



University of Pennsylvania
ScholarlyCommons

Technical Reports (CIS)

Department of Computer & Information Science

November 1992

Generalized Quantifiers and Logical Reducibilities

Anuj Dawar
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/cis_reports

Recommended Citation

Anuj Dawar, "Generalized Quantifiers and Logical Reducibilities", . November 1992.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-92-85.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/cis_reports/348
For more information, please contact repository@pobox.upenn.edu.

Generalized Quantifiers and Logical Reducibilities

Abstract

We consider extensions of first order logic (FO) and least fixed point logic (LFP) with generalized quantifiers in the sense of Lindström [Lin66]. We show that adding a finite set of such quantifiers to LFP fails to capture all polynomial time properties of structures, even over a fixed signature. We show that this strengthens results in [Hel92] and [KV92a]. We also consider certain regular infinite sets of Lindström quantifiers, which correspond to a natural notion of logical reducibility. We show that if there is any recursively enumerable set of quantifiers that can be added to FO (or LFP) to capture P, then there is one with strong uniformity conditions. This is established through a general result, linking the existence of complete problems for complexity classes with respect to the first order translations of [Imm87] or the elementary reductions of [LG77] with the existence of recursive index sets for these classes.

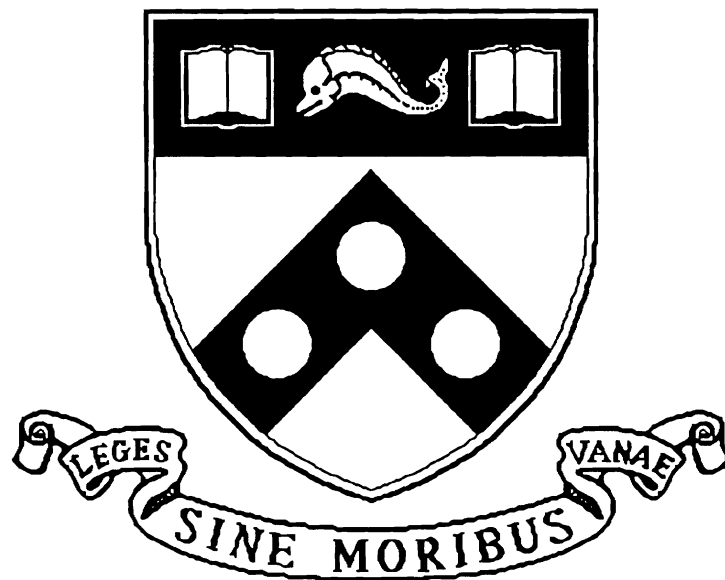
Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-92-85.

Generalized Quantifiers and Logical Reducibilities

MS-CIS-92-85
LINC LAB 240

Anuj Dawar



University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

December 1992

Generalized Quantifiers and Logical Reducibilities¹

Anuj Dawar²

Department of Computer and Information Science

University of Pennsylvania

Philadelphia, PA 19104.

November 30, 1992

¹Research supported in part ARO grant DAAL 03-89-C-0031.

²I am extremely grateful to Scott Weinstein and Steven Lindell for valuable discussions on the material presented in this paper.

Abstract

We consider extensions of first order logic (FO) and least fixed point logic (LFP) with generalized quantifiers in the sense of Lindström [Lin66]. We show that adding a finite set of such quantifiers to LFP fails to capture all polynomial time properties of structures, even over a fixed signature. We show that this strengthens results in [Hel92] and [KV92a]. We also consider certain regular infinite sets of Lindström quantifiers, which correspond to a natural notion of logical reducibility. We show that if there is any recursively enumerable set of quantifiers that can be added to FO (or LFP) to capture P , then there is one with strong uniformity conditions. This is established through a general result, linking the existence of complete problems for complexity classes with respect to the first order translations of [Imm87] or the elementary reductions of [LG77] with the existence of recursive index sets for these classes.

1 Introduction

Computational complexity measures the complexity of a problem in terms of the resources, such as time, space, or hardware, required to solve the problem relative to a given machine model of computation. In contrast, descriptive complexity analyzes the complexity of a problem in terms of the logical resources, such as number of variables, kinds of quantifier, or length of formula (even through infinitary formulas), required to define the problem. An interesting outcome of investigations in descriptive complexity has been the discovery of a close connection between descriptive and computational complexity. Fagin [Fag74] showed that the classes of finite relational structures definable in existential second order logic are exactly those classes that are in the class NP. Immerman [Imm86] and Vardi [Var82] showed that in the presence of a linear order on the domain of every structure, the classes of finite structures in P are exactly those that are definable in the extension of first-order logic with a least-fixed-point operation (LFP). Similar results have been obtained for a variety of other complexity classes (see, for instance, [Imm89]).

However, for every complexity class below NP, the known characterizations of the class in descriptive terms rely on the presence of a linear order on the domains of the structures. For instance, there is no known logical characterization of the collection of classes of finite structures that are recognizable in polynomial time. Indeed, it is an open question whether this collection has a recursively enumerable index set. Since it is known that relational structures over an arbitrary signature can be encoded as graphs (by an encoding that is first-order definable, see, for instance, [Lin87]), this question is equivalent to asking if there is a recursive enumeration of the polynomial time recognizable properties of graphs.

The logic LFP proves too weak to express all the polynomial time recognizable properties of finite structures that are not necessarily ordered. One approach to increasing the expressive power of this logic is to add to the language generalized quantifiers, in the sense of Lindström [Lin66]. Associated with each quantifier is its arity n . Recently, Hella [Hel92] has shown that for any set \mathbf{Q} of generalized quantifiers whose arities are bounded by n , there is a polynomial time recognizable class of finite structures C_n that is not expressible in $\text{LFP}(\mathbf{Q})$ – the extension of LFP with all the quantifiers in the set \mathbf{Q} . One important consequence of this result is that there is no finite set of generalized quantifiers that can be added to LFP to yield a logic that captures P. The class C_n constructed by Hella is in

a signature that contains a relation of arity $n + 1$. If we confine ourselves to structures over a fixed signature, such as the language of graphs, the result vanishes. Indeed, there is a collection \mathbf{Q} of *binary* quantifiers such that $\text{LFP}(\mathbf{Q})$ expresses all the polynomial time properties of graphs. This leaves open the question of whether there is a finite set with this property.

