

Subclasses of Binary NP

December 5, 1996

Arnaud Durand
Université de Caen

Clemens Lautemann Thomas Schwentick
Johannes Gutenberg-Universität Mainz

Abstract

Binary NP consists of all sets of finite structures which are expressible in existential second order logic with second order quantification restricted to relations of arity 2. We look at semantical restrictions of binary NP, where the second order quantifiers range only over certain classes of relations. We consider mainly three types of classes of relations: unary functions, order relations and graphs with degree bounds.

We show that many of these restrictions have the same expressive power and establish a 4-level strict hierarchy, represented by sets, permutations, unary functions and arbitrary binary relations, respectively.

1 Introduction

It is a well-known fact that many results, tools, and techniques of mathematical logic break down when only finite structures are considered as models (cf. [Gur88]). Although this restriction to finite structures is of little interest in most areas of traditional mathematics the situation is quite different in computer science. It was therefore mainly with applications to computer science in mind that finite model theory was developed in the last few years, and a number of fascinating connections between logic and, in particular, database theory and complexity theory have been discovered. For instance, most computational complexity classes have been characterized in terms of *descriptive* complexity, i.e., as classes of model sets of sentences in some syntactically defined logic.

One of the earliest such results, and the most influential, was Fagin's Theorem [Fag74], which characterizes the complexity class **NP** by existential second order logic. This result created hopes that it might be possible to attack some of the main open problems of complexity theory with methods of logic.

The one powerful proof method of model theory which survives the transfer to the finite realm is Fraïssé's theorem [Fra54], mainly used in the form of Ehrenfeucht games [Ehr61]. In an attempt to tackle the **NP** vs. **coNP** problem Fagin considered monadic **NP** that is the class of those sets of structures that can be characterized by existential second order sentences, in which second order quantifiers range only over *monadic* relations, i.e., sets. He then used an Ehrenfeucht game to prove that the set of connected graphs (which is in monadic **coNP**) is not in monadic **NP**, thus showing that monadic **NP** is not closed under complement [Fag75].

Since then more and more powerful methods for using Ehrenfeucht games have been developed and have led to various extensions and strengthenings of this result [AF90, dR87, FSV95, Sch94, Sch95, KS96]. However, this success has been confined to the monadic fragment of existential second order logic, and new techniques seem necessary to analyse the expressive power of higher fragments such as binary **NP** (where second order quantifiers range over *binary* relations).

On the other hand, we know that the hierarchy induced by restricting the arity of quantified relations is strict: Ajtai showed in [Ajt83] with involved combinatorial arguments that for all k , quantification over $(k+1)$ -ary relations is strictly more powerful than quantification over relations of arity at most k . However, the separating example in Ajtai's proof is a set of $(k+1)$ -ary hypergraphs; it is not known whether the hierarchy is strict for a

fixed signature. In particular, we are not able to prove for any graph class in **NP** that it is not in binary **NP**. In fact one quantifier over binary relations suffices to express many natural graph properties.

It was shown by Lynch [Lyn82, Lyn92] that, for every $k > 1$, k -ary **NP**, the fragment of existential second order logic in which quantification is allowed over at most k -ary relations, captures at least all sets of strings that are accepted by a nondeterministic Turing machine in time $O(n^k)$. In particular, **BinNP** captures nondeterministic quadratic time on strings. It is an interesting open problem to find a (natural) graph problem that is not in **BinNP**. As a **BinNP** formula can be evaluated in quadratic space and every problem on strings can be easily translated into a problem on graphs, it follows directly from the space hierarchy theorem that there exist graph problems in **PSPACE** that are not in **BinNP**. On the other hand, the evaluation of a **BinNP** formula on a nondeterministic polynomial-time bounded Turing machine needs only a quadratic amount of nondeterminism. Hence, by an analogue argument, if every **NP** graph problem were in **BinNP**, then every **NP** problem would need only a quadratic amount of nondeterminism.

In order to gain insight into the expressive power of binary quantifiers, we look at *semantical* restrictions, where these quantifiers range only over certain classes of relations. For instance, in order to express that a graph has a Hamiltonian cycle, it is enough to quantify over one successor relation and to check that successive elements are connected by an edge.

Some such restrictions have been considered in the literature: subsets of the set of edges [Tur84], unary functions (which suffice for capturing nondeterministic linear time) [Gra90, DR94] and certain pairing relations on strings [LST95].

Of course, when restricting second order quantification in this way, in order to obtain interesting results, we have to use interesting classes of relations.¹

In this paper we will mainly consider three types of classes of binary relations:

- unary functions;
- order relations;

¹Any set S of graphs can easily be defined by the sentence $\exists E' \forall x, y (E(x, y) \leftrightarrow E'(x, y))$, where E' ranges over those binary relations which are the edge set of a graph in S .

- graphs with degree bounds.

Furthermore, we look at addition, equivalence relations and graphs with at most linearly many arcs.

As a result we obtain the hierarchy indicated in Figure 1, where classes within one box are of the same expressive power, and (strict) inclusion is upwards.

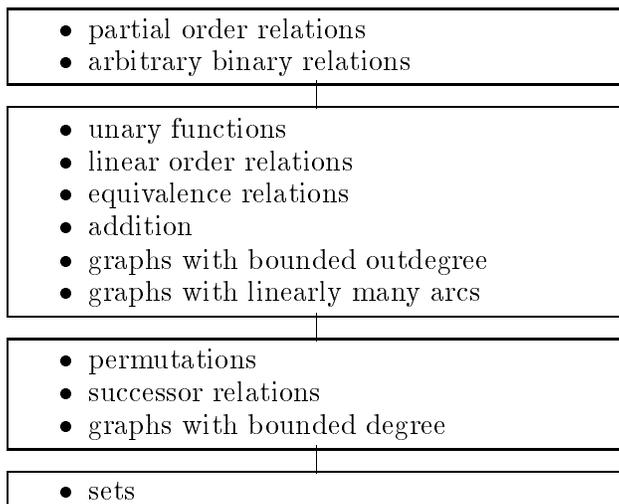


Figure 1: The inclusion structure of some fragments of binary **NP**. All inclusions are strict

Counting from the bottom, box #1 is monadic **NP** and box #4 is binary **NP**.

It should be pointed out that we do not limit the number of second order quantifiers. For a recent result on restricting the number of quantified unary functions see [Loe96].

The separation at the bottom is well-known: structures with an even number of elements are easily defined with permutations, but impossible to define with sets. The separation at the top can be proved directly using Ajtai's main lemma from [Ajt83]; we will prove it with a reduction. For the separation between #2 and #3 we will show in fact that the set of graphs in which the number of vertices equals the number of arcs can not be expressed by permutations but by unary functions. The proof of this result combines the mentioned lemma of Ajtai with a winning strategy argument for Ehrenfeucht games (cf. [Sch94]).

It should be noted that in the case of ordered structures the situation is different. As Etienne Grandjean [Gra] pointed out, levels #2 and #3 coincide in this case (because any linear order can be represented as the image of the given linear order under a permutation).

It was shown by Shelah and Baldwin that in the case of infinite structures there are only four different classes of first-order definable second-order quantifiers, represented by binary relations, permutations, unary relations and empty relations, respectively [She73, Bal85, BS85, She86].

All our equivalence results are based on techniques for expressing one kind of relation by another. We systematize this argument with the notion of representability of a class of relations by another class of relations.

This definition, together with the other necessary formal preliminaries, is given in Section 2. In the following sections we move from the top to the bottom of Figure 1:

- In Section 3 we investigate level 4 and separate it from level 3;
- In Section 4 we investigate level 3;
- In Section 5 we investigate level 2 and separate it from level 1;
- In Section 6 we separate level 2 from level 3.

