Definability of Second Order Generalized Quantifiers

Juha Kontinen

Abstract

We study second order generalized quantifiers on finite structures. One starting point of this research has been the notion of definability of Lindström quantifiers. We formulate an analogous notion for second order generalized quantifiers and study definability of second order generalized quantifiers in terms of Lindström quantifiers.

1 Introduction

The expressive power of first order logic, FO, is rather limited on finite structures. Therefore, many extensions of first order logic have been studied in the context of finite model theory. *Lindström quantifiers* [3] give direct means to define extensions of various logics. 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

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

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

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

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

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

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 $Q_1, \ldots, Q_n$ if there is a way to express every formula of the form $Qx\,\varphi(x)$ using the quantifiers $Q_1, \ldots, Q_n$. For example, all the counting quantifiers $\exists^{\geq n}$ can be defined using just $\exists$ in the obvious way. Actually, for definability, it suffices to require that the class Q is axiomatizable in the logic $FO(Q_1, \ldots, Q_n)$. Indeed, if $Q = Mod(\psi)$ 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 notion of *second order generalized quantifier* was introduced in [1]. 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 this paper we study definability of second order generalized quantifiers. The notion of definability we study is analogous to the notion of definability of Lindström quantifiers. Let us take an example.

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

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

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

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) 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 first order quantifier variables. It turns out that a defining sentence $\psi \in \mathcal{L}'$ can be then converted into a uniform definition of $\mathcal{Q}$ in the logic $\mathcal{L}$.

We shall show that the notion of definability studied in this paper has the desired properties. In particular, we shall show that definability of $\mathcal{Q}$ in a logic $\mathcal{L}$, where $\mathcal{L}$ is of the form $FO(\mathcal{B})$ for a collection $\mathcal{B}$ of Lindström quantifiers, implies that $\mathcal{L}(\mathcal{Q}) \equiv \mathcal{L}$. We can now 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. In the last part of this paper we 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.

1.1 Notations

Vocabularies τ are finite sets consisting of relation symbols and constant symbols. All structures are assumed to be finite. The universe of a structure $\mathbb{M}$ is denoted by M. For $\tau = \emptyset$, we denote τ-structures by their universes. The cardinality of a set X is denoted by $|X|$. 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 $Str(\tau)$, and $Str(\tau, n)$ is the restriction of $Str(\tau)$ to structures having cardinality n. If φ is a τ-sentence then the class of τ-models of φ is denoted by $Mod(\varphi)$.

1.2 Lindström quantifiers

Lindström [3] defined the general notion of (first order) generalized quantifier as a class of structures of a fixed finite relational vocabulary such that the class is closed under isomorphisms.

Definition 1.2. Let $s = (l_1, \ldots, l_r)$ be a tuple of positive integers. A *Lindström quantifier* of type s is a class Q of structures of vocabulary $\tau_s = \{P_1, \ldots, P_r\}$ such that P_i is l_i-ary for $1 \leq i \leq r$, and Q is closed under isomorphisms.

Example 1.3. Let us look at some examples of Lindström quantifiers.

$$\begin{aligned} \forall &= \{(M, P) \mid P = M\} \\ Q_{even} &= \{(M, P) \mid P \subseteq M \text{ and } |P| \text{ even}\} \\ I &= \{(M, P, S) \mid P, S \subseteq M \text{ and } |P| = |S|\} \end{aligned}$$

Quantifiers can be also defined over a fixed set M.

Definition 1.4. A quantifier Q of type s over M is a subset of $\mathcal{P}(M^{l_1}) \times \cdots \times \mathcal{P}(M^{l_r})$ 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 $(A_1, \dots, A_r) \in \mathcal{P}(M^{l_1}) \times \cdots \times \mathcal{P}(M^{l_r})$ and $f \in S_M$, where S_M is the set of permutations of M.

Example 1.5. Let Q be a Lindström quantifier of type s and M 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.