In one of the results in this paper, we provide a negative answer to this last question. In Section 3, we show that for any fixed signature σ and any finite collection \mathbf{Q} of generalized quantifiers, there is a polynomial time recognizable set of finite structures (or even a log-space recognizable set) over σ that is not expressible in $L_{\infty\omega}^\omega(\mathbf{Q})$. The logic $L_{\infty\omega}^\omega$, infinitary logic with a bounded number of variables, is a powerful extension of LFP. This result is proved by showing that the properties of complete graphs (or, more generally, of complete structures over any signature) that are expressible in $L_{\infty\omega}^\omega(\mathbf{Q})$ are eventually finitely many. This also strengthens a result in [KV92a].

In Section 4 we define certain kinds of infinite sets of generalized quantifiers of unbounded arity, with a strong uniformity condition. These *uniform sequences* of generalized quantifiers correspond to a natural notion of logical reducibility. We establish that there is such a uniform sequence \mathbf{Q} such that LFP (or FO) enriched with the quantifiers in \mathbf{Q} expresses all the properties of structures that are computable in polynomial time (or logarithmic space), if and only if, there is a property that is complete for P (respectively, L) with respect to first order definable reductions. Moreover, this occurs, if and only if, the properties in P are recursively enumerable. We show, thus, that if there is any recursively enumerable set of generalized quantifiers which can be added to FO to capture exactly P, there is a uniform sequence with this property.

2 Background

A signature $\sigma = \langle R_1 \dots R_m \rangle$ is a finite sequence of relation symbols, R_i , each with an associated arity n_i . A structure $\mathfrak{A} = \langle A, R_1^\mathfrak{A} \dots R_m^\mathfrak{A} \rangle$ over signature σ , consists of a universe, A , and relations $R_i^\mathfrak{A} \subseteq A^{n_i}$ interpreting the relation symbols in σ . Unless otherwise stated, we will assume that the universe of every structure considered is finite.

We will write FO, LFP, *etc.* both to denote logics (*i.e.*, sets of formulas) and the collections of classes of finite structures that are expressible in the respective logics. By a class

of structures, we mean a collection of structures that is closed under isomorphisms of the structures. We also use L to denote a logic, in general, whereby we mean an extension of first order logic that satisfies reasonable closure properties. It suffices, for instance, if L is a regular logic in the sense of [Ebb85].

2.1 Logics for Complexity Classes

We say that a logic L captures a complexity class \mathcal{C} if every class of structures definable in L is in \mathcal{C} , and *vice versa*. Fagin proved the following result:

Theorem 1 ([Fag74]) *Existential second order logic (Σ_1^1) captures NP.*

Definition 1 *A complexity class \mathcal{C} is said to be recursively indexable, if there is a recursive set \mathcal{I} and a Turing machine M such that: on input $i \in \mathcal{I}$, M produces the code for a machine $M(i)$; the class of structures C accepted by $M(i)$ is a class in \mathcal{C} and $M(i)$ witnesses the membership of C in \mathcal{C} (i.e., $M(i)$ runs within the complexity bounds defining \mathcal{C} ; and for each class of structures $C \in \mathcal{C}$, there is an $i \in \mathcal{I}$ such that $M(i)$ accepts C .*

Gurevich [Gur88] defines the notion of a complexity class having a logic that captures it. This notion is essentially the same as the definition of a class being recursively indexable, given above. Recursive index sets for a complexity class may be generated by effective listings of formulas or machines. For instance, Theorem 1 provides an indexing of NP. It is an open question whether any class below NP is recursively indexable. Thus, while the classes of linearly ordered structures in P are known to be recursively enumerable (see Theorem 2 below), it is not known if the properties of graphs in P are.

We will need the following definition in Theorem 7 below.

Definition 2 *A complexity class \mathcal{C} is defined by a machine model, some resource R (such as space or time) and a family of functions T . A problem is in \mathcal{C} if it is recognized by a machine whose use of resource R on inputs of size n is bounded by $t(n)$ for some function $t \in T$. We say that \mathcal{C} is bounded just in case there is a function $s \in T$ such that for every $t \in T$, there is a k such that $t(n)$ is eventually bounded by $s(n^k)$.*

We will also assume that the function $s(n)$ in this definition is at least n (for sequential time) and $\log n$ (for space). The complexity classes L, NL, P, NP, PSPACE, *etc.* are all bounded, under this definition.

2.2 Least Fixed Point Logic

Let $\phi(R, x_1, \dots, x_k)$ be a first-order formula. On a structure \mathfrak{A} ϕ defines the operator, $\Phi_{\mathfrak{A}}(R^{\mathfrak{A}}) = \{\langle a_1, \dots, a_k \rangle \mid \langle \mathfrak{A}, R^{\mathfrak{A}} \rangle \models \phi[a_1, \dots, a_k]\}$. If ϕ is an R -positive formula, $\Phi_{\mathfrak{A}}$ is monotone. We may view ϕ as determining an induction on \mathfrak{A} the stages of which are defined as follows: $\phi_{\mathfrak{A}}^0 = \emptyset$; $\phi_{\mathfrak{A}}^{m+1} = \Phi_{\mathfrak{A}}(\phi_{\mathfrak{A}}^m)$. The *closure ordinal* of ϕ on \mathfrak{A} , denoted $\|\phi\|_{\mathfrak{A}}$, is the least m such that $\phi_{\mathfrak{A}}^m = \phi_{\mathfrak{A}}^{m+1}$. The m^{th} stage of the induction determined by ϕ can be uniformly defined over all structures by a first-order formula which we denote by ϕ^m . The set inductively defined by ϕ on \mathfrak{A} , denoted $\phi_{\mathfrak{A}}^{\infty}$, is the least fixed point of the operator $\Phi_{\mathfrak{A}}$, that is, $\phi_{\mathfrak{A}}^{\infty} = \phi_{\mathfrak{A}}^m$, where $m = \|\phi\|_{\mathfrak{A}}$.