Finally, Section 7 gives a short discussion.

We would like to thank Etienne Grandjean, Ron Fagin, John Baldwin, Yuri Gurevich and an anonymous referee for a lot of helpful suggestions.

2 Preliminaries

2.1 Definitions

We make use of the following notations.

A finite structure A with universe U , relations R_1, \dots, R_l and constants c_1, \dots, c_m is written $\langle U, R_1, \dots, R_l, c_1, \dots, c_m \rangle$. We also refer to the universe of A by $|A|$. If A is a finite structure and S_1, \dots, S_p and d_1, \dots, d_q are additional relations and constants on $|A|$ we write $\langle A, S_1, \dots, S_p, d_1, \dots, d_q \rangle$ for the extension of A by the new relations and constants. We view unary functions f as binary relations via $f(x) = y \iff (x, y) \in f$. Unless otherwise stated, graphs are always directed and do not need to be loop-free.

In all our proofs we assume that the universes contain at least 3 elements to omit technical subtleties. Of course, smaller universes can always be dealt with in the first order part of our formulas.

We are interested in sets of finite structures which can be defined by existential second order sentences, where the second order quantifiers are restricted to range only over certain kinds of binary relations. We will look at the following types of binary relations:

- arbitrary binary relations (*BinRel*);
- unary functions (*UnF*);
- permutations, i.e. bijective unary functions (*Perm*);
- partial orders (*PartOrd*);
- linear orders (*LinOrd*);
- successor relations (*Succ*);
- equivalence relations, i.e. symmetric, reflexive, transitive relations (*Equiv*);
- ternary relations which are isomorphic to the addition on an initial segment of the integers² (*Add*);
- directed graphs with outdegree at most k (*k-OutDegGr*);
- directed graphs with total degree at most k (*k-DegGr*);
- relations with at most n arcs, where n is the size of the universe (*Linear*).

²Of course these are no binary relations. But it turns out that the resulting class coincides with a subclass of binary NP.

In this list, the names of the respective sets of finite structures are indicated in parentheses. E.g., *LinOrd* is the set of all finite structures $\langle U, R \rangle$ with universe U and a linear order R on U .

Let in the following σ denote an arbitrary relational signature and let τ be a signature with one relation symbol R .

2.1 Definition

Given a set \mathcal{C} of finite σ -structures and a set \mathcal{D} of finite structures, we say that \mathcal{C} is in $\exists\mathcal{D}$ FO, iff there is a k and a first order formula ψ over $\sigma\langle R_1, \dots, R_k \rangle$ such that for every finite structure A the following holds:

$$A \in \mathcal{C} \iff \begin{array}{l} \text{there are relations } R_1, \dots, R_k \text{ on } |A|, \text{ such} \\ \text{that} \\ \bullet \langle |A|, R_i \rangle \in \mathcal{D}, \text{ for every } i, \text{ and} \\ \bullet \langle A, R_1, \dots, R_k \rangle \models \psi. \end{array}$$

We will often say that a set of structures is *expressible by permutations* (unary functions etc.), if it is in $\exists\text{Perm FO}$ ($\exists\text{UnF FO}$ resp.).

In our applications σ will almost always consist of a single binary relation symbol.

Let **BinNP** be the class of sets of structures that are expressible by arbitrary binary relations.

To show that a class $\exists\mathcal{C}$ FO is contained in a class $\exists\mathcal{D}$ FO, we will always prove that relations of \mathcal{C} can be encoded into one or several relations of type \mathcal{D} in a first order manner. This connection is formalized in the next definition and in Lemma 2.3.

2.2 Definition

Given a set \mathcal{C} of finite τ -structures and a set \mathcal{D} of finite structures, we say that \mathcal{C} is representable by \mathcal{D} if there are l, p and a first-order formula φ over $\langle S_1, \dots, S_l \rangle$ with tuples \bar{x} and $\bar{y} = (y_1, \dots, y_p)$ of free variables, such that for every structure $\langle U, R \rangle \in \mathcal{C}$, there exists a tuple $\bar{S} = S_1, \dots, S_l$ of relations over U , fulfilling $\langle U, S_i \rangle \in \mathcal{D}$, for every i , and a tuple $\bar{b} = b_1, \dots, b_p$ of elements of U such that for every \bar{a} :

$$\bar{a} \in R \iff \langle U, \bar{S}, \bar{b}, \bar{a} \rangle \models \varphi.$$

Examples are found in Sections 3, 4 and 5.

In our setting the notion of representability plays a similar role as the notion of interpretability does in the context of first-order logic (cf. [EF95]).

The tuple \bar{y} in the definition of representability will give us an extra amount of freedom to handle special cases and will simplify some of the constructions.

2.3 Lemma

If a set \mathcal{C} of τ -structures is representable by \mathcal{D} and $C \in \exists\mathcal{D} FO$ then $\exists\mathcal{C} FO \subseteq \exists\mathcal{D} FO$.

Proof.

The proof is straightforward but notationally complicated. Nevertheless, we present it in detail because this lemma is the essential tool for proving most of our results.

Let $L \in \exists\mathcal{C} FO$ be a set of σ -structures. By Definition 2.1 there are a number k and a first order sentence ψ over $\sigma\langle R_1, \dots, R_k \rangle$ such that

$$\begin{aligned} A \in L \iff & \text{ there are relations } R_1, \dots, R_k \text{ on } |A| \text{ such that:} \\ & \langle |A|, R_i \rangle \in \mathcal{C}, \text{ for every } i = 1, \dots, k, \text{ and} \\ & \langle A, R_1, \dots, R_k \rangle \models \psi. \end{aligned} \tag{1}$$

Since $C \in \exists\mathcal{D} FO$, Definition 2.1 gives us a number m and a first-order sentence θ over $\tau\langle T_1, \dots, T_m \rangle$ such that for every A and every relation R on A :

$$\begin{aligned} \langle |A|, R \rangle \in \mathcal{C} \iff & \text{ there are relations } T_1, \dots, T_m \text{ on } |A| \text{ such that:} \\ & \langle |A|, T_j \rangle \in \mathcal{D}, \text{ for every } j = 1, \dots, m, \text{ and} \\ & \langle A, T_1, \dots, T_m \rangle \models \theta. \end{aligned}$$

For $i = 1, \dots, k$, we let θ_i be the formula obtained from θ by replacing the relation symbols R with R_i and T_1, \dots, T_m with T_{i1}, \dots, T_{im} , respectively. Then (1) can be rewritten as:

$$\begin{aligned} A \in L \iff & \text{ there are relations } R_i, T_{i1}, \dots, T_{im} \text{ on } |A|, \\ & \text{ for } i = 1, \dots, k, \text{ such that} \\ & 1. \langle |A|, T_{ij} \rangle \in \mathcal{D}, \text{ for } i = 1, \dots, k, j = 1, \dots, m, \\ & 2. \langle |A|, R_i, T_{i1}, \dots, T_{im} \rangle \models \theta_i, \text{ for } i = 1, \dots, k, \text{ and} \\ & 3. \langle A, R_1, \dots, R_k \rangle \models \psi. \end{aligned} \tag{2}$$

Since θ_i does not contain any relation symbols other than $R_i, T_{i1}, \dots, T_{im}$, and ψ contains only symbols from σ and R_1, \dots, R_k , we can combine 2. and 3. into one expression and obtain:

$$\begin{aligned}
A \in L &\iff \text{there are relations } R_i, T_{i1}, \dots, T_{im} \text{ on } |A|, \\
&\text{for } i = 1, \dots, k, \text{ such that} \\
&1. \langle |A|, T_{ij} \rangle \in D, \text{ for } i = 1, \dots, k, j = 1, \dots, m, \text{ and} \quad (3) \\
&2. \langle A, R_1, \dots, R_k, T_{11}, \dots, T_{km} \rangle \models \psi \wedge \bigwedge_{i=1}^k \theta_i.
\end{aligned}$$

We now use the representability of \mathcal{C} in order to eliminate the relations R_1, \dots, R_k . Let l, p and φ be as in Definition 2.2. We let φ_i be the formula obtained from φ by replacing the relation symbols S_1, \dots, S_l with S_{i1}, \dots, S_{il} and the variables y_1, \dots, y_p with y_{i1}, \dots, y_{im} , respectively.