Definition 1.6. The extension $FO(Q)$ 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:

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

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

1.3 Definability of Lindström quantifiers

Definition 1.7. Let $\mathcal{L}$ and $\mathcal{L}'$ be logics. The logic $\mathcal{L}'$ is *as strong as* the logic $\mathcal{L}$ ($\mathcal{L} \leq \mathcal{L}'$) if for every $\varphi \in \mathcal{L}$ over any vocabulary there exists a formula $\psi \in \mathcal{L}'$ 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 $\mathcal{L}' \leq \mathcal{L}$.

Definition 1.8. 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 $\varphi \in FO(\mathcal{B})$ of vocabulary τ_s such that for any τ_s-structure $\mathbb{M}$:

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

Proposition 1.9. *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*

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

The proof of Proposition 1.9 is based on the fact that if φ is the formula which defines Q and $\psi_1(\bar{x}_1), \ldots, \psi_r(\bar{x}_r)$ are formulas, then

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

where the formula on the right is obtained by substituting every occurrence of $P_i(\bar{x}_i)$ in φ by $\psi_i(\bar{x}_i)$. It is worth noting that since $Q = Mod(\varphi)$, where $\varphi = Q\bar{x}_1, \ldots, \bar{x}_r\,(P_1(\bar{x}_1), \ldots, P_r(\bar{x}_r))$, Proposition 1.9 holds also to the other direction.

We shall prove that the version of Proposition 1.9, where Q is replaced with a second order quantifier $\mathcal{Q}$, holds. On the other hand, the converse of Proposition 1.9 turns out to be false in the second order case.

1.4 Second order generalized quantifiers

The notion of second order generalized quantifier was introduced in [1]. Let $t = (s_1, \ldots, s_w)$ where $s_i = (l^i_1, \ldots, l^i_{r_i})$ is a tuple of positive integers for $1 \leq i \leq w$. A *second order structure* of type t is a structure of the form $(M, P_1, \ldots, P_w)$, where $P_i \subseteq \mathcal{P}(M^{l^i_1}) \times \cdots \times \mathcal{P}(M^{l^i_{r_i}})$. In the following the notations $Str(t)$ and $Str(t, n)$ are used in the obvious way.

Definition 1.10. A *second order generalized quantifier* $\mathcal{Q}$ of type t is a class of structures of type t such that $\mathcal{Q}$ is closed under isomorphisms.

Example 1.11. Let us look at some examples of second order generalized quantifiers. Suppose $S \subseteq \mathbb{N}$.

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

The first example is the correspondent of the familiar unary second order existential quantifier. 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.

Definition 1.12. The extension $FO(\mathcal{Q})$ 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 $\bar{X}_i = (X_{1,i}, \ldots, X_{r_i,i})$ is a tuple of pairwise distinct predicate variables such that $ar(X_{j,i}) = l^i_j$ for $1 \leq j \leq r_i$, then

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

  is a formula.
- Satisfaction relation of FO is extended by the rule:

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

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

2 A notion of definability for second order generalized quantifiers

Before formulating the notion of definability studied in this paper, we need to define some auxiliary logics.

Definition 2.1. Let $t = (s_1, \ldots, s_w)$ be a type, where $s_i = (l^i_1, \ldots, l^i_{r_i})$ for $1 \leq i \leq w$, and let $\mathcal{G}_1, \ldots, \mathcal{G}_w$ be first order quantifier symbols of types $s_1, \ldots, s_w$.

- The syntax of the logic $FO(\mathcal{G}_1, \ldots, \mathcal{G}_w)$ is obtained by extending the syntax of FO by quantifiers $\mathcal{G}_1, \ldots, \mathcal{G}_w$.
- The models $\mathcal{M}$ of the logic $FO(\mathcal{G}_1, \ldots, \mathcal{G}_w)$ are of the form

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

  where $\mathbb{M}$ is a first order model and $G_i \subseteq \mathcal{P}(M^{l^i_1}) \times \cdots \times \mathcal{P}(M^{l^i_{r_i}})$.
- The satisfaction relation of FO is extended by the following rule for $1 \leq i \leq w$:

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

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

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

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

Proof. Induction on φ. $\square$

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

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

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

We emphasize the fact that the logic $FO(\mathcal{B}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$ has only an auxiliary role when considering definability of a quantifier $\mathcal{Q}$ in the logic $FO(\mathcal{B})$. However, studying definability in $FO(\mathcal{B})$ amounts to working with the logic $FO(\mathcal{B}, \mathcal{G}_1, \ldots, \mathcal{G}_w)$. The next example shows, analogously to the case with Lindström quantifiers, that a quantifier $\mathcal{Q}$ can always be defined using itself in a uniform way.