We write LFP for the extension of first-order logic with the **lfp** operation which uniformly determines the least fixed point of an R -positive formula. That is, for any R -positive formula ϕ , $\text{lfp}(R, x_1, \dots, x_k)\phi$ is a formula of LFP and $\mathfrak{A} \models \text{lfp}(R, x_1, \dots, x_k)\phi[a_1, \dots, a_k]$, if and only if, $\langle a_1, \dots, a_k \rangle \in \phi_{\mathfrak{A}}^{\infty}$.

Immerman [Imm86] and Vardi [Var82] independently showed that when we include a total ordering on the domain as part of the logical vocabulary, the language LFP expresses exactly the class of polynomial time computable properties.

Theorem 2 ([Imm86],[Var82]) *LFP with ordering = P.*

2.3 Element Types

Let L^k be the fragment of first-order logic which consists of those formulas whose variables, both free and bound, are among x_1, \dots, x_k . Let $L_{\infty\omega}^k$ be the closure of L^k under the operations of conjunction and disjunction applied to arbitrary (finite or infinite) sets of formulas. Let $L_{\infty\omega}^{\omega} = \bigcup_{k \in \omega} L_{\infty\omega}^k$. The logic $L_{\infty\omega}^{\omega}$ was introduced in [Bar77]. Rubín [Rub75] showed that LFP is a fragment of this logic. Recently, $L_{\infty\omega}^{\omega}$ has been extensively studied in the context of finite models [KV92c, KV92b, DLW93].

Recall that for a structure \mathfrak{A} and a tuple $s = \langle a_1, \dots, a_l \rangle$ of elements of \mathfrak{A} , the first-order type of s in \mathfrak{A} , denoted $\text{Type}(\mathfrak{A}, s)$ is the set of formulas, ϕ , with free variables among x_1, \dots, x_l , such that $\mathfrak{A} \models \phi[a_1 \dots a_l]$. In [DLW93], we introduce the following variation of this notion.

Definition 3 Let \mathfrak{A} be a structure and let $l \leq k$ be natural numbers. For any sequence $s = \langle a_1, \dots, a_l \rangle$ of elements of \mathfrak{A} , the L^k -type of s in \mathfrak{A} , denoted $\text{Type}_k(\mathfrak{A}, s)$, is the set of formulas, $\phi \in L^k$ with free variables among x_1, \dots, x_l , such that $\mathfrak{A} \models \phi[a_1 \dots a_l]$. τ is an L^k -type, if and only if, it is the L^k -type of some tuple in some (finite or infinite) structure. If τ is an L^k -type we say that the tuple s realizes τ in \mathfrak{A} , if and only if, $\tau = \text{Type}_k(\mathfrak{A}, s)$.

We write $\langle \mathfrak{A}, s \rangle \equiv_k \langle \mathfrak{B}, t \rangle$ for $\text{Type}_k(\mathfrak{A}, s) = \text{Type}_k(\mathfrak{B}, t)$. The equivalence relation \equiv_k has an elegant characterization in terms of Ehrenfeucht-Fraïssé pebble games [Imm82, Poi82].

Definition 4 A class of structures, C , is k -compact, if and only if, the set of all L^k -types that are realized in structures in C is finite.

Theorem 3 ([DLW92]) If C is k -compact, then every formula of $L_{\infty\omega}^k$ is equivalent, over C , to a formula of L^k .

We will also need the following definition:

Definition 5 Given a structure \mathfrak{A} and a tuple of elements \bar{a} from \mathfrak{A} , the basic equality type of $\langle \mathfrak{A}, \bar{a} \rangle$ is the (unique up to equivalence) quantifier free formula ϕ , with no non-logical vocabulary, such that $\mathfrak{A} \models \phi[\bar{a}]$ and for every quantifier free formula ψ , with no non-logical vocabulary, exactly one of $\phi \models \psi$ or $\phi \models \neg\psi$ holds. We write $\bar{a} \simeq_k \bar{b}$ to denote that \bar{a} and \bar{b} are k -tuples of the same basic equality type.

Note that the number of distinct basic equality types of k -tuples in a structure of size k or greater depends only on k and not on the particular structure.

2.4 Generalized Quantifiers

Let C be any collection of structures over the signature $\sigma = \langle R_1 \dots R_m \rangle$ (where R_i has arity n_i) that is closed under isomorphism, *i.e.*, if $\mathfrak{A} \cong \mathfrak{B}$ then $\mathfrak{A} \in C$ if and only if $\mathfrak{B} \in C$. We associate with C the generalized quantifier Q_C . For a logic L , define the extension $L(Q_C)$ by closing the set of formulas of L under the following formula formation rule: if $\phi_1 \dots \phi_m$ are formulas of $L(Q_C)$ and $\bar{x}_1 \dots \bar{x}_m$ are tuples of variables with the length of \bar{x}_i being n_i , then $Q_C \bar{x}_1 \dots \bar{x}_m (\phi_1 \dots \phi_m)$ is a formula of $L(Q_C)$ with the variables in $\bar{x}_1 \dots \bar{x}_m$ bound. The semantics of the quantifier is given by the following rule: $\mathfrak{A} \models Q_C \bar{x}_1 \dots \bar{x}_m (\phi_1 \dots \phi_m)$ if and only if $\langle A, \phi_1^{\mathfrak{A}} \dots \phi_m^{\mathfrak{A}} \rangle \in C$, where A is the universe of \mathfrak{A} and $\phi_i^{\mathfrak{A}} = \{\bar{a} \mid \mathfrak{A} \models \phi_i[\bar{a}]\}$.