Now any relation R_i with $\langle |A|, R_i \rangle \in \mathcal{C}$ can be represented by \mathcal{D} -relations S_{i1}, \dots, S_{il} and elements b_{i1}, \dots, b_{ip} , in such a way that

$$\bar{a} \in R_i \iff \langle |A|, S_{i1}, \dots, S_{il}, b_{i1}, \dots, b_{ip}, \bar{a} \rangle \models \varphi_i.$$

Therefore, if we replace the atom $R_i(\bar{x})$ by the formula $\varphi_i(\bar{x})$ - thus obtaining $\tilde{\theta}_i$ from θ_i and $\tilde{\psi}$ from ψ - we can conclude from (3)

$$\begin{aligned}
A \in L &\implies \text{there are relations } S_{i1}, \dots, S_{il}, T_{i1}, \dots, T_{im} \text{ on } |A| \text{ and} \\
&\text{elements } b_{i1}, \dots, b_{ip} \in |A|, \text{ for } i = 1, \dots, k \text{ such that} \\
&\langle |A|, T_{ij} \rangle \in D, \text{ for } i = 1, \dots, k, j = 1, \dots, m, \quad (4) \\
&\langle |A|, S_{ij} \rangle \in D, \text{ for } i = 1, \dots, k, j = 1, \dots, l, \text{ and} \\
&\langle A, T_{11}, \dots, T_{km}, S_{11}, \dots, S_{kl}, b_{11}, \dots, b_{kp} \rangle \models \tilde{\psi} \wedge \bigwedge_{i=1}^k \tilde{\theta}_i.
\end{aligned}$$

Now, let the right-hand side of (4) be given. Define the relations R_1, \dots, R_k by setting

$$R_i := \{\bar{a} \mid \langle A, T_{i1}, \dots, T_{im}, S_{i1}, \dots, S_{il}, b_{i1}, \dots, b_{ip}, \bar{a} \rangle \models \varphi_i\}.$$

Then we obtain from the last condition in (4)

$$\langle A, R_1, \dots, R_k, T_{11}, \dots, T_{km}, S_{11}, \dots, S_{kl}, b_{11}, \dots, b_{kp} \rangle \models \psi \wedge \bigwedge_{i=1}^k \theta_i.$$

Since the only relation symbols in ψ are R_1, \dots, R_k and those of σ , and the only ones in θ_i are $R_i, T_{i1}, \dots, T_{im}$, we can break this up into

- (i) $\langle |A|, R_i, T_{i1}, \dots, T_{im} \rangle \models \theta_i$, for $i = 1, \dots, k$, and
- (ii) $\langle A, R_1, \dots, R_k \rangle \models \psi$.

As, furthermore, we know that $\langle |A|, T_{ij} \rangle \in \mathcal{D}$, for all i, j , this gives us the right hand side of (2), hence $A \in L$, as required. □

It should be noted that $\mathcal{C} \in \exists \mathcal{D}FO$ is an essential condition in the statement of Lemma 2.3. But most of the classes that we investigate are even first-order decidable. The only two exceptions are *Succ* and *Linear*.

Two vertices of a structure A are *adjacent* if they occur in a tuple of one of the relations of A . The distance $d(u, v)$ between vertices u and v is the minimal k such that there exist $v_0 = u, v_1, \dots, v_k = v$ and every v_i is adjacent to v_{i+1} for $i < k$. For a substructure H of A we set $d(v, H) := \min_{u \in H} d(u, v)$.

2.2 Games

The definitions of this subsection are needed only in Section 6.

As mentioned in the introduction Ehrenfeucht games [Ehr61] are important for proving inexpressibility results. The rules of a first-order (FO) Ehrenfeucht game are as follows.

There are two players, Spoiler and Duplicator. They play on two structures A_1, A_2 . Spoiler's aim is to prove a difference between A_1 and A_2 , whereas Duplicator tries to make them look alike.

They play a fixed number, k , of rounds. In every round, Spoiler chooses an element of one of the two structures. Then Duplicator chooses an element of the other structure. We write a_i for the element of A_1 , chosen in round i , and a'_i for the element of A_2 , chosen in round i .

At the end of the game, Duplicator wins if the structures induced by the chosen elements are isomorphic under an isomorphism which maps a_i to a'_i for every i .

The importance of Ehrenfeucht games results from the fact that Spoiler has a k -round winning strategy on structures A_1, A_2 iff there exists a first-order formula φ of quantifier rank (i.e., nesting depth of quantifiers) at most k which holds in A_1 but not in A_2 . For our purposes the following formulation of this connection is sufficient.

2.4 Theorem

[Ehr61, Fra54] *Let L be a set of structures. L is first order definable, if and only if there is a fixed k , such that, whenever $A_1 \in L$ and $A_2 \notin L$, then*

Spoiler has a winning strategy in the k -round FO Ehrenfeucht game on A_1 and A_2 .

For a proof of this theorem and a rigorous definition of quantifier rank see [EF95].

Ehrenfeucht games can be extended to characterize second order expressibility [Fag75, Ten75, Loe91]. For Monadic **NP** Ajtai and Fagin [AF90] invented such a game which can be easily transferred to other existential SO logics. As we are going to use Ehrenfeucht games to separate the subclasses of **BinNP** induced by permutations and unary functions respectively, we give here a version of the game for expressibility by permutations.

The *Permutation game* for a set L of graphs consists of the following steps.³

- (1) Spoiler chooses numbers k and l .
- (2) Duplicator selects a graph $A_1 \in L$.
- (3) Spoiler chooses a tuple $\bar{f} = (f_1, \dots, f_l)$ of permutations on $|A_1|$.
- (4) Duplicator selects a graph $A_2 \notin L$ and a tuple $\bar{f}' = (f'_1, \dots, f'_l)$ of permutations on $|A_2|$.
- (5) Spoiler and Duplicator play a k -round FO Ehrenfeucht game on the structures $\langle A_1, \bar{f} \rangle$ and $\langle A_2, \bar{f}' \rangle$.

Analogously to the result of Ajtai and Fagin we get the following.

2.5 Theorem

A set L of graphs is $\exists Perm$ FO, if and only if Spoiler has a winning strategy in the Permutation game over L .

The unary function game (UF game) is defined analogously in an obvious way. For the proof of our main separation we need a modified version of this game. The modified UF game has the additional feature that Duplicator has to choose a graph A_2 on the same vertex set as A_1 and is not allowed to choose any functions by himself. Instead, the first order game is played on the structures $\langle A_1, \bar{f} \rangle$ and $\langle A_2, \bar{f} \rangle$.

This makes the game more difficult for Duplicator. In fact, a winning strategy of Spoiler in the modified UF game on a set S does not imply the expressibility of S by unary functions.⁴

³The reader should keep in mind that we view permutations as binary relations.

⁴For the corresponding modified set game, Fagin defines an equivalent logic in [Fag96].

