Journal of Universal Computer Science, vol. 14, no. 10 (2008), 1654-1677
submitted: 2/11/07, accepted: 27/4/08, appeared: 28/5/08 © *J.UCS*

Expressibility in Σ_1^1

Walid Gomaa
(Alexandria University
Department of Computer and Systems Engineering
Alexandria - Egypt
wgomaa@alex.edu.eg)

Abstract: Inspired by Fagin's result that $NP = \Sigma_1^1$, we have developed a partial framework to investigate expressibility inside Σ_1^1 so as to have a finer look into NP. The framework uses interesting combinatorics derived from second-order Ehrenfeucht-Fraïssé games and the notion of game types. Some of the results that have been proven within this framework are: (1) for any k, divisibility by k is not expressible by a Σ_1^1 sentence where (1.i) each second-order variable has arity at most 2, (1.ii) the first-order part has at most 2 first-order variables, and (1.iii) the first-order part has quantifier depth at most 3, (2) adding one more first-order variable makes the same problem expressible, and (3) inside this last logic the parameter k creates a proper hierarchy with varying the number of second-order variables.

Key Words: Ehrenfeucht-Fraïssé games, divisibility, expressibility, first-order, second-order
Category: F.4, F.1.3

1 Introduction

The birth of finite model theory is often identified with Trakhtenbrot's result from 1950 stating that logical validity over finite models is not recursively enumerable, that is, completeness fails over finite structures [Libkin 2004]. In 1974, R. Fagin proved his celebrated theorem that NP can be exactly captured by existential second-order logic [Fagin 1974]. This opened up a new area of research called descriptive complexity. It is a branch of complexity theory that views the hardness of problems in terms of the complexity of their logical expressiveness such as the number of object variables, quantifier depth, type, and alternation, sentences length (finite/infinite), etc.

Fagin's result has been generalized in [Stockmeyer 1977] to show that the whole of the polynomial hierarchy is exactly captured by second-order logic.

Inspired by the above results we have developed a partial framework to investigate expressibility inside Σ_1^1. Currently this framework encompasses sublogics of Σ_1^1 defined as follows.

Definition 1.1 *1.* Existential second-order logic*, or Σ_1^1, is defined to be the class of sentences of the form*

$$\exists X_1 \ldots \exists X_l \varphi \tag{1.1.1}$$

where the X_i's are second-order relational variables of arbitrary finite arities and φ is a first-order sentence.

2. *Let $mon\Sigma_1^1$ be the sublogic of Σ_1^1 obtained by restricting the arities of the X_i's to be at most 1 (hence the prefix mon).*

3. *Let $bin\Sigma_1^1$ be the sublogic of Σ_1^1 obtained by restricting the arities of the X_i's to be at most 2 (hence the prefix bin). Note that any sentence in $bin\Sigma_1^1$ is equivalent to a sentence of the form*

$$\exists R_1 \ldots \exists R_n \exists S_1 \ldots \exists S_m \varphi \tag{1.1.2}$$

where the R_i's and the S_i's are binary and unary second-order variables respectively. For simplicity of discussion we will assume that $bin\Sigma_1^1$ consists exactly of sentences of the form (1.1.2).

4. *Let $bin\Sigma_1^1(p,r)$ be the sublogic of $bin\Sigma_1^1$ obtained by restricting φ to have at most p first-order variables and quantifier depth at most r. Define $mon\Sigma_1^1(p,r)$ similarly.*

Within this framework we plan to study expressibility of some number-theoretic properties. In this paper we started by studying divisibility.

Definition 1.2 *For every integer $k \geq 2$, let DIV_k denote the problem of deciding whether a positive integer is divisible by k. Let $\overline{DIV_k}$ denote the complement problem, that is non-divisibility by k.*

Example 1.3 *Consider DIV_2 which is the famous $EVEN$ problem. It was shown that $EVEN$ is not expressible in first-order logic (e.g., see [Libkin 2004]). However, $EVEN$ can be expressed by the following $bin\Sigma_1^1$ sentence.*

$$\sigma \overset{def}{=} \exists R \left(\varphi_1(R) \wedge \varphi_2(R) \wedge \varphi_3(R)\right) \tag{1.3.1}$$

where

$$\varphi_1(R) \overset{def}{=} \forall x \neg R(x,x)$$

$$\varphi_2(R) \overset{def}{=} \forall x \forall y \left(R(x,y) \longleftrightarrow R(y,x)\right)$$

$$\varphi_3(R) \overset{def}{=} \forall x \exists y \left(R(x,y) \wedge \forall z \left(R(x,z) \longrightarrow z = y\right)\right)$$

Notice that σ defines the class of finite simple graphs with isolated edges (1-regular graphs). The number of vertices in these graphs must be even.