Example 1

1. The existential quantifier (\exists) can be defined as the generalized quantifier associated with the class of structures C over the signature with one unary relation symbol R given by $C = \{\langle A, R^A \rangle \mid R^A \text{ is not empty}\}$.
2. The universal quantifier (\forall) is the generalized quantifier associated with the class $C = \{\langle A, A \rangle\}$.
3. The Hartig (or equicardinality) quantifier is given by the class $C = \{\langle A, S_1, S_2 \rangle \mid S_1, S_2 \subseteq A \text{ and } |S_1| = |S_2|\}$.
4. The Rescher (or majority) quantifier is given by the class $C = \{\langle A, S_1, S_2 \rangle \mid S_1, S_2 \subseteq A \text{ and } |S_1| \geq |S_2|\}$.
5. The unary counting quantifiers are those associated with the classes $C_i = \{\langle A, S \rangle \mid S \subseteq A \text{ and } |S| \geq i\}$, for each $i \in \omega$.

For a quantifier Q associated with a class of structures over the signature $\langle R_1 \dots R_m \rangle$, define the arity of Q to be $\max(n_1, \dots, n_m)$, where n_i is the arity of R_i . Hella [Hel92] has established the following result:

Theorem 4 ([Hel92]) *Given any set \mathbf{Q} of generalized quantifiers of bounded arity, there is a signature σ and a polynomial time recognizable class of structures C of signature σ that is not definable in $\text{LFP}(\mathbf{Q})$.*

It follows immediately that the addition of a finite number of generalized quantifiers to LFP will not allow us to express all classes of structures recognizable in polynomial time:

Corollary 1 ([Hel92]) *If \mathbf{Q} is a finite set of generalized quantifiers, there is a signature σ and a polynomial time recognizable class of structures C of signature σ that is not definable in $\text{LFP}(\mathbf{Q})$.*

Note that in Theorem 4, the signature σ depends on \mathbf{Q} . In particular, σ must contain a relation symbol of arity greater than the bound on the arities of the quantifiers in \mathbf{Q} . If we only consider classes of structures over a fixed signature, Theorem 4 fails. Consider, for instance, graphs, *i.e.*, structures over the signature with one binary relation. If we add to

LFP a quantifier for each polynomial time property of graphs, each of these properties is then trivially definable. Moreover, all the quantifiers have arity 2.

While Theorem 4 fails when we fix the signature, it is still possible to establish Corollary 1 about finite sets of generalized quantifiers. This is what we show in the next section.

In what follows, we will generally not distinguish between a generalized quantifier and the class of structures with which it is associated, where this will not result in any confusion.

3 Finitely Many Quantifiers

In this section, we establish our first main result, *i.e.*, for every signature σ , and any finite collection of generalized quantifiers \mathbf{Q} there is a polynomial time recognizable property of structures over signature σ that is not expressible in $L_{\infty\omega}^\omega(\mathbf{Q})$. We begin by showing that it suffices to consider the case where \mathbf{Q} consists of a single quantifier and that quantifier is associated with a class of structures C over a signature with just one relation R .

Lemma 1 *For every finite collection \mathbf{Q} of generalized quantifiers and any logic L , there is a single quantifier Q such that every property expressible in $L(\mathbf{Q})$ is expressible in $L(Q)$.*

Proof: Let $\mathbf{Q} = \{Q_1, \dots, Q_n\}$ We can assume, without loss of generality, that all the quantifiers in \mathbf{Q} are over the same signature σ . If this is not the case, let σ be the union of the signatures of the Q_i and let Q'_i be the set of all possible expansions of structures in Q_i to the signature σ . It is easily verified that $L(Q'_1, \dots, Q'_n)$ expresses the same properties as $L(\mathbf{Q})$.

Let k be a natural number such that the number of basic equality types of k -tuples (in a structure of size at least k) is greater than $\log_2 n$. Let ϕ_1, \dots, ϕ_r be an enumeration of these basic equality types, and let g be a fixed *one-to-one* map $g : \{1, \dots, n\} \rightarrow \mathcal{P}(\{\phi_1, \dots, \phi_r\})$. This map is used to encode the indices of the quantifiers as subsets of $\{\phi_1, \dots, \phi_r\}$.

Let U be a new k -ary relation symbol and Q a quantifier over the signature $\sigma \cup \{U\}$ such that $\mathfrak{A} \in Q$ if and only if:

1. $U^{\mathfrak{A}}$ is closed under the equivalence relation \simeq_k (thus, $U^{\mathfrak{A}}$ can be identified with a subset τ of $\{\phi_1, \dots, \phi_r\}$);
2. $\tau = g(i)$ for some i ; and

3. the reduct of \mathfrak{A} to the signature σ is in Q_i .