In Section 6 we need the following model-theoretic notions. We say that two structures A_1 and A_2 are k -equivalent if for every first-order formula φ of quantifier rank at most k it holds that $A_1 \models \varphi \iff A_2 \models \varphi$. From the remark before Theorem 2.4 it follows that this is the case if and only if Duplicator has a winning strategy in the k -round Ehrenfeucht game on A_1 and A_2 . The k -type, $\tau_k(A)$, of a structure A is its equivalence class with respect to k -equivalence.

We will make use of the fact that for fixed k and a fixed signature (i.e., fixed number and arities of the relations of the structure) the number of different k -types is finite (cf. [EF95]).

Finally we state a version of the Weak Extension Theorem from [Sch95]. It says that under certain circumstances a winning strategy of Duplicator on substructures H_1 of A_1 and H_2 of A_2 can be extended to a winning strategy on A_1 and A_2 .

Let the e -neighbourhood of H_1 in A_1 be defined as the set of all vertices a with $d(a, H_1) \leq e$. We say that Duplicator has a distance respecting winning strategy on neighbourhoods of H_1 and H_2 , if he can play in such a way that $d(a_i, H_1) = d(a'_i, H_2)$ for every i .

2.6 Theorem

Let $k > 0$.

Let A_1, A_2 be two structures and let H_1 and H_2 be induced substructures of A_1 and A_2 , respectively.

Let $N(H_1)$ ($N(H_2)$) denote the 2^k -neighbourhood of H_1 in A_1 (H_2 in A_2).

Duplicator has a winning strategy in the k -round FO Ehrenfeucht game on A_1 and A_2 , if the following conditions are fulfilled.

- (i) Duplicator has a distance respecting winning strategy in the k -round Ehrenfeucht game on $N(H_1)$ and $N(H_2)$.
- (ii) There is a distance respecting isomorphism α from $A_1 - H_1$ to $A_2 - H_2$ (i.e., $d(a, H_1) = d(\alpha(a), H_2)$ for every $a \in A_1 - H_1$).

3 The Power of Partial Orders

In this section we show that quantification over partial order relations already has the same power as **BinNP**. In fact, it follows from our proof that partial orders of maximal depth 1 are sufficient.

3.1 Theorem

- (a) $\exists UnF FO \subset \exists BinRel FO$.
- (b) $\exists BinRel FO = \exists PartOrd FO$.

Proof.

- (a) The inclusion holds because every unary function is also a binary relation.

Let E_2 and E_4 be a 2-ary, (resp. 4-ary) relation symbol.

To obtain a *strict* inclusion we show that the set of graphs $P_2 = \{\langle V, E_2 \rangle : |E_2| \text{ is even}\}$ is in the class $\exists BinRel FO$ but not in $\exists UnF FO$.

In [Ajt83], it is shown that a binary relation and a linear ordering on vertices are sufficient to express the evenness of the number of edges in a graph (in fact, one binary relation is enough, but we will not prove that here). So P_2 is clearly in $\exists BinRel FO$. The negative part will also be derived from results stated in [Ajt83].

Let $P_4 = \{\langle V, E_4 \rangle : |E_4| \text{ is even}\}$. We denote by $BinF$ the class of structures with a binary function.

Claim *If P_2 is in $\exists UnF FO$ then P_4 is in $\exists BinF FO$.*

The idea is to view pairs of elements of the universe V as elements of a universe $V' = V \times V$. One unary function on V' can be encoded by two binary functions on V . On the other hand, the number of 4-tuples over V equals the number of edges over V' .

Hence, if the parity of the number of edges over V' can be expressed in $\exists UnF FO$ (where the unary functions are functions over V'), then the parity of the number of 4-tuples over V can be expressed in $\exists BinF FO$.

But Theorem 2.1 in [Ajt83] tells us that P_4 cannot even be expressed by ternary relations. Hence P_2 is not in $\exists UnF FO$.

- (b) We show that $BinRel$ is representable by $PartOrd$. This gives the inclusion from left to right. The opposite inclusion is obvious.

In fact we show a bit more, in that the partial orders we use are of depth one. I.e. there are no $x \neq y \neq z$ fulfilling $x \leq y \leq z$.

Let, for the moment, E be a binary relation over a universe of even size.

Our representation makes use of the following simple idea: we take a bijection from one half of the universe into the other half of the universe. This bijection is of course a partial order. We call the bijection $<_0$ and the two sets it induces A and B .

These sets induce a partition of the arcs of E into arcs from A to A , from A to B , from B to A and from B to B . Each of these four sets of arcs is represented by a different partial order relation. Two of these partial orders are easily obtained:

- $x <_1 y \iff x \in A$ and $y \in B$ and $E(x, y)$,
- $x <_2 y \iff x \in B$ and $y \in A$ and $E(x, y)$.

Obviously $<_1$ and $<_2$ are partial orders.

To represent the arcs from A to A and from B to B we make again use of $<_0$. We represent an arc from x to y , where both x and y are in A by an arc from x to the image of y under $<_0$, which is a vertex in B . Analogously we represent the arcs from B to B . More formally:

$$\begin{aligned}
 x <_3 y &\iff x \in A \text{ and there is } z \in A \text{ such that } E(x, z) \text{ and} \\
 &\quad z <_0 y, \text{ and} \\
 x <_4 y &\iff x \in B \text{ and there is } z \in B \text{ such that } E(x, z) \text{ and} \\
 &\quad y <_0 z.
 \end{aligned}$$

We note that in both cases $z = x$ is allowed. This represents arcs (x, x) .

It is easy to see that $<_1, <_2, <_3$ and $<_4$ are partial orders of depth one.

This completes the description of the representation in case the universe is of even size. If the universe is of odd size there remains a single vertex x_0 that is not matched by $<_0$. We represent the arcs from A to x_0 within $<_1$, the arcs from B to x_0 within $<_2$, the arcs from x_0 to A within $<_4$ and the arcs from x_0 to B within $<_3$. To represent an arc from x_0 to x_0 we make use of one additional element y_1 in the definition of representability. If $x_0 = y_1$ then there is a self-loop in x_0 , if they are different there is no such self-loop.

It is easy to see that a first order formula is sufficient to decode the arcs of E from $\langle_0, \dots, \langle_4, y_1$.

□

4 The Power of Unary Functions

When considering the expressive power of unary functions it is worthwhile to note that unary functions can be used to give an exact characterization of nondeterministic linear time on RAMs, if inputs are encoded as functional structures, e.g. graphs are represented by a universe which is divided into two disjoint parts $V \cup E$ (V for vertices, E for edges) and two functions from E to V , which encode the edges of the graph. A set L of such structures, closed under isomorphism, is decidable in linear time on RAMs iff $L \in \exists UnFFO(\forall)$, where $\exists UnFFO(\forall)$ is the restriction of $\exists UnFFO$ to formulas with only one first-order universal quantifier (see [Gra90, GO94] for more details).

In this section we show the following theorem.

4.1 Theorem

For every $k \geq 1$, the following classes coincide.

- $\exists UnF FO$
- $\exists Equiv FO$
- $\exists LinOrd FO$
- $\exists Add FO$
- $\exists k\text{-OutDegGr} FO$
- $\exists Linear FO$

Proof. The theorem follows from the next two lemmas. We note that, with the only exception *Linear* all the sets of structures are first-order definable. \square

4.2 Lemma

- (a) *UnF* is representable by *Equiv*.
- (b) *Equiv* is representable by *LinOrd*.
- (c) *LinOrd* is representable by *Add*.
- (d) *Add* is representable by *UnF*.

Proof.