Notation 1.4 *Throughout the remaining part of this paper if the variable k is occurring free (unquantified) in a result, this indicates that the result holds for every value of k.*

Assuming the empty vocabulary we will prove the following results:

1. $DIV_k, \overline{DIV_k}$ are neither in $mon\Sigma_1^1$ nor in $mon\Pi_1^1$
2. $DIV_k \notin bin\Sigma_1^1(1,r)$ for any r
3. $DIV_k \notin bin\Sigma_1^1(2,2)$ and $DIV_k \notin bin\Sigma_1^1(2,3)$
4. $DIV_k \in bin\Sigma_1^1(3,3)$. More specifically, given $\Gamma \subseteq bin\Sigma_1^1(3,3)$ where every $\sigma \in \Gamma$ has at most l binary variables then $DIV_k \in \Gamma$ for every $k \leq (4^l - 1)$. Furthermore, $DIV_k \in \Gamma$ for only finitely-many k, hence DIV_k creates a proper hierarchy inside the logic $bin\Sigma_1^1(3,3)$.
5. An immediate consequence of the above is that $mon\Sigma_1^1 \subset bin\Sigma_1^1$.
6. $DIV_k \notin bin\Sigma_1^1$ when the sizes of the interpretations of the binary variables are bounded from above by some linear function of the size of the universe.

The main tool used in this paper to obtain the above results is *Ehrenfeucht-Fraïssé (EF)* games. A particular version of these games characterizes expressibility in some corresponding logical formalism. In this paper we extend the traditional first-order version of the game to a second-order one that matches our framework. Based on the *locality* of strategies in first-order games, we define the notion of *game types* which is used to characterize the winner in the first-order phase of the extended game. Hence, no need to actually play it which often involves complicated combinatorial arguments. Our definition of game types is inspired by that given in [Koucký et al. 2006]. This latter article uses game types in a combinatorially involved argument (employing the switching lemma) to give an EF proof of the fact that $PARITY$ is not in the circuit class AC^0. Although this result had already been known, the motivation was to give an easier proof using EF games. However, it turned out that their EF argument resembles the classical one. Their context in general is different from ours (circuit-based vs. algorithmic-based models of computation).

T. Schwentick [Schwentick 1995] gives a similar definition of game types. In this article the author focuses on the specific problem of expressibility of graph connectivity. It represented a contribution towards solving the still open conjecture that graph connectivity is not expressible in $mon\Sigma_1^1$ even in the presence of arbitrary built-in relations. Game types are defined for the class of *Gaifman graphs* of graphs with built-in relations with degree at most $n^{o(1)}$. Game types are then used in a rather simple combinatorial argument to show that graph connectivity is not expressible in $mon\Sigma_1^1$ even in the presence of such built-in relations. The main part of his argument is the proof of what he calls the *extension theorem*. This theorem allows us to extend, under certain circumstances, a winning strategy of the duplicator on small local parts of the given structures to a global winning strategy. This can be thought of as a sort of homogeneity property of the given structures. More elaborate treatment of this work, in particular the extension theorem and its applications, can be found in [Schwentick 1996].