Example 2.4. 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 2.3 does have the desired properties.

Theorem 2.5. *Let* $\mathcal{Q}$ *be definable in* $FO(\mathcal{B})$ *with* $\varphi_{\mathcal{Q}}$. *Suppose* τ *is a vocabulary and* $\theta_1(\bar{P}_1), \ldots, \theta_w(\bar{P}_w)$ *are* τ*-formulas. Then there is* $\theta^* \in FO(\mathcal{B})$ *such that*

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

Proof. We may assume that θ_i does not contain any free variables for $1 \leq i \leq w$. Define a mapping f as follows:

$$\begin{array}{rcl} f(x = y) & = & x = y \\ f(\neg\varphi) & = & \neg f(\varphi) \\ f(\varphi \wedge \psi) & = & f(\varphi) \wedge f(\psi) \\ f(Q\bar{x}_1, \ldots, \bar{x}_r\,(\varphi_1, \ldots, \varphi_r)) & = & Q\bar{x}_1, \ldots, \bar{x}_r\,(f(\varphi_1), \ldots, f(\varphi_r)) \\ f(\mathcal{G}_i\bar{x}_1, \ldots, \bar{x}_{r_i}\,(\varphi_1, \ldots, \varphi_{r_i})) & = & \theta_i(P_1/f(\varphi_1), \ldots, P_{r_i}/f(\varphi_{r_i})) \end{array}$$

The bound variables of θ_i are replaced with new ones so that no free variable of $f(\varphi_i)$ becomes bound in the substitution.

Lemma 2.6. *Let $\mathbb{M}$ be a $\tau \setminus \bigcup_{i=1}^{w}\{P_{i,1}, \dots, P_{i,r_i}\}$-model and $G_i = \{\bar{Z} \mid (\mathbb{M}, \bar{Z}) \models \theta_i(\bar{P}_i)\}$. Then for all $\varphi(\bar{x}) \in FO(\mathcal{B}, \mathcal{G}_1, \dots, \mathcal{G}_w)[\emptyset]$ and $\bar{b} = b_1 \dots b_n \in M$:*

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

Proof. The claim holds for $\varphi = (x = y)$. The cases $\varphi = \neg\psi(\bar{x})$, $\varphi = Q\bar{x}_1, \dots, \bar{x}_r\,(\psi_1, \dots, \psi_r)$, and $\varphi = \psi \wedge \chi$ are trivial. Suppose that $\varphi = \mathcal{G}_i\bar{x}_1, \dots, \bar{x}_{r_i}\,(\psi_1, \dots, \psi_{r_i})(\bar{y})$. Then

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

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

$$\mathbb{M} \models \theta_i(P_1/f(\psi_1), \dots, P_{r_i}/f(\psi_{r_i}))(\bar{b}).$$

□

We claim now that

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

Assume $\mathbb{M}$ is a $\tau \setminus \bigcup_{i=1}^{w}\{P_{i,1}, \dots, P_{i,r_i}\}$-model and $G_i = \{\bar{Z} \mid (\mathbb{M}, \bar{Z}) \models \theta_i\}$ for $1 \le i \le w$. By the assumption that $\varphi_{\mathcal{Q}}$ defines $\mathcal{Q}$ and Lemma 2.6:

$$\begin{aligned} \mathbb{M} \models \mathcal{Q}P_1, \dots, P_w\,(\theta_1, \dots, \theta_w) \quad &\Leftrightarrow \quad (\mathbb{M}, G_1, \dots, G_w) \models \varphi_{\mathcal{Q}} \\ &\Leftrightarrow \quad \mathbb{M} \models f(\varphi_{\mathcal{Q}}). \end{aligned}$$

This completes the proof of Theorem 2.5. □

Theorem 2.5 implies the following.