- (a) The proof proceeds in two steps. We define an intermediate class, *PartUnF*. It consists of all finite structures with one binary relation which is the graph of a partial unary function f fulfilling the following condition: if $f(x) = y$ for some x and y then either $f(y) = y$ or $f(y)$ is not defined. In other words, the graph of f has no directed paths of length 2 or more. We show that *UnF* is representable by *PartUnF* and that *PartUnF* is representable by *Equiv*.

1. *UnF* is representable by *PartUnF*

Let $\langle U, f \rangle$ be a structure with a unary function f (represented as a binary relation). The components of f consist of trees the roots of which are connected to a directed cycle. Therefore f is 3-colourable (we do not worry about loops). Given such a colouring with colours, say, red, green and blue, we define the partial unary function f_r by

$$f_r(x) = y \iff f(x) = y \text{ and } y \text{ is coloured red.}$$

Correspondingly we define f_g and f_b . We get $f = f_r \cup f_g \cup f_b$ and all of $\langle U, f_r \rangle, \langle U, f_g \rangle, \langle U, f_b \rangle$ are in *PartUnF*.

2. *PartUnF* is representable by *Equiv*

Let $\langle U, f \rangle$ be from *PartUnF*. We construct two equivalence relations, E_1 and E_2 . The equivalence classes of E_1 are all non-empty sets $S_y := \{x \mid f(x) = y\}$ and the set of all remaining elements (those x such that $f(x)$ is undefined). E_2 has as its equivalence classes all sets $S_y \Delta \{y\}$, where S_y is nonempty and all sets $\{z\}$ of remaining elements. Here, Δ denotes the symmetric difference. Of course the sets in E_1 are pairwise disjoint. The sets in E_2 are pairwise disjoint because whenever $f(x) = y$ then $f(y) = y$ or $f(y)$ is undefined. Hence E_1 and E_2 induce equivalence relations. From these equivalence relations f is easily recovered in a first-order manner.

- (b) Let E be an equivalence relation on $\{1, \dots, n\}$. Let $<_1$ be a linear order in which the equivalence classes of E constitute connected intervals. I.e., for any two elements x, y of the same class there is no z of a different class with $x <_1 z <_1 y$. Let $m(C)$ denote, for every

equivalence class C its maximal element with respect to $<_1$. Let $<_2$ be a linear order in which these maximal elements are the smallest elements and choose y_1 such that $z \leq_2 y_1$ iff $z = m(C)$ for some C .

Then two elements x, y are in the same equivalence class if there exists no z such that $x \leq_1 z <_1 y$ and $z \leq_2 y_1$.

- (c) Of course an addition relation on an initial segment of the natural numbers induces a linear order.
- (d) Let Bit_n be the binary relation on the set $\{0, \dots, n\}$ which fulfils

$$Bit_n(x, y) \iff \text{the } y\text{-th bit of } x \text{ is one.}$$

A relation R on a set U is called a bit-relation, if $\langle U, R \rangle$ is isomorphic to $\langle \{0, \dots, n\}, Bit_n \rangle$. Of course, addition on U can be defined in a first order manner from a bit-relation R and the corresponding linear order. From [Gra90] it is known that $LinOrd$ is representable by UnF . It remains only to show how Bit_n can be encoded in a first order manner by a unary function on $\{0, \dots, n\}$.

We define, for every $x \in \{0, \dots, n\}$,

$$h(x) := \max\{i \mid 2^i \text{ divides } x\},$$

the number of zeros at the end of the binary representation of x .

The i -th bit of x is on, if and only if for the maximal number $y \leq x$ which has $h(y) \geq i$ it holds that $h(y) = 1$.

It is easy to see that, given h , we can define Bit explicitly by a first-order formula.

□

4.3 Lemma

- (a) For every k , UnF is representable by k - $OutDegGr$.
- (b) For every k , k - $OutDegGr$ is representable by $Linear$.
- (c) $Linear$ is representable by UnF and $Linear \in \exists UnF FO$.

Proof.

- (a) Of course, every unary function is a graph of outdegree one.

- (b) A graph of total degree at most k has at most kn edges. Hence such graphs are easily representable by k graphs of at most n edges.
- (c) Let $\langle U, R \rangle$ be a structure with n vertices and at most n edges. We represent the edges of R by two unary functions f_1, f_2 and two vertices y_1, y_2 in the following way. y_1 and y_2 are chosen such that $(y_1, y_2) \notin R$. We assign to every edge $e = (x, y)$ of R a unique vertex a and define

$$f_1(a) := x, \quad f_2(a) := y.$$

We say that a represents e . For all remaining vertices b we define $f_1(b) = y_1$ and $f_2(b) = y_2$.

This shows that *Linear* is representable by UnF .

That a structure $\langle U, R \rangle$ is in *Linear* can be tested similarly, by using two unary functions and checking that every edge is represented by at least one vertex. Of course, functions f_1, f_2 with this property do not exist if R has more than n edges.

□

In [DR94] it is shown that k unary functions can be represented by one single graph of bounded out-degree (where the out-degree depends only on k). In particular this shows that all classes mentioned above can be expressed by one single binary quantifier.

5 The Power of Permutations

5.1 Theorem

- (a) *Monadic NP* \subseteq \exists *Perm FO*.
- (b) For every $k \geq 1$, the following classes coincide.

- \exists *Perm FO*
- \exists *Succ FO*
- $\exists k$ -*DegGr FO*

Proof.

- (a) First we show that structures with a unary relation are representable by *Perm*, which implies *Monadic NP* \subseteq \exists *Perm FO*.

Let $\langle U, R \rangle$ be a structure with a unary relation R . Every permutation on U defines the set of its fixed points. On the other hand, for every subset $R \subseteq U$ there is a permutation f which leaves exactly the elements of R fixed, unless $|U \setminus R| = 1$. We handle the latter case by making use of two additional elements, y_1, y_2 . We set $y_1 := y_2 := y$, if $U \setminus R = \{y\}$. Otherwise we choose different elements y_1 and y_2 .

Therefore it holds $x \in R \iff f(x) = x \wedge \neg(x = y_1 \wedge x = y_2)$

On the other hand the inclusion is proper, because

- even cardinality of the universe is not expressible in *Monadic NP*, but
- a permutation f which fulfils $f(f(x)) = x$ and $f(x) \neq x$, for every x , exists iff the universe has even cardinality.

- (b) is shown by Lemmas 5.2 and 5.3 below.

□

5.2 Lemma

- (a) For every $k \geq 2$, *Perm* is representable by k -*DegGr*.
- (b) *Perm* is representable by 1-*DegGr*.
- (c) For every $k \geq 1$, k -*DegGr* is representable by *Perm*.

Proof.

- (a) Of course, permutations have total degree two.
- (b) The edges of the graph of a permutation can be properly coloured with 3 colours. The edges of every colour class can be represented by one graph of total degree at most one.
- (c) Let G be a graph of total degree at most k .

We represent the loops of G by a function f_0 in the same way as we represented sets in the proof of Theorem 5.1. Let G' be the remaining graph, i.e., G without loops.

We make use of an extended version of Vizing's theorem (see for example [WW90]) which says that in a loop-free graph of maximal vertex degree k , in which between any two vertices there are at most d edges, the edges can be coloured with $k + d$ colours in such a way that no two edges of the same colour share a common vertex.

In our setting we have $d = 2$, so we need at most $k + 2$ colours for the edges of G' .

Let the edges of G' be coloured with colours $1, \dots, k + 2$. We represent the edges of G' by defining $f_i(x) = y$ and $f_i(y) = x$, for every edge (x, y) of colour i . For all remaining z and i we set $f_i(z) = z$. Of course this doesn't tell us the orientation of the edges. To get this information we choose f'_i such that $f'_i(x) = x$ just in case there is an edge (x, y) coloured with i for some y .