Inexpressibility of graph connectivity in $mon\Sigma_1^1$ is further investigated in [Kreidler et al. 1997] and [Kreidler et al. 1998]. In the former $mon\Sigma_1^1$ is enriched with a built-in relation that has the shape of a forest. The authors show that this new logic is not powerful enough to capture graph connectivity. Their technique is based on finding a winning strategy for the duplicator by searching for a vertex with a large degree in the built-in relation such that sufficiently large number of its neighbors are game-theoretically indistinguishable (have the same game types). This stands in contrast to the traditional Hanf's method which looks for vertices that have identical game types but lie sufficiently far away from each other. Both methods in turn stand in contrast to ours which completely depends on the notion of game types and the ability of the duplicator to choose the two structures and to color her own in such a way both of them have the same game type. In [Kreidler et al. 1998], the same authors took more advanced step by showing the inexpressibility of graph connectivity in $mon\Sigma_1^1$ even in the presence of a built-in relation of arbitrary degree that does not have the complete graph K_m as a minor for arbitrary but fixed m. However, the argument in this article is much more combinatorially involved.

Padding techniques are well-known from computational complexity theory. T. Schwentick [Schwentick 1997] applies the notion of padding to the context of descriptive complexity, where it can be viewed as a logical resource that adds expressibility power to the formalism under consideration. A graph H is a *padded version* of G if H consists of an isomorphic copy of G and some additional isolated vertices. The main idea is that these isolated vertices can be used, for example, in a $mon\Sigma_1^1$ sentence to enumerate the elements of the original graph (with the help of built-in relations). Unlike our approach which investigates Σ_1^1 by getting separation results, hence proper hierarchy of sublogics of Σ_1^1 defined by restricting the arities of the second-order variables, Schwentick uses padding and built-in relations to reduce Σ_1^1 (hence NP) to $mon\Sigma_1^1$. The notion of *weak expressibility* is defined to denote expressibility in the presence of sufficiently large padding. His approach implies that to negatively answer the $NP = coNP$? question (consequently $P = NP$?) we need to find a $coNP$ set that is not weakly expressible inside $mon\Sigma_1^1$. In other words we need to find a $coNP$ set that is not expressible in ($mon\Sigma_1^1$ + *padding+built-in relations*). The article classifies different variations of $mon\Sigma_1^1$ based on how padding affects the expressibility power of the underlying logic. Second-order EF games (Ajtai-Fagin games) are employed in a direct way in this work.

Section 2 gives axiomatization of a type of colored graphs which will be the main structures throughout the rest of the paper. Section 3 introduces the

Ehrenfeucht-Fraïssé (EF) game. We define a specific version called $bin\Sigma_1^1(p,r)$-game which will be applied to study the expressibility of DIV_k in $bin\Sigma_1^1(p,r)$. In Section 4 we prove that DIV_k and its complement are neither in $mon\Sigma_1^1$ nor in $mon\Pi_1^1$. In Section 5, the notion of game types is defined which is a combinatorial concept based on the locality of first-order logic, it is used to provide necessary and sufficient conditions for winning EF-games without actually playing them. In Sections 6 through 9 we prove the other expressibility results mentioned above. Section 10 concludes the paper with some insights for future work.

2 Colored Graphs

We study expressibility by sentences of the following form

$$\exists R_1 \ldots \exists R_{n'} \exists S_1 \ldots \exists S_{m'} \varphi \tag{2.0.1}$$

where the R_i's and S_i's are binary and unary second-order variables respectively and φ is a first-order sentence whose vocabulary is exactly the R_i's and the S_i's.

Such sentences will be modeled by first-order structures of the following form

$$G' = (V, U_1, \ldots, U_{m'}, E_1, \ldots, E_{n'})$$

V is a finite set of vertices. The U_i's are unary relations over V, these represent the interpretations of the S_i's in (2.0.1). The E_i's are binary relations over V which represent the interpretations of the R_i's. Consider a vertex $u \in V$. For each $j = 1, \ldots, m'$, either $u \in U_j$ or $u \notin U_j$. Hence, there are a total of $m = 2^{m'}$ different membership possibilities of u in all of the $U_j's$. Similarly, consider a pair of vertices $u, v \in V$. For each $k = 1, \ldots, n'$, exactly one of the following cases must hold: (1) $(u,v) \in E_k$ and $(v,u) \notin E_k$, (2) $(u,v) \notin E_k$ and $(v,u) \in E_k$, (3) $(u,v) \in E_k$ and $(v,u) \in E_k$, and (4) $(u,v) \notin E_k$ and $(v,u) \notin E_k$. Hence, there are a total of $n = 4^{n'}$ different membership possibilities of the pair u, v in all of the $E_k's$. It is simpler then to rethink of the structure G' as a graph G where the vertices are m-colored and the edges are n-colored. We denote this graph