Theorem 2.7. *Let $\mathcal{Q}$ be definable in $FO(\mathcal{B})$. Then*

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

Proof. Fix a vocabulary τ. By induction on $\varphi(\bar{x}, \bar{X}) \in FO(\mathcal{Q}, \mathcal{B})$ we prove that there is a formula $\varphi'(\bar{x}) \in FO(\mathcal{B})$ with an extra predicate P for every free predicate variable X in φ, such that for all $\mathbb{M}$, $\bar{a} = a_1 \dots a_n \in M$, and relations $\bar{R}$:

$$\mathbb{M} \models \varphi(\bar{a}, \bar{R}) \Leftrightarrow (\mathbb{M}, \bar{R}) \models \varphi'(\bar{a}).$$

If φ is atomic, and not of the form $X(\bar{y})$, we set $\varphi' = \varphi$. If $\varphi = X(\bar{y})$, then we set $\varphi' = P(\bar{y})$. Assume then that

$$\varphi = \mathcal{Q}\bar{X}_1, \dots, \bar{X}_w\,(\psi_1(\bar{X}_1), \dots, \psi_w(\bar{X}_w)).$$

By the induction assumption, the formula $\psi_i(\bar{X}_i)$ is equivalent to some $\psi_i' \in FO(\mathcal{B})$ with extra predicates. Let $\bar{P}_i$ be the predicates corresponding to $\bar{X}_i$ for $1 \leq i \leq w$. By Theorem 2.5, there is $\varphi' \in FO(\mathcal{B})$ such that for all $\mathbb{M}$, $\bar{a}$, and $\bar{R}$:

$$(\mathbb{M}, \bar{R}) \models \mathcal{Q}\bar{P}_1, \dots, \bar{P}_w\,(\psi_1', \dots, \psi_w')(\bar{a}) \Leftrightarrow (\mathbb{M}, \bar{R}) \models \varphi'(\bar{a}).$$

Therefore, for all $\mathbb{M}$, $\bar{a}$, and $\bar{R}$:

$$\mathbb{M} \models \varphi(\bar{a}, \bar{R}) \Leftrightarrow (\mathbb{M}, \bar{R}) \models \varphi'(\bar{a}).$$

□

From now on, we shall be mainly concerned with definability of second order quantifiers in the logic $FO(\mathcal{B})$, where $\mathcal{B}$ is the collection of all Lindström quantifiers of all types.

Definition 2.8. Let $\mathcal{B}$ be the collection of all Lindström quantifiers of all types. Denote the logic $FO(\mathcal{B})$ by $\mathcal{L}^*$. Also, $FO(\mathcal{B}, \mathcal{G}_1, \dots, \mathcal{G}_w)$ 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 2.7 implies the following.

Corollary 2.9. *Let $\mathcal{Q}$ be definable in the logic $\mathcal{L}^*$ with a sentence φ. Then*

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

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 2.4 we show how to treat quantifiers of arbitrary types.

2.1 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.

Example 2.10.

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

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

Example 2.11. Let $M = \{a_1, \ldots, a_n\}$ be a set and $G \subseteq \mathcal{P}(M)$. Choose variables $x_1, \ldots, x_n$ and, for $X = \{a_{i_1}, \ldots, a_{i_m}\} \subseteq M$, define

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

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

Example sentences ψ_1, ψ_2, ψ_3, and ψ_4 indicate that we need at least $n+1$ variables to define sets of size n. This is not entirely correct but still motivates the following definition.

Definition 2.12. 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 $G(n)$.

Theorem 2.13. *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. Define a formula φ and a Lindström quantifier Q as follows:

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

where ψ_1 is the formula

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

and ψ_2 is the formula

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

Finally, ψ_3 is the formula

$$(\mathcal{G}z\,(z \neq z) \to x \neq x)$$

and ψ_4 is the formula

$$(\neg\mathcal{G}z\,(z = z) \to y \neq y).$$

Let Q be defined as follows:

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

where R'_1 is

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

and R'_2 is

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

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

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

Lemma 2.14. *The class Q is closed under isomorphisms.*

Proof. Let $f\colon (M, R_1, \ldots, R_4) \cong (N, S_1, \ldots, S_4)$. 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

$$(M, R_1, \ldots, R_4) \in Q \Leftrightarrow (N, S_1, \ldots, S_4) \in Q.$$

□

By Lemma 2.14, the class Q is a Lindström quantifier. Hence, the logic $\mathcal{L}^*$ is closed under Q and $\varphi \in \mathcal{L}^*(\mathcal{G})$. Now φ defines $\mathcal{Q}$ since we have for any structure (M, G):

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

□

2.2 Proving undefinability

In the previous section we proved that using suitable Lindström quantifiers we can define all second order quantifiers $\mathcal{Q}$ 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 2.19). Therefore, we concentrate now on developing tools for proving undefinability. One simple property of $\mathcal{L}^*(\mathcal{G})$ which can be used to prove undefinability is that every formula of the logic $\mathcal{L}^*(\mathcal{G})$ contains finitely many variables.

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

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

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

Definition 2.15. Let $(\mathbb{M},G)$ be a structure and $a_1,\ldots,a_n \in M$. A relation $X \subseteq M^l$ is *definable with parameters* $a_1,\ldots,a_n$ if for some $\varphi(\bar{y},x_1,\ldots,x_n)$,

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

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

Definition 2.16. Let $(\mathbb{M},G)$ be a model and $a_1,\ldots,a_n \in M$. A relation $X \subseteq M^l$ is called $(a_1,\ldots,a_n)$-*invariant* if $f[X] = X$ for all $f\colon (\mathbb{M},G) \cong (\mathbb{M},G)$ such that $f(a_i) = a_i$ for $1 \le i \le n$.

Proposition 2.2 implies that all $(a_1,\ldots,a_n)$-definable relations are $(a_1,\ldots,a_n)$-invariant. Note, however, that there might not be any permutations fixing the collection G.

Example 2.17. 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 $f \neq id_M$.

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.

Lemma 2.18. *Let (M,G) be a structure such that G is permutation invariant. Then a set $X \subseteq M$ is $(a_1,\ldots,a_n)$-invariant if and only if $X \subseteq \{a_1,\ldots,a_n\}$ or $M \setminus X \subseteq \{a_1,\ldots,a_n\}$.*

The next example shows that Lemma 2.18 is false if G is not required to be permutation invariant.

Example 2.19. 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 and G' implying elementary equivalence up to k variables.

Lemma 2.20. *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 $\varphi(x_1, \ldots, x_n)$ that

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

for all $a_1, \ldots, a_n \in M$. We consider the case $\varphi = \mathcal{G}x\, \psi(x, \bar{y})$ only since other cases are trivial. Assume $(M, G) \models \varphi(a_1, \ldots, a_n)$. Then $X \in G$, where $X = \{c \in M \mid (M, G) \models \psi(c, a_1, \ldots, a_n)\}$. The formula φ contains at most $k + 1$ variables, hence $|\{a_1, \ldots, a_n\}| \leq k$. By Lemma 2.18, we have either $X \subseteq \{a_1, \ldots, a_n\}$ or $M \setminus X \subseteq \{a_1, \ldots, a_n\}$. Since $G(k) = G'(k)$, we have that $X \in G'$. By the induction hypothesis, $X = \{c \in M \mid (M, G') \models \psi(c, a_1, \ldots, a_n)\}$ and thus $(M, G') \models \varphi(a_1, \ldots, a_n)$. □

Lemma 2.20 can be used to show many quantifiers undefinable in the logic $\mathcal{L}^*$.

Theorem 2.21. *The following quantifiers are undefinable in $\mathcal{L}^*$:*

- $\{(M, G) \mid G$ *contains only singletons*$\}$
- $\exists^2$

Proof. We prove the claim for the quantifier $\exists^2$. For $k \in \mathbb{N} \setminus \{0\}$, we set $(M, G)_k = (\{1, \ldots, 3k\}, \emptyset)$ and $(M, G')_k = (\{1, \ldots, 3k\}, B_k)$, where $B_k = \{A \subseteq \{1, \ldots, 3k\} \mid |A| > k \text{ and } |M \setminus A| > k\}$. Suppose that φ defines $\exists^2$. Then

$$(M, G)_k \not\models \varphi \text{ and } (M, G')_k \models \varphi \tag{1}$$

for all $k \in \mathbb{N} \setminus \{0\}$. Let $r \in \mathbb{N}$ be the number of variables in φ. By Lemma 2.20, we have $(M, G)_{r-1} \equiv^r (M, G')_{r-1}$, and thus $(M, G)_{r-1} \models \varphi$ iff $(M, G')_{r-1} \models \varphi$ contradicting (1). □

There are quantifiers which cannot be shown to be undefinable using Lemma 2.20. 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 do not exist. Our next goal is to modify Lemma 2.20 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 2.19. 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.