Again it is easy to see that there is a first order formula that decides $(x, y) \in E$, given the f_i and f'_i .

□

5.3 Lemma

- (a) *Perm* is representable by *Succ*.
- (b) *Succ* is representable by *Perm* and $Succ \in \exists Perm FO$.

Proof.

- (a) Let $\langle U, f \rangle \in Perm$. We first show how we can represent a subset $R \subseteq U$ by two successor relations. Let a_1, \dots, a_m be the elements of R and b_1, \dots, b_l be the elements of $U \setminus R$. Let s be the successor relation induced by the sequence $a_1, \dots, a_m, b_1, \dots, b_l$ (i.e., $s(a_i) = a_{i+1}$ for

$i < m$, $s(a_m) = b_1$ and $s(b_i) = b_{i+1}$ (for $i < l$). Analogously let the successor relation s' be induced by the sequence $b_l, b_{l-1}, \dots, b_1, a_1, \dots, a_m$. Let furthermore $y_1 := a_m$. Then

$$x \in R \iff s(x) = s'(x) \text{ or } x = y_1.$$

We represent the fixed points of f by a unary relation R_0 . The remaining part of f , respresented as a graph, consists of disjoint directed cycles. It is easy to see that there exist two successor relations s_1 and s_2 such that if $f(x) = y \neq x$ it holds that $s_1(x) = y$ or $s_2(x) = y$. E.g., from any cycle of length n we can put $n - 1$ edges into one successor relation and the remaining one into the other.

Of course $s_i(x) = y$ now also holds for some pairs (x, y) that are not edges of f . Therefore, we make use of two additional subsets, R_1 and R_2 , of the universe, each encoded by two successor relations as described above. We get $f(x) = y$ iff $(s_1(x) = y \text{ and } x \in R_1)$ or $(s_2(x) = y \text{ and } x \in R_2)$ or $(x = y \text{ and } x \in R_0)$.

- (b) A successor relation s can be represented by a permutation f and a constant a by defining $f(x) := y$, if $s(x) = y$ and $f(a) := b$, where a and b are the maximal and minimal element of the successor relation respectively. So *Succ* is easily representable by *Perm*.

We show next that the set of structures $\langle U, f \rangle$ in which f is a connected permutation is in $\exists Perm FO$. From this we can easily conclude that $Succ \in \exists Perm FO$.

In order to express the fact that a permutation f is connected, we are going to use another permutation t together with a set A of $\lfloor \frac{n}{2} \rfloor$ elements (which can be represented as the set of fixed points of yet another permutation).

If f is indeed connected, we name the elements of the universe $1, \dots, n$ such that, for every i , $f(i) = i + 1$ modulo n . We choose $A = \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$, and define $t(i) := 2i$ for $i \in A$, the values of $t(i)$ for $i \notin A$ are irrelevant. We note that the set $t(A)$ consists exactly of the even elements of $\{1, \dots, n\}$. In particular, at most one of x and $f(x)$ can be in $t(A)$, for every x .

Figure 2 illustrates the definition of t and A for $n = 9$.

Then the system (f, t, A) has the following first-order expressible properties.

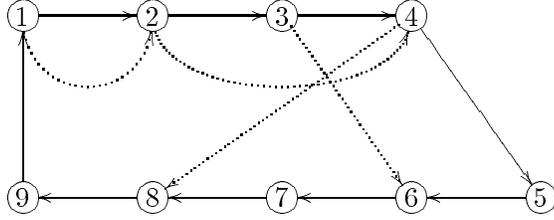


Figure 2: The definition of t and A for $n = 9$. $A = \{1, 2, 3, 4\}$, $a = 9$, f is indicated by solid edges, t by dotted edges (only for vertices in A).

- (1) there is exactly one $a \notin A$ with $f(a) \in A$;
- (2) for every $x \neq a$ it holds that $x \in t(A)$ if and only if $f(x) \notin t(A)$.
- (3) for every $x \in A$ it holds that $t(f(x)) = f(f(t(x)))$.
- (4) $t(f(a)) = f(f(a))$.

In order to show that (1)-(4) imply that f is connected, consider G_f , the graph of f . As f is a permutation, it consists of a number of disjoint cycles. Property (1) implies that, for all but two x , it holds that $x \in A \iff f(x) \in A$. Hence each of these cycles contains either only vertices in A (call such cycles A -cycles) or only vertices in \bar{A} , except for the cycle C_0 which contains a . (1) also implies that C_0 consists of one consecutive A -part (beginning with $f(a)$), followed by a consecutive \bar{A} -part (ending with a).

Properties (3) and (4) assure that t maps the A -part of C_0 to C_0 .

Let us now assume the existence of another cycle $C \neq C_0$. From (2) it follows that $t(A) \cap C \neq \emptyset$. Hence there are a cycle C' (either an A -cycle or C_0) and a $x \in C'$ such that $t(x) \in C$. (3) implies that $t(C') \subseteq C$, therefore $C' \neq C_0$. Applying (2) again, we can conclude that $|C| \geq 2|C'|$. But this means that for every $C \neq C_0$ there is a $C' \neq C_0$ with at most half as many elements, yielding a decreasing chain, hence no $C \neq C_0$ can exist.

□

From this lemma we get the following corollary.

5.4 Corollary

The set of connected undirected graphs is in $\exists Perm FO$.

Proof. Sekanina [Sek60] has shown that a graph G is connected if and only if its cube contains a hamiltonian cycle. In the cube of G two vertices x and y are adjacent iff their distance in G is at most 3.

Hence a graph G is connected, iff there exists a successor relation on the universe of G such that the distance between any two successive vertices (and between the maximum and minimum vertex) is at most 3. \square

6 Unary Functions are Stronger than Permutations

In this section we show that $\exists Perm FO$ is strictly included in $\exists UnF FO$.

Let $HalfEdge$ be the set of graphs in which the number of edges equals $\lfloor \frac{n^2}{2} \rfloor$, where n is the number of vertices and let $nEdge$ be the set of all graphs in which the number of edges equals the number of vertices. We show that $nEdge$ is in $\exists UnF FO$ but not in $\exists Perm FO$.

Remember that graphs are always directed.

As an intermediate step we first show that Duplicator has a winning strategy in the modified UF game (as defined in Section 2) on $HalfEdge$.

We make use of Ajtai's result [Ajt83] that first order logic fails dramatically in distinguishing between sets of even and odd size. If A is a σ -structure, T an additional relation symbol and φ a first order formula over $\sigma\langle T \rangle$, we write $S_\varphi^{even}(A)$ for the number of relations T with an even number of elements such that φ holds in $\langle A, T \rangle$ (analogously $S_\varphi^{odd}(A)$).

Ajtai formulated his theorem for unary relations, but it is easy to deduce a version for any arity from it. For our purposes we use the following.

6.1 Theorem ([Ajt83])

Let T be a binary relation symbol, φ a first order formula and $\epsilon > 0$. Then for all but finitely many n and every structure A of size n it holds that

$$|S_\varphi^{even}(A) - S_\varphi^{odd}(A)| \leq 2^{n^2 - n^{2-\epsilon}}.$$

We note that the proof of Ajtai's original result is rather involved.

From this theorem Ajtai concluded that the set of graphs with an even number of edges is not (weakly⁵) expressible in Monadic **NP** even in the presence of arbitrary built-in relations. It follows immediately that Duplicator has a winning strategy in the (Monadic **NP**) Ajtai-Fagin game over this set of graphs. The following theorem strengthens this result slightly with $HalfEdge$ instead of the set of graphs with an even number of edges. The proof is similar to the original one of Ajtai [Ajt83].