$$G = (V, C_1, \ldots, C_m, D_1, \ldots, D_n) \tag{2.0.2}$$

where G is a complete undirected graph, each vertex has a self-edge, each C_i is a unary relation (for a vertex color), and each D_i is a binary relation (for an edge color). G must satisfy the following axioms:

$$\forall u \bigvee_{1\le i\le m} C_i(u)$$

$$\forall u(C_i(u) \longrightarrow \bigwedge_{1\le j\le m, j\ne i} \neg C_j(u)), \qquad for\ every\ 1\le i\le m$$

$$\forall u \bigvee_{1\le i\le n} D_i(u,u)$$

$$\forall u(D_i(u,u) \longrightarrow \bigwedge_{1\le j\le n, j\ne i} \neg D_j(u,u)), \qquad for\ every\ 1\le i\le n$$

$$\forall u\forall v \bigwedge_{1\le i\le n} (D_i(u,v) \longleftrightarrow D_i(v,u))$$

$$\forall u\forall v(u\ne v \longrightarrow \bigvee_{1\le i\le n} D_i(u,v))$$

$$\forall u\forall v(u\ne v \wedge D_i(u,v) \longrightarrow \bigwedge_{1\le j\le n, j\ne i} \neg D_j(u,v)) \qquad for\ every\ 1\le i\le n$$

The first two axioms indicate that every vertex u must have a unique color from the color list $C_1,\dots,C_m$. The third and fourth axioms indicate that the self-edge of every vertex u must have a unique color from the color list $D_1,\dots,D_n$. The last three axioms indicate that the graph is undirected and every edge (u,v) must have a unique color from the color list $D_1,\dots,D_n$. It can easily be observed that the axioms for self-edges can be combined into the last two axioms, however, they are separated since for the rest of the paper they are treated differently from the other edges.

Notation 2.1 *1. Let $\mathcal{G}_{m,n}$ be the class of graphs with exactly m vertex colors and n edge colors. Let $\mathcal{G} = \bigcup_{m,n} \mathcal{G}_{m,n}$.*

2. Let $\mathscr{C}$ be the set of m vertex colors and let $\mathscr{D}$ be the set of n edge colors.

3 Ehrenfeucht-Fraïssé Games

Ehrenfeucht-Fraïssé (EF) games are used to characterize expressibility in some logical formalism. They were invented by Ehrenfeucht [Ehrenfeucht 1961] based on work done by Fraïssé [Fraïssé 1954]. In our context we apply it to study expressibility in $bin\Sigma_1^1(p,r)$ for positive integers p and r.

3.1 Pebble first-order *EF*-games

In this section we briefly review pebble first-order EF-games. A pebble first-order EF-game (for detailed discussion see [Libkin 2004, Immerman 1998]) is played over two structures of the same kind, for example two linear orderings. There are two players: the spoiler denoted by $\mathcal{S}$ and the duplicator denoted by $\mathcal{D}$. The game has two parameters: the number of rounds r and the number of pebbles $p \leq r$. Intuitively, the goal of $\mathcal{S}$ is to show that the two structures can be distinguished in at most r steps using only p pebbles, whereas $\mathcal{D}$ wants to show that this can not be done.

Definition 3.1 (Partial isomorphism) *Let $\mathcal{A}$ and $\mathcal{B}$ be two first-order structures with vocabulary τ. Assume $\bar{a} = \langle a_1, \ldots, a_n \rangle \in A^n$ and $\bar{b} = \langle b_1, \ldots, b_n \rangle \in B^n$. We say that there is a* partial isomorphism *from $\bar{a}$ onto $\bar{b}$ if for every m, for every first-order quantifier-free formula $\varphi(x_1, \ldots, x_m)$ over τ, and for every $\{i_1, \ldots, i_m\} \subseteq \{1, \ldots, n\}$ the following holds*

$$\mathcal{A} \models \varphi(a_{i_1}, \ldots, a_{i_m}) \iff \mathcal{B} \models \varphi(b_{i_1}, \ldots, b_{i_m})$$