We are now in a position to translate any formula of $L(\mathbf{Q})$ into an equivalent formula of $L(Q)$. We proceed by induction on the structure of the formula. Only one case is of interest: consider the formula $Q_i \bar{x}(\psi_1, \dots, \psi_r)$. First, observe that there can be only finitely many structures of size less than k in Q_i . We can therefore write a *first order* formula α which says “there are fewer than k elements in the universe A , and the structure $\langle A, \psi_1^A, \dots, \psi_r^A \rangle$ is in Q_i .” Secondly, let β be the first order sentence that says “there are at least k elements in the universe.” Then, the required translation is $\alpha \vee (\beta \wedge Q_i \bar{x}(\bigvee_{\phi \in g(i)} \phi, \psi'_1, \dots, \psi'_r))$, where $\psi'_j (1 \leq j \leq r)$ is the translation of ψ_j obtained by induction hypothesis. ■

Lemma 2 *For any generalized quantifier Q_C , there is a quantifier $Q_{C'}$ associated with a class, C' , of structures over a signature with only one relation such that for any logic L , $L(Q_C) = L(Q_{C'})$.*

Proof: If C is a class of structures over the signature $\langle R_1 \dots R_m \rangle$, let R be a relation symbol of arity $n_1 + \dots + n_m$, where n_i is the arity of R_i . Let $C' = \{ \langle A, R_1 \times \dots \times R_m \rangle \mid \langle A, R_1 \dots R_m \rangle \in C \}$. ■

Fix a signature σ , and let \mathcal{K}^σ be the class of *complete* structures over σ , *i.e.* for every relation symbol R (of arity a) in σ and every structure $\mathfrak{A} \in \mathcal{K}^\sigma$, $R^{\mathfrak{A}} = A^a$, where A is the universe of \mathfrak{A} . We will write K_n^σ for the unique (up to isomorphism) structure of size n in \mathcal{K}^σ (or just K_n and \mathcal{K} , when σ is understood). This notion is a direct generalization of the notion of a complete graph. Note that the class \mathcal{K} is first-order definable.

Each isomorphism-closed sub-class, C , of \mathcal{K} is determined by a set of natural numbers S such that $C = \{K_n \mid n \in S\}$. The polynomial time recognizable sub-classes of \mathcal{K} are just those corresponding to polynomial time recognizable sets of tally natural numbers. Among these are the sets $\mu_i = \{i \cdot n \mid n \in \omega\}$ (multiples of i), for every $i \in \omega$.

The class \mathcal{K} is k -compact for all k , and therefore, by Theorem 3, $L_{\infty\omega}^\omega$ collapses to FO on this class. A simple induction on the structure of formulas shows that this remains true when a finite collection of generalized quantifiers is added:

Lemma 3 *For any finite collection of generalized quantifiers, \mathbf{Q} , any formula of $L_{\infty\omega}^\omega(\mathbf{Q})$ is equivalent, over the class of complete structures to a formula of FO(\mathbf{Q}).*

Thus to show that a sub-class of \mathcal{K} is not definable in $L_{\infty\omega}^\omega(\mathbf{Q})$, we need only show that it is not in $\text{FO}(\mathbf{Q})$.

Fix a generalized quantifier Q and let $\phi_0, \dots, \phi_i, \dots$ be an enumeration of the sentences of $\text{FO}(Q)$. For each i , let $f_i = \{n \mid K_n \models \phi_i\}$. The next result shows that there are eventually only finitely many distinct sets f_i :

Theorem 5 *For any finite collection, \mathbf{Q} , of generalized quantifiers there is a finite collection, $q_0, \dots, q_t \subseteq \omega$ such that for every i , there is a j and an n_0 such that for all $n > n_0$, $n \in f_i$ if and only if $n \in q_j$.*

Proof: As we have seen, we can assume, without loss of generality that \mathbf{Q} consists of a single quantifier Q which is associated with a class of structures over a signature containing just one relation symbol, R . Let k be the arity of R .

For any structure \mathfrak{A} with universe A and any $\text{FO}(\mathbf{Q})$ formula $\phi(x_1 \dots x_k)$ in the signature of \mathfrak{A} , the relation $\phi^{\mathfrak{A}} \subseteq A^k$ defined by ϕ on \mathfrak{A} is closed under isomorphisms of \mathfrak{A} . In particular, if \mathfrak{A} is a complete structure, where every permutation on A is an isomorphism of \mathfrak{A} , $\phi^{\mathfrak{A}}$ is closed under the equivalence relation \simeq_k .

There is a bound m , depending only on k such that any structure realizes at most m distinct basic equality types of k -tuples. Therefore, among structures of a given size n , there are (up to isomorphism) at most 2^m distinct structures $\langle A, R \rangle$ such that R is closed under the equivalence relation \simeq_k , each corresponding to a subset S of the set $\{\phi_1, \dots, \phi_m\}$ of basic equality types of k -tuples. Define the function $G : \omega \rightarrow \mathcal{P}(\mathcal{P}(\{\phi_1, \dots, \phi_m\}))$ such that $S \in G(n)$ if and only if the structure $\langle A, R \rangle$ corresponding to the subset S of $\{\phi_1, \dots, \phi_m\}$ is in Q . Let p_0, \dots, p_t be an enumeration of subsets of $\mathcal{P}(\mathcal{P}(\{\phi_1, \dots, \phi_m\}))$, and let q_0, \dots, q_t be the subsets of ω such that $n \in q_i$ if and only if $G(n) \in p_i$.

We now show that for any sentence ϕ of $\text{FO}(Q)$, the set $f = \{n \mid K_n \models \phi\}$ is eventually equal to one of the q_i . Let n_0 be the number of distinct variables that appear in ϕ and let n_1 and n_2 be natural numbers such that $n_1, n_2 > n_0$. We first establish the following lemma:

Lemma 4 *If $G(n_1) = G(n_2)$, then $K_{n_1} \models \phi$, if and only if, $K_{n_2} \models \phi$.*

Proof of Lemma: Assume $G(n_1) = G(n_2)$. We show, by induction, that for any subformula ψ of ϕ and any l -tuples of elements \bar{a} from K_{n_1} and \bar{b} from K_{n_2} such that $\bar{a} \simeq_l$

\bar{b} , $K_{n_1} \models \psi[\bar{a}]$ if and only if $K_{n_2} \models \psi[\bar{b}]$: For the basis, if ψ is a first order formula (*i.e.* it contains no occurrences of Q), this is established by a straightforward pebble game argument (using the fact that $n_1, n_2 > n_0$). If ψ is of the form $Q\bar{x}(\psi')$ and the induction hypothesis holds of ψ' , it follows that it holds for ψ from the fact that $G(n_1) = G(n_2)$. For the inductive case of the other first order connectives, observe that the subformulas for which we have established the induction hypothesis can be replaced by formulas of the form $\bigvee_{i \in I} \phi_i$, where the ϕ_i are basic equality types. This does not change the interpretation of ψ on K_{n_1} and K_{n_2} , and it does not increase the total number of variables. We are therefore back in the basis case.