Lemma 2.22. *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 $\varphi(\bar{x})$ that for all $f \in S_M$ and $\bar{a} = a_1 \ldots a_n \in M$,

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

We consider only the 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 2.18) we have either $X \subseteq \{a_1, \ldots, a_n\}$ or $(M \setminus X) \subseteq \{a_1, \ldots, a_n\}$. 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. □

Lemma 2.22 now allows us to prove the undefinability of the quantifier $Three$ in the logic $\mathcal{L}^*$.

Theorem 2.23. *The following quantifiers are undefinable in $\mathcal{L}^*$:*

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

Proof. Similar to the proof of Theorem 2.21. Let us consider the quantifier $Three$. For $k \geq 3$, we set $(M, G)_k = (\{1, \ldots, 3k\}, \emptyset)$ and $(M, G')_k = (\{1, \ldots, 3k\}, \{A_1, A_2, A_3\})$, where $A_i = \{1, \ldots, k+i\}$ for $1 \leq i \leq 3$. By the same arguments as in the proof of Theorem 2.21 and Lemma 2.22, the quantifier $Three$ is undefinable in the logic $\mathcal{L}^*$. □

However, there are intuitively undefinable quantifiers which cannot be shown to be undefinable in the logic $\mathcal{L}^*$ using Lemma 2.22. Let k be a positive integer. The following quantifier is an obvious example.

$$\{(M, G) \mid G \text{ contains exactly } k \text{ set of size } n \text{ for } 1 \leq n \leq |M| - 1\}.$$

2.3 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 2.19. 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 2.13. The reason is that φ contains nesting of the quantifier $\mathcal{G}$. We shall next show that the characterization of Theorem 2.13 is complete with respect to sentences not containing nesting of $\mathcal{G}$.

Definition 2.24. A formula ψ is *flat* if it does not contain nesting of $\mathcal{G}$. The collection of all flat formulas of $\mathcal{L}^*(\mathcal{G})$ is denoted by $\mathcal{F}$. The collection of flat formulas with at most k variables is denoted by $\mathcal{F}_k$.

Lemma 2.25. *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 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 2.18, 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 2.25 implies that Theorem 2.13 characterizes definable second order quantifiers completely with respect to sentences $\varphi \in \mathcal{F}$.

Theorem 2.26. *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 2.13.*

Proof. The defining sentence of $\mathcal{Q}$ constructed in the proof of Theorem 2.13 is flat. On the other hand, by Lemma 2.25, a quantifier $\mathcal{Q}$ defined by a flat sentence φ satisfies the assumption of Theorem 2.13 with $r - 1$, where r is the number of variables appearing in φ. $\square$

The task of characterizing $\mathcal{L}^*$-definable quantifiers without any restricting assumptions seems difficult.

2.4 Quantifiers of arbitrary types

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

Let $(M, G_1, \ldots, G_w)$ be a t-structure and $n \in \mathbb{N}$. Denote the set $\{(X_1, \ldots, X_{r_i}) \in G_i : |X_j| < n \text{ or } |M \setminus X_j| < n \text{ for } 1 \le j \le r_i\}$ by $G_i(n)$. Now, Lemma 2.20 can be reformulated as follows.

Lemma 2.27. *Let $(M, G_1, \ldots, G_w)$ and $(M, G'_1, \ldots, G'_w)$ be structures of type t such that G_i and G'_i are permutation invariant and $G_i(k) = G'_i(k)$ for $1 \le i \le w$. Then*

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

Again, the assumption of permutation invariance can be avoided (cf. Lemma 2.22).

