq j \leq r_i$, is fixed by π_{s_i} . Hence, we have $s(\pi_{s_i}) > m(n-2)$. If $l_1^i = q > 1$, then $\pi_q: M^q \rightarrow M^q$ 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 (15.7) the terms of the sequence in (15.11) are dominated by

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

and thus by

$$(15.12) \quad \sum_{\pi \in S'_n} 2^{-(s(\pi)(n-2))/2}.$$

Sequence (15.12) is identical to sequence (15.8) which converges to 0. \square

16. AN ORDERING OF SECOND ORDER TYPES

In this section we define an ordering for second order types arising naturally from Theorem 15.4. The following results concerning first order types and structures are taken from [8].

For a first order type $s = (l_1, \dots, l_r)$, the *pattern* of s is the tuple $p(s) = (p_1, \dots, p_a)$, where $p_i = |\{1 \leq j \leq r : l_j = i\}|$ and $a = \max\{l_j : 1 \leq j \leq r\}$. Let S denote the set of finite sequences $p = (p_1, \dots, p_a)$ of natural numbers such that $p_a \neq 0$. We denote a by $l(p)$.

16.1. Definition. 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 $p_i \neq q_i$.

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

Proof. See [8]. \square

We denote the order-isomorphism between $(S, <)$ and $(\omega^\omega, \epsilon)$ by Φ .

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

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

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

Proof. See [8]. \square

Now we turn to second order types. We shall define an ordering for second order types analogously to Definition 16.3. We shall first define an ordering for a certain collection B of functions $j: \omega^\omega \rightarrow \omega$ and then associate each second order type t with a unique element of B .

For a function $j: \omega^\omega \rightarrow \omega$, denote by $\text{spt}(j)$ the support of j

$$\text{spt}(j) = \{\alpha \in \omega^\omega \mid j(\alpha) \neq 0\}.$$

We shall restrict attention to functions j with finite support.

16.5. Definition. Let B denote the set

$$\{j: \omega^\omega \rightarrow \omega : |\text{spt}(j)| \in \omega \setminus \{0\}\}.$$

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.

16.6. Proposition. *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 $\{\alpha_1, \dots, \alpha_n\}$ ($\alpha_i < \alpha_j$ iff $i < j$) be the set of those ordinals α for which $j(\alpha) \neq 0$. We divide j into component functions j_1, \dots, 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, \dots, j_n be the component functions of j . Define

$$(16.7) \quad F(j) = F(j_n) + F(j_{n-1}) + \dots + F(j_1).$$

Since every ordinal $\alpha < \omega^{\omega^\omega}$ has a unique representation in the form (16.7), 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

16.8. Definition. Let $t = (s_1, \dots, s_w)$ be a second order type. Define $j_t: \omega^\omega \rightarrow \omega$ 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 $j_t \prec j_{t'}$.

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

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

Proof. By Theorem 15.4,

$$(16.10) \quad \frac{b_n(t')}{b_n(t)} \sim 2^{((2^{P_{d_1}(n)} + \dots + 2^{P_{d_e}(n)}) - (2^{P_{s_1}(n)} + \dots + 2^{P_{s_w}(n)}))}.$$

After subtraction and by the assumption $t' \prec t$, the right-hand side of (16.10) is dominated by

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

where $d \in \{d_1, \dots, d_e\}$, $s \in \{s_1, \dots, s_w\}$, and $d < s$. By Lemma 16.4, we have $\lim_{n \rightarrow \infty} 2^{P_d(n)}/2^{P_s(n)} = 0$. Thus

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

□

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

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

17. A HIERARCHY OF DEFINABILITY

In this section we prove that on every level t of the ordering \prec there is a quantifier \mathcal{Q} 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 17.1 is analogous to the proof of Theorem 5 in [8].