6.2 Theorem

Duplicator has a winning strategy in the modified UF game on $HalfEdge$.

Proof. Let $HalfEdge_n$ denote the set of graphs from $HalfEdge$ with universe $\{1, \dots, n\}$. Correspondingly, let $\overline{HalfEdge}_n$ be the set of graphs with

⁵Here weak expressibility essentially means that there exists a formula which expresses the property for infinitely many sizes.

universe $\{1, \dots, n\}$ that are not in HalfEdge . The proof is by contradiction. Let us assume that Spoiler has a winning strategy in the game on HalfEdge .

This means that there are k and l such that for every n and for every graph $G_1 \in \text{HalfEdge}_n$ there exist unary functions $\bar{f} = f_1, \dots, f_l$ such that for all graphs $G_2 \in \overline{\text{HalfEdge}_n}$ it holds that Spoiler has a winning strategy in the k -round game on $\langle G_1, \bar{f} \rangle$ and $\langle G_2, \bar{f} \rangle$, (i.e., $\tau_k(\langle G_1, \bar{f} \rangle) \neq \tau_k(\langle G_2, \bar{f} \rangle)$).⁶

The number of graphs in HalfEdge_n is $\binom{n^2}{\lfloor \frac{1}{2}n^2 \rfloor} > \frac{2^{n^2}}{n^2} = 2^{n^2 - 2 \log n}$, a number which is, for large n , much greater than the number of l -tuples of unary functions on $\{1, \dots, n\}$, which is $n^{ln} = 2^{ln \log n}$.

Hence, for every n , there is some tuple \bar{f} and a set $\mathcal{G} \subseteq \text{HalfEdge}_n$ of size at least $2^{n^2 - (ln+2) \log n}$ such that, for every $G_1 \in \mathcal{G}$ and $G_2 \in \overline{\text{HalfEdge}_n}$, Spoiler has a winning strategy in the k -round Ehrenfeucht game on $\langle G_1, \bar{f} \rangle$ and $\langle G_2, \bar{f} \rangle$.

That Spoiler has a winning strategy implies that there is a set A of k -types such that $\tau_k(\langle G_1, \bar{f} \rangle) \in A$, for all $G_1 \in \mathcal{G}$ and $\tau_k(\langle G_2, \bar{f} \rangle) \notin A$, for all $G_2 \in \overline{\text{HalfEdge}_n}$.

For every n , we fix such a set $A(n)$ of k -types. As there is only a finite number of k -types there exists a set A_0 such that $A_0 = A(n)$, for infinitely many n . Let θ be a first order formula such that $\tau_k(\langle G, \bar{f} \rangle) \in A_0$ if and only if $\langle G, \bar{f} \rangle \models \theta$ (see [EF95]). Hence for infinitely many n there is some \bar{f} such that

- for at least $2^{n^2 - (ln+2) \log n}$ graphs $G \in \text{HalfEdge}_n$ it holds that $\langle G, \bar{f} \rangle \models \theta$, but
- for all graphs $G \in \overline{\text{HalfEdge}_n}$ it holds that $\langle G, \bar{f} \rangle \not\models \theta$.

Hence we obtain: for infinitely many n there is a structure $G_n = \langle \{1, \dots, n\}, \bar{f} \rangle$ such that

- $S_\theta^{\text{even}}(G_n) \geq 2^{n^2 - (ln+2) \log n}$, but $S_\theta^{\text{odd}}(G_n) = 0$, if $\lfloor \frac{n}{2} \rfloor$ is even, and
- $S_\theta^{\text{odd}}(G_n) \geq 2^{n^2 - (ln+2) \log n}$, but $S_\theta^{\text{even}}(G_n) = 0$, if $\lfloor \frac{n}{2} \rfloor$ is odd.

This contradicts Theorem 6.1. □

Now we are ready to prove the separation between $\exists \text{Perm FO}$ and $\exists \text{UnF FO}$. It follows from the proof of Lemma 4.3 (c) that $\text{nEdge} \in \exists \text{UnF FO}$.

⁶For the definition of k -types $\tau_k(\cdot)$ refer to subsection 2.2.

It is important that the functions in the proof of Lemma 4.3 are allowed to map many vertices to the same vertex. We will show that, in general, this behaviour cannot be simulated by permutations. On the other hand it is easy to show by a similar proof that the restriction of nEdge to graphs with a fixed degree bound k is expressible by permutations.

6.3 Lemma

$\text{nEdge} \notin \exists \text{Perm FO}$.

Proof. Let a number, k , of rounds, and a number, l , of permutations be given and let $p := (2l)^{2^k+1}$. Theorem 6.2 says that Duplicator has a winning strategy in the modified UF game on HalfEdge . In particular, Duplicator wins, if Spoiler chooses p^{2l} unary functions and $(2^k + 1)p$ unary relations⁷. Let $H_1 \in \text{HalfEdge}$ be a graph with vertices $1, \dots, m$ that can be chosen by Duplicator according to this winning strategy. In particular, H_1 has $\lfloor \frac{m^2}{2} \rfloor$ edges.

Let $n := \lfloor \frac{m^2}{2} \rfloor$.

In the permutation game Duplicator chooses a graph G_1 with vertices $1, \dots, n$ which equals H_1 on $\{1, \dots, m\}$ and has no other edges. By definition $G_1 \in \text{nEdge}$.

Let the permutations $\bar{f} = f_1, \dots, f_l$ be chosen by Spoiler. With $N(H_1)$ we denote the 2^k -neighbourhood of H_1 in the structure $\langle G_1, \bar{f} \rangle$.

Claim Duplicator can define a graph G_2 on $\{1, \dots, n\}$ such that

- (1) G_2 has graph edges only on $\{1, \dots, m\}$ (we call this subgraph H_2);
- (2) H_2 (and therefore G_2) has a number of edges different from $n = \lfloor \frac{m^2}{2} \rfloor$.
- (3) Duplicator has a distance respecting winning strategy in the k -round Ehrenfeucht game on the substructures of $\langle G_1, \bar{f} \rangle$ and $\langle G_2, \bar{f} \rangle$ that are induced by $N(H_1)$ and $N(H_2)$ respectively. (In the following we use the notation $N(H_1)$ also for the substructure that is induced by the vertices of $N(H_1)$.)

It is most important for the following that in (3) both structures are equipped with the same permutations.

We encode the structures $N(H_1)$ into p^{2l} additional unary functions and $(2^k + 1)p$ additional unary relations on H_1 as follows.

⁷We can encode these relations by functions, so as to be able to apply Theorem 6.2. However, the proof will be more transparent, if we use relations here.

Because the number of vertices in $N(H_1)$ is at most pm (this is where we use that \bar{f} consists of *permutations*), there exists a function h which maps the vertices of $N(H_1)$ in a one to one manner to pairs (y, i) where $y \in H_1$ and $i \leq p$ is a natural number.

Let the unary relations A_{ij} , for every $i \leq p$ and every $j \leq 2^k$, be defined by

$$y \in A_{ij} \iff h(x) = (y, i) \text{ for some } x \in N(H_1) \text{ and } d(x, H_1) = j.$$

Finally, let the unary functions $g_{ji_1i_2}$ for every $j \leq l$ and $i_1, i_2 \leq p$ be defined as follows.

$$g_{ji_1i_2}(y_1) = y_2,$$

if for some $x_1, x_2 \in N(H_1)$ it holds that

- $f_j(x_1) = x_2$,
- $h(x_1) = (y_1, i_1)$,
- $h(x_2) = (y_2, i_2)$.