Lemma 2.28. *Let $(M, G_1, \ldots, G_w)$ and $(M, G'_1, \ldots, G'_w)$ 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 $1 \le i \le w$. Then*

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

It is also straightforward to prove versions of Theorems 2.13 and 2.26 in this context.

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. Elementary equivalence up to $k+1$ variables with respect to $\mathcal{L}^*(\mathcal{G})$, where the type of $\mathcal{G}$ is (n), follows if G and G' contain the same relations which can be defined with k parameters in the empty vocabulary. Since G and G' are permutation invariant, we have that $(a_1, \ldots, a_n)$-definable relations are fixed by all permutations of M which fix the parameters $a_1, \ldots, a_n$. However, the picture is not so simple as with unary relations. In particular, there are difficulties with describing the collection $G(n)$ since it is supposed to include all relations which are $(a_1, \ldots, a_n)$-invariant for some $(a_1, \ldots, a_n) \in M^n$.

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

Proposition 2.29. *The quantifier $\exists^2_n$ is undefinable in the logic $\mathcal{L}^*$.*

Proof. For $k \in \mathbb{N} \setminus \{0\}$ set $(M, G)_k = (\{1, \ldots, 3k\}, \emptyset)$ and $(M, G')_k = (\{1, \ldots, 3k\}, B_k)$, where

$$B_k = \{\{(c, \ldots, c) \mid c \in A\} \mid A \subseteq \{1, \ldots, 3k\},\ |A| > k \text{ and } |M \setminus A| > k\}.$$

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

3 Further undefinability 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 $Q_1, \ldots, Q_n$ such that

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

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 $FO(Q_1, \ldots, Q_n)$ but is still undefinable in $FO(Q_1, \ldots, Q_n)$.

The construction of the quantifier $\mathcal{Q}_0$ in Theorem 3.1 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, Lemma 3.3 is proved in the same way as in [1].

Theorem 3.1. *There is a quantifier $\mathcal{Q}_0$ of type $((1))$ 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 $\mathcal{Q} \subseteq \mathcal{Q}_0$,*

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

Proof. 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 $\varphi_0, \ldots, \varphi_{m-1}$ without 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 $h\colon \mathbb{N} \to \mathbb{N}$ such that if $|M| \geq h(m)$, then the number of collections $G \subseteq \mathcal{P}(M)$ which are $\varphi_{<m}$-undefinable over M is greater $\frac{1}{2}b_{|M|}(t)$, where $t = ((1))$ and $b_n(t)$ 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 of cardinality n is at most $2^{p_{\tau_{<m}}(n)}$. Over any structure $\mathbb{M}$, sentences $\varphi_0, \dots, \varphi_{m-1}$ define 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

$$\{(M, G) \mid (\forall m \in \mathbb{N})(h(m) \leq |M| \Rightarrow G \text{ is } \varphi_{<m}\text{-undefinable over } M\}.$$

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 $n \in \mathbb{N} \setminus \{0\}$.

Lemma 3.2. *The class $\mathcal{Q}_0$ is closed under isomorphisms.*

Proof. Let M and N be sets, $f\colon M \to N$ a bijection, and $G \subseteq \mathcal{P}(M)$. Assume also that 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. Let $\mathbb{N}$ be such that $f\colon \mathbb{M} \cong \mathbb{N}$. Then

$$f[[G]] = \{B \mid (\mathbb{N}, B) \models \varphi_i\}.$$

This holds since for all $B \subseteq M$, $f\colon (\mathbb{M}, B) \cong (\mathbb{N}, f[B])$, and also $f^{-1}\colon (\mathbb{N}, B) \cong (\mathbb{M}, f^{-1}[B])$ for all $B \subseteq N$. From the above, it is easy to see that $\mathcal{Q}_0$ is closed under isomorphisms. $\square$

Lemma 3.3. *Let $\mathcal{Q} \subseteq \mathcal{Q}_0$ be a quantifier. Then $FO \equiv FO(\mathcal{Q})$.*