17.1. Theorem. *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 $\text{SO}(\mathbb{Q})$.*

Proof. Let $\tau = \emptyset$ and let $(\psi_i)_{i \in \omega}$ be a list of τ -sentences of the logic $\text{SO}(\mathbb{Q}, \mathcal{G}_1, \dots, \mathcal{G}_w)$ such that every τ -sentence of $\text{SO}(\mathbb{Q}, \mathcal{G}_1, \dots, \mathcal{G}_w)$ is similar (see Definition 12.4) to some ψ_i . Suppose n_0, \dots, n_{j-1} and classes K_0, \dots, K_{j-1} have been defined already. Let $\mathcal{Q}_1, \dots, \mathcal{Q}_y$ be the quantifiers in \mathbb{Q} appearing in ψ_j . The type of \mathcal{Q}_i is denoted by t_i . Let $\mathcal{Q}'_1, \dots, \mathcal{Q}'_y$ be quantifiers such that the type of \mathcal{Q}'_i is t_i for $1 \leq i \leq y$. If

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

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

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

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 16.11, 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

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

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

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

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

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

which contradicts the definition of \mathcal{Q} . \square

This completes the proof of Theorem 17.1 . \square

Next we show that the diagonalization argument used in the proof of Theorem 17.1 can be modified so that the resulting non-definable quantifier can be extended to a non-definable monotone quantifier. The proof is adapted from Theorem 7 in [8] where an analogous result is proved for Lindström quantifiers.

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

The quantifiers Most^2 and \exists_k^2 are examples of monotone quantifiers and the quantifiers Even and Three are obviously not monotone.

17.3. Theorem. *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 $\text{SO}(\mathbb{Q})$.*

Before going to the proof of Theorem 17.3, we introduce an auxiliary concept [8]. Let $M = \{1, \dots, n\}$ and let (M, G_1, \dots, 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 $\sigma: \omega \rightarrow \omega$. We say that (M, G_1, \dots, G_w) is σ -balanced if $\sum_{i=1}^w |G_i| = \sigma(n)$. Define a function $f: \omega \rightarrow \omega$ as follows: $f(n)$ is the least y such that

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

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

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

$$\lim_{n \rightarrow \infty} \frac{b_n(t')}{I_n(\text{Str}_f(t))} = 0.$$

Proof. By the choice of f , we have

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

Therefore,

$$(17.5) \quad \frac{b_n(t')}{I_n(\text{Str}_f(t))} \leq \frac{b_n(t')(\sum_{i=1}^w 2^{P_{s_i}(n)} + 1)}{b_n(t)}.$$

By the estimate in Lemma 16.9, the right-hand side of (17.5) is eventually dominated by

$$(17.6) \quad c \left(\sum_{i=1}^w 2^{P_{s_i}(n)} + 1 \right) 2^{(e2^{P_d(n)} - 2^{P_s(n)})},$$

where $c > 1$ is a constant and $d < s$. By Lemma 16.4, $\lim_{n \rightarrow 0} 2^{P_d(n)} / 2^{P_s(n)} = 0$, hence there are $u, h, r \in \omega$ such that, for $n \geq r$, the sequence in (17.6) is dominated by

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

□

Proof of Theorem 17.3. We modify the proof of Theorem 17.1 so that the diagonalization is realized with respect to the class $\text{Str}_f(t)$. In other words, we ensure that

$$K_j \neq \text{Mod}(\psi_j(\mathcal{Q}_1, \dots, \mathcal{Q}_y)) \cap \text{Str}_f(t, n_j),$$

for all sequences $(\mathcal{Q}_1, \dots, \mathcal{Q}_y)$ of quantifiers. This is possible by Lemma 17.4, since the number of collections of the form

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

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

$$\{(M, G_1, \dots, G_w) \mid \text{there is } (M, G'_1, \dots, G'_w) \in \mathcal{Q} \text{ such that } G'_i \subseteq G_i\}.$$

Since $\mathcal{Q}' \cap \text{Str}_f(t) = \mathcal{Q}$ and $\mathcal{Q} \neq (\text{Mod}(\varphi) \cap \text{Str}_f(t))$ for all $\varphi \in \text{SO}(\mathbb{Q}, \mathcal{G}_1, \dots, \mathcal{G}_w)$, the quantifier \mathcal{Q}' is not definable in the logic $\text{SO}(\mathbb{Q})$. □

Theorem 17.1 can be also improved to yield an non-definable *recursive quantifier*, i.e., there is a *Turing machine* which accepts a coding of a structure (M, G_1, \dots, G_w) if and only if $(M, G_1, \dots, G_w) \in \mathcal{Q}$.