Given $\mathcal{A}$ and $\mathcal{B}$, the pebble EF-game goes as follows. The players start the game each having a fixed number of p pebbles. At each round $\mathcal{S}$ does the following: (i) she chooses an element x from one of the two structures and (ii) then she either removes a pebble that has been placed on a previously chosen element and places it on x or placing a new pebble, if she still has any, on x. $\mathcal{D}$ then responds to the challenge by choosing an element from the other structure and does the same pebbling so as to preserve the partial isomorphism among the pebbled elements chosen so far from $\mathcal{A}$ and $\mathcal{B}$. At the beginning the pebbles are not placed on any elements (we can assume having extra pebbles always placed on the distinguished elements of the structure such as the group identity, even before the game starts). Assume that at the end of the game p pebbles are placed on $\bar{a} = \langle a_1, \ldots, a_p \rangle$ from the structure $\mathcal{A}$ and correspondingly p pebbles are placed on $\bar{b} = \langle b_1, \ldots, b_p \rangle$ from the structure $\mathcal{B}$. Notice that these are in general subsets of the elements chosen during the course of the game. $\mathcal{D}$ wins the game if $\bar{a}$ and $\bar{b}$ are partially isomorphic, otherwise $\mathcal{S}$ wins.

Pebble first-order EF-games characterize expressibility in bounded variable logic. Let $\mathcal{L}^p$ denote first-order logic with at most p variables. For a formula $\varphi \in \mathcal{L}^p$, let $qr(\varphi)$ denote the quantifier rank (depth) of φ.

Definition 3.2 (Elementary equivalence) *Assume $\mathcal{A}$ and $\mathcal{B}$ are two structures over a vocabulary τ. We say that $\mathcal{A}$ and $\mathcal{B}$ are* (p, r)-elementarily equivalent*, denoted by $\mathcal{A} \equiv_r^p \mathcal{B}$ if and only if for every sentence $\varphi \in \mathcal{L}^p$ such that $qr(\varphi) \leq r$ we have*

$$\mathcal{A} \models \varphi \iff \mathcal{B} \models \varphi \tag{3.2.1}$$

The following theorem gives the relationship between pebble games and expressibility in $\mathcal{L}^p$.

Theorem 3.3 *The following are equivalent:*

i. $\mathcal{A} \equiv_r^p \mathcal{B}$

ii. $\mathcal{D}$ *has a winning strategy in the pebble first-order* EF*-game over* $\mathcal{A}$ *and* $\mathcal{B}$ *with* r*-rounds and* p*-pebbles*

This theorem basically says that no sentence in $\mathcal{L}^p$ of quantifier rank at most r can distinguish $\mathcal{A}$ and $\mathcal{B}$ if and only if the duplicator has a winning strategy in the EF-game over $\mathcal{A}$ and $\mathcal{B}$ with r rounds and p pebbles.

3.2 Second-order *EF*-games

As seen above the first-order game is played over two structures that are fixed apriori. In contrast the second-order game is played over a class of structures and consists of two phases: (i) the second-order phase played over a class of structures $\mathscr{K}$ where the duplicator gets to choose two structures $\mathcal{A} \in \mathscr{K}$ and $\mathcal{B} \in \overline{\mathscr{K}}$ (the complement of $\mathscr{K}$) and (ii) the first-order phase which is the regular pebble first-order game played over $\mathcal{A}'$ and $\mathcal{B}'$ where $\mathcal{A}'$ and $\mathcal{B}'$ are expansions of $\mathcal{A}$ and $\mathcal{B}$ as described below. These games are used to study expressibility in second-order logic.

The second-order game was introduced by Fagin in [Fagin 1975] and then modified in [Ajtai et al. 1990] to what is called the *Ajtai-Fagin* game (also called *monadic* Σ_1^1 game). In our context we slightly modify the Ajtai-Fagin game to a new game we call $bin\Sigma_1^1(p, r)$. The new game has four parameters m, n, p, and r and has the following rules.