Proof. The claim is proved using induction on the number of occurrences of $\mathcal{Q}$ in a formula φ. We prove that any formula of $FO(\mathcal{Q})$ is equivalent to a first order formula over the same vocabulary. Let φ be a formula with $n+1$ many 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 $\tau_\psi \cup \{X, \bar{x}\}$ 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 $|M| > h(m+1)$ then $\mathbb{M} \not\models \mathcal{Q}X\, \psi'(X, \bar{a})$. There are only finitely many different isomorphism types of structures of cardinality less than $h(m+1)$. In structures of cardinality greater than $h(m+1)$, the formula $\mathcal{Q}X\, \psi'(X, \bar{x})$ is equivalent to the formula $\exists x(x \neq x)$. Therefore, φ is equivalent to a first order formula. □

This completes the proof of Theorem 3.1. □

Theorem 3.4. *Let $\mathcal{B} = \{Q_i\}_{i \in \mathbb{N}}$ be a collection of Lindström quantifiers. Then there is a quantifier $\mathcal{Q}$ of type $((1))$ such that $FO(\mathcal{B}) \equiv FO(\mathcal{B}, \mathcal{Q})$ but $\mathcal{Q}$ is not definable in the logic $FO(\mathcal{B})$.*

Proof. As Andersson [1] points out, Theorem 3.1 can be generalized to the case where FO is replaced with $FO(\mathcal{B})$. Let $\mathcal{Q}_0$ denote the quantifier given by this generalization of Theorem 3.1. Then, for all $S \subseteq \mathbb{N}$

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

where $\mathcal{Q}_S = \{(M, G) \mid (M, G) \in \mathcal{Q}_0 \text{ and } |M| \in S\}$. On the other hand, the logic $FO(\mathcal{B}, \mathcal{G})$ is countable. Thus, $\mathcal{Q}_S$ is undefinable in $FO(\mathcal{B})$ for some $S \subseteq \mathbb{N}$. □

Note that the quantifier $\mathcal{Q}_S$ is definable in the logic $FO(\mathcal{B}, \mathcal{Q}_S)$ with the sentence $\mathcal{Q}_S X\, (\mathcal{G}x\, (X(x)))$. Theorem 3.4 can be improved to yield an in-expressive quantifier $\mathcal{Q}$ which is not definable in the logic $\mathcal{L}^*$. The cardinality argument used in the proof of Theorem 3.4 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 [2].

Definition 3.5. Let φ be a formula and let $Q_1, \ldots, Q_n$ be the Lindström quantifiers appearing in φ. Then we write $\varphi(Q_1, \ldots, Q_n)$. If $Q'_1, \ldots, Q'_n$ 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 $\varphi(Q'_1, \ldots, Q'_n)$ denotes the formula obtained from φ by changing Q_i everywhere to Q'_i. Formulas obtained from each others by a change of quantifiers are called *similar*.

Theorem 3.6. *There is a quantifier $\mathcal{Q}$ of type $((1))$ with the following properties:*

1. *$FO(\mathcal{Q}) \equiv FO$.*
2. *The quantifier $\mathcal{Q}$ is undefinable in the logic $\mathcal{L}^*$.*

Proof. Let $\mathcal{Q}_0$ be as in Theorem 3.1. 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 $n_0, \ldots, n_{j-1}$ and the classes $K_0, \ldots, K_{j-1}$ have been defined already. Suppose $(Q_1, \ldots, Q_y)$ are the Lindström quantifiers appearing in ψ_j. The type of Q_i is denoted by s_i. Let $(Q'_1, \ldots, Q'_y)$ be Lindström quantifiers such that the type of Q'_i is s_i for $1 \le i \le y$. If

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

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

$$Mod(\psi_j(Q_1, \ldots, Q_y)) \cap Str(t, n) = Mod(\psi_j(Q'_1, \ldots, Q'_y)) \cap Str(t, n),$$

where $t = ((1))$. The number of τ_{s_i}-structures with universe $\{1, \ldots, n\}$ 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

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

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

$$Mod(\psi_j(Q_1, \ldots, Q_y)) \cap Str(t, n_j),$$

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

Lemma 3.7. *The quantifier $\mathcal{Q}$ is not definable in the logic $\mathcal{L}^*$.*

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

This completes the proof of Theorem 3.6. □

Acknowledgments

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

References

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

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

[3] P. Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966.

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