17.7. Theorem. *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 $\text{SO}(\mathbb{Q})$.*

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 [8].

Proof of Theorem 17.7. Let $\tau = \emptyset$ and let $(\psi_i)_{i \in \omega}$ be an effective enumeration of t -sentences of $\text{SO}(\mathbb{Q}, \mathcal{G}_1, \dots, \mathcal{G}_w)$ up to similarity.

Input: $\mathcal{M} = (M, G_1, \dots, G_w)$ of type t and cardinality n .

Algorithm:

```

 $j \leftarrow 0; n_{-1} \leftarrow 0;$ 
repeat  $\{n_j \leftarrow \min\{n' > n_{j-1} \mid \text{Cond}(n', j)\}; j \leftarrow j + 1\}$ 
until  $n_j \geq n;$ 
if  $n_j > n$ 
then reject  $\mathcal{M}$ 
else
   $\{\text{FindQuant}(K, n, j);$ 
  if  $\mathcal{M} \in K$ 
  then accept  $\mathcal{M}$ 
  else reject  $\mathcal{M}\}$ .

```

The condition $\text{Cond}(n, j)$ 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

$$(17.8) \quad \text{Mod}(\psi_j(\mathcal{Q}_1, \dots, \mathcal{Q}_y)) \cap \text{Str}(t, n_j),$$

for any sequence $(\mathcal{Q}_1, \dots, \mathcal{Q}_y)$ of quantifiers. If $n = n_j$ for some j then the procedure $\text{FindQuant}(K, n, j)$ returns a canonical class $K(n, j)$ of t -structures of cardinality n which cannot be represented in the form (17.8). 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
   $\{K \leftarrow \text{Next}(K);$ 
  repeat
     $\{\overline{\mathcal{Q}} \leftarrow (\emptyset, \dots, \emptyset);$ 
     $b \leftarrow \text{Definable}(K, \overline{\mathcal{Q}}, j);$ 
     $\overline{\mathcal{Q}} \leftarrow \text{NextSeq}(\overline{\mathcal{Q}})\}$ 
  until  $b$  or  $\overline{\mathcal{Q}} = (\emptyset, \dots, \emptyset)\}$ 
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 $\text{Str}(t, M)$, where $\text{Str}(t, M)$ is the set of t -structures with universe M . The procedure $\text{Next}(K)$ returns the least quantifier $K' \subseteq \text{Str}(t, M)$ which is greater than K with respect to $<_t$. If such K' does not exist, $\text{Next}(K) = \emptyset$. $\text{NextSeq}(\overline{\mathcal{Q}})$ is a similar function for sequences of quantifiers of types lower than t . The condition

$\text{Definable}(K, \overline{\mathcal{Q}}, j)$ is satisfied just in case $K = \{\mathcal{M}' \in \text{Str}(t, M) \mid \mathcal{M}' \models \psi_j(\overline{\mathcal{Q}})\}$. \square

Suppose $((1)) \prec t$ and \mathcal{Q} is a quantifier of type t as in Theorem 17.1. Although \mathcal{Q} is not definable in the logic $\text{SO}(\mathbb{Q})$, where \mathbb{Q} is the collection of quantifiers of types lower than t , we are bound to have

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

The reason is that any class of finite first order structures is axiomatizable in $\text{SO}(M_1^1)$, 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 17.1 holds.