1. $\mathcal{D}$ selects a member $\mathcal{A} \in \mathscr{K}$.
2. Using the domain of $\mathcal{A}$ as a set of vertices, $\mathcal{S}$ forms a complete undirected graph in which each vertex has a self-edge.
3. $\mathcal{S}$ colors the vertices using colors from $\mathscr{C}$ such that each vertex has exactly one color. She then colors the edges using colors from $\mathscr{D}$ such that each edge has exactly one color. Let $\mathcal{A}'$ be the new expanded colored structure.
4. $\mathcal{D}$ selects a member $\mathcal{B} \in \overline{\mathscr{K}}$.
5. Using the domain of $\mathcal{B}$ as a set of vertices, $\mathcal{D}$ forms a complete undirected graph with each vertex has a self-edge.
6. $\mathcal{D}$ colors the vertices from $\mathscr{C}$ such that each vertex has exactly one color. She then colors the edges from $\mathscr{D}$ such that each edge has exactly one color. Let $\mathcal{B}'$ be the new expanded colored structure.

7. $\mathcal{S}$ and $\mathcal{D}$ play a pebble first-order game over $\mathcal{A}'$ and $\mathcal{B}'$ with parameters r rounds and p pebbles.

This new game is used to study expressibility in $bin\Sigma_1^1(p,r)$. The relation is indicated in the following theorem.

Theorem 3.4 *Let $\mathscr{K}$ be a class of structures of the same vocabulary. Then $\mathscr{K}$ is $bin\Sigma_1^1(p,r)$ if and only if there are positive integers m, n, p and r such that $\mathcal{S}$ has a winning strategy in the $bin\Sigma_1^1(p,r)$-game with parameters m, n, p and r.*

Proof. The proof is very similar to that of Theorem 4.5 in [Ajtai et al. 1990]. The only difference is that both players now have the additional ability of coloring the edges as well as the vertices. This should make the game harder for $\mathcal{S}$ to win. However, if $\mathscr{K}$ is expressible in $bin\Sigma_1^1(p,r)$, then the defining sentence should indicate to $\mathcal{S}$ how to color the vertices and the edges in addition to a winning strategy in the first-order phase of the game.

Remark 3.5 *1. If the coloring is restricted to the vertices (no edge coloring), then we would call the resulting game $mon\Sigma_1^1(p,r)$, this is actually a pebbled version of the Ajtai-Fagin game.*

2. *In the definition of the $bin\Sigma_1^1(p,r)$-game, the ordering of the coloring of the vertices and/or edges (by either of the players) does not matter since the ordering of the corresponding second-order existential quantifiers is irrelevant as long as it does not alternate with universal quantifiers.*

3. *Notice that in the rules of the $bin\Sigma_1^1(p,r)$-game, the spoiler has to color the vertices and the edges of $\mathcal{A}$ before she knows what the other structure $\mathcal{B}$ is or how it will be colored by the duplicator. However, this does not make the game harder for her since if $\mathscr{K} \in bin\Sigma_1^1(p,r)$, then the coloring is predetermined completely by the sentence that defines $\mathscr{K}$.*

4. *In the following discussion we will always assume, unless otherwise stated, classes of structures over the empty vocabulary (the base language does not contain any non-logical symbols) so the structure is just a domain of elements; however, relations are defined over the domains during the course of the second-order EF-game. More specifically, the pebble first-order games are played over structures in $\mathcal{G}$.*

4 $\mathbf{DIV_k}, \overline{\mathbf{DIV_k}} \notin \mathbf{mon}\boldsymbol{\Sigma}_\mathbf{1}^\mathbf{1}\mathbf{(p,r)}$

Theorem 4.1 *$DIV_k \notin mon\Sigma_1^1(p,r)$ for any positive integers p and r.*