It follows immediately that $K_{n_1} \models \phi$, if and only if, $K_{n_2} \models \phi$. ■

It follows from the lemma that we can identify a set p_i such that for all $n > n_0$, $K_n \models \phi$, if and only if, $G(n) \in p_i$. But, this is to say that $n \in f$ if and only if $n \in q_i$. ■

In the case where \mathbf{Q} is empty, that is, we are just considering first order logic, the collection consists of just two sets, \emptyset and ω . This is a restatement of the classic result that in the pure language of equality, first-order logic can only express finite or co-finite sets of finite structures. Theorem 5 can therefore be seen as a generalization of this result.

We can now prove the following:

Theorem 6 *For any signature σ and any finite collection \mathbf{Q} of generalized quantifiers, there is a class C of finite structures of signature σ that is recognizable in logarithmic space but is not expressible in $L_{\infty\omega}^\omega(\mathbf{Q})$.*

Proof: By Lemma 3 $\{f_i \mid i \in \omega\}$ represents all subclasses of \mathcal{K} expressible in $L_{\infty\omega}^\omega(\mathbf{Q})$. Let q_0, \dots, q_t be the sets derived from \mathbf{Q} , as in Theorem 5. Since this collection is finite, there is a μ_i (indeed, infinitely many of them) that does not eventually coincide with any of the q_j . Let $C = \{K_n \mid n \in \mu_i\}$. C is clearly recognizable in log-space. ■

Indeed, Theorem 6 remains true if we replace L by any complexity class which does not eventually collapse to finitely many problems (and this includes all standard complexity classes). Theorems 5 and 6 also generalize a result in [KV92a] which showed that if \mathbf{Q} is a finite collection of *unary* quantifiers, then $L_{\infty\omega}^\omega(\mathbf{Q})$ is strictly weaker than $L_{\infty\omega}^\omega(\mathbf{C})$, where \mathbf{C} is the collection of all counting quantifiers (see Example 1). This follows from our results, because $L_{\infty\omega}^\omega(\mathbf{C})$ can express every subset of \mathcal{K} .

4 Quantifiers and Reducibilities

We now establish some connections between results on generalized quantifiers and the notion of logical reducibilities. By logical reducibilities, we refer to reductions between problems that are determined, not by resource-bounds on the computation of the reduction, but by the definability of the reduction in a logical language. The notion is derived from the idea of interpretations between theories (see, for instance, [End72]), and was used in [LG77] and [Imm87]. The following definitions are based on those in [Imm87].

Definition 6 *Let σ and τ be two signatures, where $\tau = \langle R_1, \dots, R_r \rangle$ and the arity of R_i is n_i (for $1 \leq i \leq r$) and let L be a logic. An L -interpretation of τ in σ is a sequence, $\langle \pi_{\forall}, \pi_1, \dots, \pi_r \rangle$ of formulas of L in the signature σ , such that the free variables of π_{\forall} are among x_1, \dots, x_k (for some k) and the free variables of π_i (for each i) are among x_1, \dots, x_{k+n_i} . The width of the interpretation is k .*

An interpretation of τ in σ , of width k , can be seen as a map, π , from structures over the signature σ to structures over τ . If \mathfrak{A} is structure over σ , with universe A , then $\pi(\mathfrak{A}) = \langle B, R_1^B, \dots, R_r^B \rangle$, where $B = \{\bar{a} \in A^k \mid \mathfrak{A} \models \pi_{\forall}[\bar{a}]\}$ and for each i , $R_i^B = \{\bar{a}_1 \dots \bar{a}_{n_i} \mid \bar{a}_1, \dots, \bar{a}_{n_i} \in B \text{ and } \mathfrak{A} \models \pi_i[\bar{a}_1 \dots \bar{a}_{n_i}]\}$. In the following, we will use π both for the interpretation and for the map it defines when no confusion would result.

Definition 7 *Given C_1 - a class of structures over σ , C_2 - a class of structures over τ , and π , an L -interpretation of τ in σ , π is an L - m -reduction of C_1 to C_2 if and only if $\mathfrak{A} \in C_1 \Leftrightarrow \pi(\mathfrak{A}) \in C_2$. If such a π exists, we say that C_1 is L - m -reducible to C_2 .*

In the case where L is first order logic, the notion of an L - m -reduction is essentially the same as that of a *first order translation* in [Imm87] or an *elementary reduction* in [LG77].

Definition 8 *An L - m -reduction, π , is a linear reduction if it has width 1.*

The following straightforward lemma links the notion of linear reduction with generalized quantifiers:

Lemma 5 *For any class of structures C (over a signature $\tau = \langle R_1, \dots, R_r \rangle$), there is a generalized quantifier Q such that every class that is linearly L - m -reducible to C is expressed by a sentence of $L(Q)$.*

Proof: Let C' be the class of structures over the signature $\langle U, R_1, \dots, R_r \rangle$ (where U is a unary relation) such that $\mathfrak{A} \in C'$ if and only if the substructure of \mathfrak{A} generated by the set $U^{\mathfrak{A}}$ is in C . Then, if $\pi = \langle \pi_{\forall}, \pi_1, \dots, \pi_r \rangle$ is a linear L - m -reduction from any class D to C , the sentence $Q_{C'} \bar{x}(\pi_{\forall}, \pi_1, \dots, \pi_r)$ expresses D . ■

With the aid of Lemma 5, the following is a direct corollary of Theorem 6:

Corollary 2 *There are no problems that are hard for L with respect to linear $L_{\infty\omega}^{\omega}$ -reductions.*

The situation is different when we consider reductions that are not linear. Immerman shows in [Imm87] that there are problems that are complete for P (and for L) via FO- m -reductions in the ordered case. That is, in these constructions, it is assumed that there is a linear order on the domain of every structure, and this order is available as a logical relation. Lovàsz and Gàcs [LG77] show that SAT is complete for NP via FO- m -reductions, with a weaker requirement on structures than a linear order. They also show that a number of other problems are NP-complete via FO- m -reductions when an ordering is present. We show below that there is a problem that is complete for NP via FO- m -reductions, without any requirements on the domain of the structures. We also establish that there are properties that are complete for P and L via FO- m -reductions, if and only if, these classes have recursively enumerable index sets. This is done by a general construction linking the existence of complete problems for a complexity class to the existence of a recursive indexing of that class.

To establish the link between generalized quantifiers and non-linear reductions, we need the following definition:

Definition 9

1. *Given a class of structures, C (over a signature $\tau = \langle R_1, \dots, R_r \rangle$, where the arity of R_i is n_i), for each $k \in \omega$, let C_k be a class of structures over the signature $\langle U_k, R_{k,1}, \dots, R_{k,r} \rangle$ (where the arity of U_k is k and the arity of $R_{k,i}$ is $k \cdot n_i$) such that a structure \mathfrak{A} is in C_k if and only if the structure with universe $U_k^{\mathfrak{A}}$ and relations R'_1, \dots, R'_r (with arity of R'_i being n_i) given by $R'_i = \{ \langle \bar{a}_1 \dots \bar{a}_{n_i} \rangle \mid \bar{a}_j = \langle a_{j,1} \dots a_{j,k} \rangle \in U_k^{\mathfrak{A}} \text{ and } \langle a_{1,1} \dots a_{n_i,k} \rangle \in R_{k,i}^{\mathfrak{A}} \}$ is in C .*

2. If Q_k is the generalized quantifier associated with C_k , we say that the sequence of quantifiers $\{Q_k \mid k \in \omega\}$ is uniformly generated by C .
3. A countable collection of quantifiers, \mathbf{Q} , is a uniform sequence if there is a class of structures C , such that \mathbf{Q} is uniformly generated by C .

The following definition is motivated by the view that generalized quantifiers perform a role with respect to formulas similar to that of oracles with respect to machines. This is further justified in Theorem 8 below.

Definition 10 *A class of structures C_1 is L-T-reducible to a class C_2 if C_1 is expressible in $L(\mathbf{Q})$, where \mathbf{Q} is the sequence of generalized quantifiers uniformly generated by C_2 .*

The following lemma is a direct extension of Lemma 5.

Lemma 6 *If C_1 is L-m-reducible to C_2 , then C_1 is L-T-reducible to C_2 .*

We are now in a position to establish the following result:

Theorem 7 *If \mathcal{C} is any bounded complexity class that is closed under FO-m-reductions, then the following are equivalent:*

1. *there is a complete problem for \mathcal{C} with respect FO-m-reductions;*
2. *there is an index set for \mathcal{C} in P; and*
3. *there is a recursively enumerable index set for \mathcal{C} .*

Proof:

1 \Rightarrow 2 Let Q be the \mathcal{C} -complete problem and let \mathbf{Q} be the sequence of generalized quantifiers uniformly generated by Q . Since \mathcal{C} is closed under first order operations, it is captured exactly by $\text{FO}(\mathbf{Q})$. The sentences of $\text{FO}(\mathbf{Q})$ of the form $Q\bar{x}(\phi_1, \dots, \phi_r)$, where ϕ is first-order form an index set for \mathcal{C} . This set of sentences is clearly in P.

2 \Rightarrow 3 is trivial.

3 \Rightarrow 1 We will construct a class of structures Q that is complete for graph problems in \mathcal{C} via FO-m-reductions. Since, for any signature σ , there is an isomorphism preserving

first order translation from structures over σ to graphs (see, for instance, [Lin87]), this suffices. It is easily verified that if there is an r.e. index set for a class \mathcal{C} , there is a recursive index set.

Let \mathcal{I} be a recursive index set for \mathcal{C} . Since \mathcal{C} is bounded, there is a function t such that for each $i \in \mathcal{I}$, there is an associated k_i such that the complexity of the class determined by i is bounded by $t(n^{k_i})$. Also, since \mathcal{I} is a recursive set, there is a machine M and a recursive function g such that M accepts $i \in \mathcal{I}$ in time (and space) less than $g(i)$.

We now define the class Q – a class of structures over the signature $\langle V, E, \preceq, I \rangle$, where V and I are unary and E and \preceq are binary. A structure $\mathfrak{A} = \langle A, V, E, \preceq, I \rangle$ is in Q if and only if:

1. \preceq is a linear pre-order on A ;
2. if $a, b \in I$, $a \preceq b$ and $b \preceq a$, i.e. I picks out one equivalence class from the pre-order (say the i^{th});
3. i is in \mathcal{I} ;
4. $|A| \geq |V|^{k_i}$;
5. the graph $\langle V, E \rangle$ is in the class determined by i ; and
6. $g(i) \leq t(|A|)$.

We verify that Q is in \mathcal{C} . On input \mathfrak{A} , conditions 1, 2 and 4 are easily checked (in $\log n$ space and linear time). Condition 3 is checked by running M on input i . If the machine exceeds resource bounds $t(|A|)$, it is halted and the input is rejected, since it violates condition 6. Finally, we check condition 5, which by virtue of condition 4 and the definition of k_i can be done in resource bound $t(|A|)$.