17.9. Theorem. *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 $\text{SO}(\mathbb{Q})$ and $\text{FO} \equiv \text{FO}(\mathcal{Q})$.*

Proof. Let \mathcal{Q}_0 be as in Theorem 12.1 when applied to $\mathcal{B} = \emptyset$ and t . We modify the diagonalization argument in the proof of Theorem 17.1 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 Theorem 12.1. \square

18. A HIERARCHICAL QUANTIFIER

In this section we show that the binary second order existential quantifier cannot be defined in the logic $\text{FO}(\mathbb{Q})$, where \mathbb{Q} is the collection of monadic second order generalized quantifiers, i.e., quantifiers of types less than $((2))$. Since the logic $\text{FO}(\mathbb{Q})$ is capable of defining all classes of first order structures (cf. Theorem 6.2 in [1]), we also get a concrete example of a quantifier \mathcal{Q} and a logic \mathcal{L} extending MSO such that $\mathcal{L}(\mathcal{Q}) \equiv \mathcal{L}$ and \mathcal{Q} is not definable in \mathcal{L} .

18.1. Theorem. *The quantifier \exists_2^2 is not definable in the logic $\text{FO}(\mathbb{Q})$.*

Denote by $\mathcal{L}_{(l,r)}$ the fragment of $\text{FO}(\mathbb{Q}, \mathcal{G})$ containing only formulas, whose free and bound variables are among x_1, \dots, x_l and X_1, \dots, X_r . We write $\mathcal{M} \equiv_{l,r} \mathcal{M}'$ if the models \mathcal{M} and \mathcal{M}' satisfy the same sentences of $\mathcal{L}_{(l,r)}$ of vocabulary $\tau = \emptyset$.

Let M be a set and let $G \subseteq \mathcal{P}(M^2)$ be defined as follows:

$$G = \{R \subseteq M^2 \mid \forall a \neq b \in M (R(a, b) \leftrightarrow \neg R(b, a))\}.$$

The set G is the collection of so-called *tournaments* over M . It is easy to see that G is permutation invariant.

18.2. Lemma. *Let $l, r \in \mathbb{N}$, $k = l + r$, and let M be a set such that $|M| > 2^k$. Suppose $a_1, \dots, a_l \in M$ and $A_1, \dots, A_r \subseteq M$. Then no relation $R \in G$ is (\bar{a}, \bar{A}) -invariant.*

Proof. The elements a_1, \dots, a_l and sets A_1, \dots, A_r induce a partition of M which has at most 2^k different classes. Since $|M| > 2^k$, one of the classes contain at least two elements, say, c and d . Let f be a permutation of M such that $f(c) = d$, $f(d) = c$, and $f(b) = b$ for $b \in M \setminus \{c, d\}$. Then any (\bar{a}, \bar{A}) -invariant relation R is fixed by f . This implies that either $R(c, d)$ and $R(d, c)$ or $\neg R(c, d)$ and $\neg R(d, c)$. Thus, we have that $R \notin G$. \square

Theorem 18.1 now follows from Lemma 18.3.

18.3. Lemma. *Let $l, r \in \mathbb{N}$, $k = l + r$, and let M be a set such that $|M| > 2^k$. Then $(M, G) \equiv_{l,r} (M, G')$, where $G' = \emptyset$.*

Proof. We prove using induction on $\varphi(\bar{x}, \bar{X}) \in \mathcal{L}_{(l,r)}[\emptyset]$ that

$$(M, G) \models \varphi(\bar{a}, \bar{A}) \Leftrightarrow (M, G') \models \varphi(\bar{a}, \bar{A}),$$

for all $\bar{a} = a_1 \dots a_l \in M$ and $\bar{A} = A_1 \dots A_r \subseteq M$. We consider the case $\varphi = \mathcal{G}xy \psi(x, y, \bar{z}, \bar{X})$ only since other cases are trivial. By Lemma 18.2, we have $(M, G) \not\models \varphi(\bar{a}, \bar{A})$ since the relation

$$\{(b_1, b_2) \in M^2 \mid (M, G) \models \psi(b_1, b_2, \bar{a}, \bar{A})\}$$

is (\bar{a}, \bar{A}) -invariant and hence not in G . On the other hand, since $G' = \emptyset$ we have $(M, G') \models \varphi(\bar{a}, \bar{A})$. \square