All other values of $g_{ji_1i_2}$ are defined arbitrarily. We will see below how the $g_{ji_1i_2}$ help to translate a winning strategy of Duplicator on H_1 and H_2 (with the $g_{ji_1i_2}$ and A_{ij}) into a winning strategy on $N(H_1)$ and $N(H_2)$ (without the $g_{ji_1i_2}$ and A_{ij}). (We note that for the following argument it does not hurt if these functions encode more information than \bar{f} . Notice also that these functions do not need to be permutations, because we only want to apply Theorem 6.2.)

By the choice of H_1 Duplicator can find H_2 such that

- H_2 has a number of edges which is different from n , and
- he has a k -round winning strategy on the structures $\langle H_1, \bar{A}, \bar{g} \rangle$ and $\langle H_2, \bar{A}, \bar{g} \rangle$.

It is easy to see that this winning strategy induces a distance respecting k -round winning strategy on $N(H_1)$ and $N(H_2)$. E.g., if Spoiler chooses a vertex $x_1 \in N(H_1)$ then the answer of Duplicator can be computed as follows:

Let $(y_1, i) = h(x_1)$. Let y_2 be the vertex which Duplicator would answer if Spoiler chose y_1 in the game on structures $\langle H_1, \bar{A}, \bar{g} \rangle$ and $\langle H_2, \bar{A}, \bar{g} \rangle$. As $h(x_1) = (y_1, i)$ it holds that $y_1 \in A_{ij}$, for some j , hence $y_2 \in A_{ij}$. Therefore, there exists an x_2 with $h(x_2) = (y_2, i)$ and $d(x_2, H_2) = j$. Duplicator chooses this x_2 . The definition of the $g_{ji_1i_2}$ assures that this gives rise to a winning strategy for Duplicator.

On the other hand⁸ $\langle G_1 - H_1, \bar{f} \rangle$ and $\langle G_2 - H_2, \bar{f} \rangle$ are of course isomorphic via an isomorphism which respects the distance from H_1 (resp. H_2).

By Theorem 2.6 it follows that Duplicator has a k -round winning strategy on G_1 and G_2 . \square

From this lemma we conclude

6.4 Theorem

$\exists Perm FO \subset \exists UnF FO$.

⁸Here only those "function edges" are considered, which have both vertices outside H_i .

7 Discussion

Our investigations have revealed a strict hierarchy within **BinNP**: quantification over successor relations, linear order relations, and partial order relations, respectively, gives us increasing expressive power. Furthermore, we showed for a number of other restricted classes of binary relations that quantification over each of these classes coincides with quantification over either one of the three classes of order relations.

Given these results, the following topics for further research suggest themselves:

- Identify other interesting classes of binary relations, and compare their expressive power with those of the ones considered here.
- Analyse the fine structure of our classes: how does quantification over k successor relations compare with quantification over m permutations, etc.
- Show for some concrete graph problem L that it is not in **BinNP**. Although we know that such an L must exist within **PSPACE**, no concrete example is known. An example L in **NP** or **coNP** would be of particular interest.

The latter problem is the most important open problem in this context. Even with our most powerful methods for proving non-expressibility, it seems that such a result is not possible at present. It was a combination of these tools which gave us the most interesting result of this paper, the separation of $\exists Succ FO$ from $\exists LinOrd FO$. We hope that further refinement of such methods will eventually lead to nonexpressibility proofs for **BinNP** and beyond.

References

- [AF90] M. Ajtai and R. Fagin. Reachability is harder for directed than for undirected finite graphs. *Journal of Symbolic Logic*, 55(1):113–150, 1990.
- [Ajt83] M. Ajtai. Σ_1^1 formulae on finite structures. *Ann. of Pure and Applied Logic*, 24:1–48, 1983.
- [Bal85] J.T. Baldwin. Definable second order quantifiers. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 445–477. Springer-Verlag, 1985.

- [BS85] J.T. Baldwin and S. Shelah. Second order quantifiers and the complexity of theories. *Notre Dame Journal of Formal Logic*, pages 229–302, 1985.
- [dR87] M. de Rougemont. Second-order and inductive definability on finite structures. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 33:47–63, 1987.
- [DR94] A. Durand and S. Ranavoson. First-order spectra with one binary predicate. In *CSL '94*, pages 177–189, 1994.
- [EF95] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer-Verlag, 1995.
- [Ehr61] A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fund. Math.*, 49:129–141, 1961.
- [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.
- [Fag75] R. Fagin. Monadic generalized spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:89–96, 1975.
- [Fag96] R. Fagin. Easier ways to win logical games. In *Proceedings of the DIMACS Workshop on Finite Models and Descriptive Complexity*. American Mathematical Society, 1996. Preliminary version appeared in 1993 IEEE Conference on Structure in Complexity Theory, pp. 19-30.
- [Fra54] R. Fraïssé. Sur quelques classifications des systèmes de relations. *Publ. Sci. Univ. Alger. Sér. A*, 1:35–182, 1954.
- [FSV95] R. Fagin, L. Stockmeyer, and M. Vardi. On monadic NP vs. monadic co-NP. *Information and Computation*, 120:78–92, 1995.
- [GO94] E. Grandjean and F. Olive. Monadic logical definability of NP-complete problems. In *Proc. 1994 of the Annual Conference of the EACSL*, pages 190–204, 1994. Extended version submitted.
- [Gra] E. Grandjean. personal communication.
- [Gra90] E. Grandjean. First-order spectra with one variable. *Journal of Computer and System Sciences*, 40:136–153, 1990.

- [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.
- [KS96] M. Kreidler and D. Seese. Monadic NP and built-in trees. presented at CSL 96, 1996.
- [Loe91] B. Loescher. Begründung, Verallgemeinerung und Anwendung der Ehrenfeucht-Spiele in der Relationentheorie von Roland Fraïssé. Informatik-bericht 2/91, Institut für Informatik, Universität Mainz, 1991.
- [Loe96] B. Loescher. One unary function says less than two in existential second order logic. 1996.
- [LST95] C. Lautemann, T. Schwentick, and D. Thérien. Logics for context-free languages. In *CSL '94*, pages 205–216, 1995.
- [Lyn82] J. F. Lynch. Complexity classes and theories of finite models. *Mathematical System Theory*, 15:127–144, 1982.
- [Lyn92] J. F. Lynch. The quantifier structure of sentences that characterize nondeterministic time complexity. *Computational Complexity*, 2:40–66, 1992.
- [Sch94] T. Schwentick. Graph connectivity and monadic NP. In *Proc. 35th IEEE Symp. on Foundations of Computer Science*, pages 614–622, 1994.
- [Sch95] T. Schwentick. Graph connectivity, monadic NP and built-in relations of moderate degree. In *Proc. 22nd International Colloq. on Automata, Languages, and Programming*, pages 405–416, 1995.
- [Sek60] M. Sekanina. On an ordering of the set of vertices of a connected graph. *Spisy Přírod. Fak. Univ. Brno*, pages 137–141, 1960.
- [She73] S. Shelah. There are just four possible second order quantifiers. *Israel Journal of Mathematics*, pages 282–300, 1973.
- [She86] S. Shelah. Classifying generalized quantifiers. In *Around Classification Theory of Models*, pages 1–46. Springer-Verlag, 1986. Springer Lecture Notes 1182.

- [Ten75] R. Tenney. Second-order Ehrenfeucht games and the decidability of the second-order theory of an equivalence relation. *Journal of the Australian Mathematical Society*, 20:323–331, 1975.
- [Tur84] G. Turán. On the definability of properties of finite graphs. *Discrete Mathematics*, 49:291–302, 1984.
- [WW90] R. Wilson and J. Watkins. *Graphs - An introductory approach*. Wiley, 1990.