Proof. We will show that for large enough graphs $\mathcal{D}$ has a winning strategy in the $mon\Sigma_1^1(r,r)$-game. Fix $k \geq 2$. Assume m vertex colors. $\mathcal{D}$ starts by choosing a graph G such that $|G| \pmod k = 0$ and $|G| \geq mr$. $\mathcal{S}$ then colors the vertices of G using the given m colors. By the pigeonhole principle there must be at least r vertices having the same color $c \in \mathscr{C}$, let Γ be the set of all such vertices. $\mathcal{D}$ then chooses a graph $G' = (G \cup \{w\})$ with a new vertex w and does the following: (i) color $G \subseteq G'$ exactly as $\mathcal{S}$ did and (ii) color w with c. Let $\Gamma' = (\Gamma \cup \{w\})$. Now the first-order phase of the EF-game with r rounds. Assume the $(i+1)^{st}$ round of the game $(i+1 \leq r)$ and assume $\langle u_1, \ldots, u_i\rangle \subseteq G$ and $\langle v_1, \ldots, v_i\rangle \subseteq G'$ have been chosen such that for every $1 \leq j \leq i$, u_j and v_j have exactly the same color. Assume $\mathcal{S}$ chooses $u_{i+1} \in G$. If $u_{i+1} \notin \Gamma$, then $\mathcal{D}$ responds with the corresponding vertex in G' $(\notin \Gamma')$. If $u_{i+1} \in \Gamma$ then

- if $u_{i+1} = u_j$ for some $j \leq i$, then $\mathcal{D}$ responds with v_j,
- otherwise $\mathcal{D}$ responds with an arbitrary $v_{i+1} \in \Gamma'$ that has not been chosen before, this is possible since $|\Gamma'| \geq r$.

The case when $\mathcal{S}$ chooses $v_{i+1} \in G'$ is symmetric.

Theorem 4.2 $\overline{DIV_k} \notin mon\Sigma_1^1(p,r)$ *for any positive integers* p *and* r.

Proof. The proof is very similar to that of Theorem 4.1. $\mathcal{D}$ starts the game by choosing a graph G such that $|G| \pmod k \neq 0$ and $|G| \geq mr$. $\mathcal{S}$ does her coloring and then $\mathcal{D}$ responds by choosing a graph $G' = (G \cup W)$, where W is a new set of vertices such that $|G'| \pmod k = 0$. $\mathcal{D}$ colors all the vertices of W with c and let $\Gamma' = (\Gamma \cup W)$. The game then proceeds exactly as in Theorem 4.1.

Corollary 4.3 *1.* $DIV_k \notin mon\Pi_1^1$

2. $\overline{DIV_k} \notin mon\Pi_1^1$

5 Game Types

The definition of game types given in this section is inspired by a similar one given in [Koucký et al. 2006].

Definition 5.1 (Isomorphism types) *Let* $u, v, w \in G \in \mathcal{G}$.

1. *Define the* isomorphism type *of* u *in* G *as*

$$I(u;G) = \langle c, d\rangle\,, \qquad c \in \mathscr{C} \text{ and } d \in \mathscr{D} \tag{5.1.1}$$

where c *is the color of* u *and* d *is the color of its self-edge.*

2. *Define the* isomorphism type *of the pair u, v in G as*

$$I(u, v; G) = \langle I(u; G), I(v; G), d, eq(u, v)\rangle , \qquad d \in \mathscr{D} \tag{5.1.2}$$

   where d is the color of the edge (u, v) and $eq(u, v)$ is true if they are the same vertex otherwise false.

3. *Define the* isomorphism type *of the triple u, v, w in G as*

$$\begin{aligned} I(u, v, w; G) = \langle & I(u; G), I(v; G), I(w; G), \\ & I(u, v; G), I(u, w; G), I(v, w; G), eq(u, v), \\ & eq(u, w), eq(v, w)\rangle \end{aligned} \tag{5.1.3}$$

Remark 5.2 *The isomorphism type of any set of vertices corresponds to the first-order quantifier-free type of these vertices in G over the empty set of parameters.*

Definition 5.3 (Game types) *Let $u \in G \in \mathcal{G}$.*

1. *Define the* $(1, r)$-game type of u inside G *as*

$$\zeta_{1,r}(u; G) = I(u; G) \tag{5.3.1}$$

2. *Define the* $(2, r)$-game type of u inside G *inductively as*

$$\begin{aligned} &\zeta_{2,1}(u; G) = I(u; G) \\ &\zeta_{2,r}(u; G) = \langle I(u; G), \{\langle I(u, v; G), \zeta_{2,r-1}(v, G)\rangle : v \in G\}\rangle \end{aligned} \tag{5.3.2}$$