Next, we verify that Q is complete for \mathcal{C} . Let P_i be the class in \mathcal{C} determined by i and let k' be a natural number such that there are at least i distinct basic equality types of k' -tuples. There are only finitely many structures in P_i of size at most $\max(k', g(i))$. Since each finite structure is determined up to isomorphism by a first order sentence, we can write a first order formula that picks out exactly these structures and maps them to a selected structure in Q . For larger structures, we define the translation as

follows: let $k = \max(k', k_i)$. A graph $\langle V, E \rangle$ is mapped to $\langle V^k, V', E', \preceq, I \rangle$, where $V' = \{\langle v \dots v \rangle \mid v \in V\}$, E' is the natural extension of E to V' , \preceq is an arbitrary ordering of the basic equality types of k -tuples (this is first order definable, since there are only finitely many such types), and I picks out the i^{th} type in this ordering. It is easily verified that all of these are first order definable. ■

Theorem 7 remains true even if we replace the notion of FO- m -reduction with the weaker notion of a projection translation as defined in [Imm87].

Corollary 3 *There is a class of structures that is complete for NP via FO- m -reductions.*

Corollary 4 *There is a recursively indexable collection of generalized quantifiers \mathbf{Q} in P (resp. L, NL) such that FO(\mathbf{Q}) captures P (resp. L, NL), if and only if, there is a uniform sequence of generalized quantifiers with this property.*

We noted above that, intuitively, generalized quantifiers play a role similar to that of oracles. This can be made precise in the cases where a logic is known to correspond exactly to a natural complexity class. Thus, in particular, the following is a direct extension of the equivalences established in [Fag74, Imm86, Var82]:

Theorem 8 *If A is a language encoding a class of structures C , and \mathbf{Q} is the sequence of quantifiers uniformly generated by C , then:*

1. $\text{NP}^A = \Sigma_1^1(\mathbf{Q})$; and
2. on ordered structures $\text{P}^A = \text{LFP}(\mathbf{Q})$.

5 Conclusion

Hella [Hel92] showed that there is no collection of generalized quantifiers \mathbf{Q} of bounded arity such that $\text{LFP}(\mathbf{Q})$ expresses every property in P. We have strengthened this result to show that even on structures over a fixed signature (such as the language of graphs), where there clearly is an infinite such collection \mathbf{Q} , there is no finite one. It remains an open question whether, in such a case, there is a recursively enumerable set of generalized quantifiers \mathbf{Q} such that $\text{LFP}(\mathbf{Q})$ expresses exactly the properties in P. However, we established that if there is such a \mathbf{Q} , there is one that satisfies strict uniformity conditions. This

is shown by establishing a close connection between the existence of complete problems for a complexity class, the indexability of the class, and the existence of certain uniform sequences of generalized quantifiers.

References

- [Bar77] J. Barwise. On Moschovakis closure ordinals. *Journal of Symbolic Logic*, 42:292–296, 1977.
- [DLW92] A. Dawar, S. Lindell, and S. Weinstein. Element types and inductive definability on finite structures. In preparation, 1992.
- [DLW93] A. Dawar, S. Lindell, and S. Weinstein. Infinitary logic and inductive definability over finite structures. *Information and Computation*, 1993. to appear.
- [Ebb85] H.-D. Ebbinghaus. Extended logics: The general framework. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 25–76. Springer-Verlag, New York, 1985.
- [End72] H. B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, 1972.
- [Fag74] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In R. M. Karp, editor, *Complexity of Computation, SIAM-AMS Proceedings, Vol 7*, pages 43–73, 1974.
- [Gur88] Y. Gurevich. Logic and the challenge of computer science. In E. Börger, editor, *Current Trends in Theoretical Computer Science*, pages 1–57. Computer Science Press, 1988.
- [Hel92] L. Hella. Logical hierarchies in PTIME. In *Proc. 7th IEEE Symp. on Logic in Computer Science*, 1992.
- [Imm82] N. Immerman. Upper and lower bounds for first-order expressibility. *Journal of Computer and System Sciences*, 25:76–98, 1982.
- [Imm86] N. Immerman. Relational queries computable in polynomial time. *Information and Control*, 68:86–104, 1986.
- [Imm87] N. Immerman. Languages that capture complexity classes. *SIAM Journal of Computing*, 16:760–778, 1987.
- [Imm89] N. Immerman. Descriptive and computational complexity. In J. Hartmanis, editor, *Computational Complexity Theory, Proc. of AMS Symposia in Appl. Math.*, volume 38, pages 75–91, 1989.
- [KV92a] Ph. G. Kolaitis and J. A. Väänänen. Generalized quantifiers and pebble games on finite structures. In *Proc. 7th IEEE Symp. on Logic in Computer Science*, pages 348–359, 1992.

- [KV92b] Ph. G. Kolaitis and M. Y. Vardi. Fixpoint logic vs. infinitary logic in finite-model theory. In *Proc. 7th IEEE Symp. on Logic in Computer Science*, pages 46–57, 1992.
- [KV92c] Ph. G. Kolaitis and M. Y. Vardi. Infinitary logics and 0-1 laws. *Information and Computation*, 98(2):258–294, 1992.
- [LG77] L. Lovász and P. Gács. Some remarks on generalized spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 23:27–144, 1977.
- [Lin66] P. Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966.
- [Lin87] S. Lindell. *The Logical Complexity of Queries on Unordered Graphs*. PhD thesis, University of California, Los Angeles, 1987.
- [Poi82] B. Poizat. Deux ou trois choses que je sais de L_n . *Journal of Symbolic Logic*, 47(3):641–658, 1982.
- [Rub75] A. Rubin. *Free Algebras in Von Neumann-Bernays-Godel Set Theory and Positive Elementary Inductions in Reasonable Structures*. PhD thesis, California Institute of Technology, 1975.
- [Var82] M. Y. Vardi. The complexity of relational query languages. In *Proceedings of the 14th ACM Symposium on the Theory of Computing*, pages 137–146, 1982.