19. CONCLUSION AND FURTHER DIRECTIONS

Definability questions of Lindström quantifiers have been studied extensively in finite model theory. Many of the questions concerning definability of Lindström quantifiers also make sense for second order generalized quantifiers. On the other hand, we have seen that genuinely new type of questions arise in the second order case.

In this thesis we have studied definability of second order generalized quantifiers in different contexts. The most obvious questions left unanswered are questions of the form "Is the quantifier \mathcal{Q} definable in $\text{FO}(\mathcal{Q}')$ ". We have touched this area in Chapter IV where we considered definability in MSO. We conjecture that the quantifier Most^2 is not definable in MSO. If one is able to show $\text{MSO}(\text{Most}^2) > \text{MSO}$, then the conjecture follows, of course. It would be also interesting to find a concrete quantifier \mathcal{Q} which is not definable in MSO and satisfies $\text{MSO}(\mathcal{Q}) \equiv \text{MSO}$.

A more general project is to study definability in terms of all quantifiers of a certain type or with some property. For example, Theorem 18.1 implies that no extension of FO by finitely many monadic second order generalized quantifiers defines the quantifier \exists_2^2 .

Let us then consider Theorem 17.1. If we replace SO by FO in Theorem 17.1, then it remains open whether there is a strict hierarchy also in the expressive power of the logics $\text{FO}(\mathbb{Q})$ with respect to types $t \prec ((2))$. Andersson showed that quantifiers of type $((1; n))$, where $(1; n)$ is a sequence of 1's of length n , can define all Lindström quantifiers of arity at most $2^n - 1$ (cf. Theorem 6.2 in [1]). So it might be possible to prove that for any type $t \prec ((2))$ there is a quantifier \mathcal{Q} of type t which is not definable in $\text{FO}(\mathbb{Q})$ and satisfies $\text{FO}(\mathbb{Q}) < \text{FO}(\mathbb{Q}, \mathcal{Q})$. It is also an open question whether for every type t any quantifier \mathcal{Q} of some type $t' \prec t$ is definable in the logic $\text{SO}(\mathbb{Q}_t)$ where \mathbb{Q}_t is the collection of all second order quantifiers of type t .

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] A. Dawar. Generalized quantifiers and logical reducibilities. *J. Logic Comput.*, 5(2):213–226, 1995.
- [4] H.-D. Ebbinghaus and J. Flum. *Finite model theory, 2nd edition*. Perspectives in Mathematical Logic. Springer-Verlag, 1999.
- [5] R. Fagin. The number of finite relational structures. *Discrete Math.*, 19(1):17–21, 1977.
- [6] L. Hella. Definability hierarchies of generalized quantifiers. *Ann. Pure Appl. Logic*, 43(3):235–271, 1989.
- [7] L. Hella. Logical hierarchies in PTIME. *Inform. and Comput.*, 129(1):1–19, 1996.
- [8] L. Hella, K. Luosto, and J. Väänänen. The hierarchy theorem for generalized quantifiers. *J. Symbolic Logic*, 61(3):802–817, 1996.
- [9] P. G. Kolaitis and J. A. Väänänen. Generalized quantifiers and pebble games on finite structures. *Ann. Pure Appl. Logic*, 74(1):23–75, 1995.
- [10] J. Kontinen. Definability of second order generalized quantifiers. *To appear in Arch. Math. Logic*.
- [11] P. Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966.
- [12] J. Väänänen. Remarks on generalized quantifiers and second order logics. *Prace Nauk. Inst. Mat. Politech. Wrocław.*, (14 Ser. Konfer. No. 1):117–123, 1977.