3. *Define the* $(3, r)$-game type of u inside G *inductively as*

$$\begin{aligned} &\zeta_{3,1}(u; G) = I(u; G) \\ &\zeta'_{3,1}(u, v; G) = I(u, v; G) \\ &\zeta'_{3,r}(u, v; G) = \langle I(u, v; G), \{\langle I(u, v, w; G), \zeta_{3,r-1}(w; G)\rangle : w \in G\}\rangle \\ &\zeta_{3,r}(u; G) = \langle I(u; G), \{\langle I(u, v; G), \\ &\qquad\qquad \zeta'_{3,r-1}(u, v; G), \zeta_{3,r-1}(v; G)\big\rangle : v \in G\big\}\big\rangle \end{aligned} \tag{5.3.3}$$

   where ζ' is a helper function and can be thought of as the game type of edges.

4. *For every $1 \leq p \leq 3$ define the* (p, r)-game type *of G as*

$$\zeta_{p,r}(G) = \{\zeta_{p,r}(u; G) \colon u \in G\} \tag{5.3.4}$$

Remark 5.4 *The (p, r)-game type of a vertex u corresponds to the first-order type of u in G over the empty set of parameters where every formula in that type has at most p variables and has quantifier rank at most r.*

The intuition behind these definitions of isomorphism and game types is the following: given $G, G' \in \mathcal{G}_{m,n}$ and given $u \in G, v \in G'$ such that $\zeta_{p,r}(u; G) = \zeta_{p,r}(v; G')$, then $\mathcal{D}$ has a winning strategy in the r-round first-order game with p pebbles which starts by placing pebbles on u and v. One can see this by induction, as $\mathcal{D}$ can maintain the invariant that the corresponding pebbled vertices have always the same game type [Koucký et al. 2006]. If furthermore we have the stronger assumption that $\zeta_{p,r}(G) = \zeta_{p,r}(G')$, then $\mathcal{D}$ can always win no matter how the game starts.

The following proposition from [Koucký et al. 2006] states the relationship between game types and first-order expressibility.

Proposition 5.5 *Assume $G, G' \in \mathcal{G}_{m,n}$. Then $\zeta_{p,r}(G) = \zeta_{p,r}(G')$ if and only if for every first-order sentence $\sigma \in \mathcal{L}^p$ such that $qr(\sigma) \leq r$ it is the case that $G \models \sigma \iff G' \models \sigma$.*

Notation 5.6 *1. We will omit the argument G from isomorphism types and game types when understood from the context.*

2. Fix m and n for vertex and edge colors respectively. Let $\Lambda(p, r; G)$ denote the maximum number of possible (p, r)-game types of graphs in $\mathcal{G}_{m,n}$ and let $\Lambda(p, r; u)$ denote the maximum number of possible (p, r)-game types of vertices in such graphs.

6 $\mathbf{DIV_k \notin bin\Sigma_1^1(1, r)}$ and $\mathbf{DIV_k \notin bin\Sigma_1^1(2, 2)}$

We show that $DIV_k \notin bin\Sigma_1^1(1, r)$ and $DIV_k \notin bin\Sigma_1^1(2, 2)$ by looking at the $(1, r)$- and $(2, 2)$-game types of graphs in $\mathcal{G}$.

Lemma 6.1 *Assume m vertex colors and n edge colors. Then $\Lambda(1, r; u) \leq mn$ and $\Lambda(1, r; G) \leq 2^{mn}$.*

Proof. Assume some vertex u. From Definition 5.3 we need only to count the number of isomorphism types of u which is at most mn. Since the game type of any $G \in \mathcal{G}_{m,n}$ is determined by the game types of its single vertices, then $\Lambda(1, r; G) \leq 2^{mn}$ (counting all possible subsets of game types of single vertices).

Lemma 6.2 *Let $G \in \mathcal{G}_{m,n}$. Then there exists $G' \in \mathcal{G}_{m,n}$ such that $|G'| = |G| + 1$ and $\zeta_{1,r}(G') = \zeta_{1,r}(G)$.*

Proof. Choose an arbitrary $u \in G$. Add to G a new vertex v, color it and its self-edge exactly as u's, and color its edges to the vertices of G arbitrarily. Let G' be the new graph. Clearly, $\zeta_{1,r}(G') = \zeta_{1,r}(G)$.

As a direct consequence of this lemma and Proposition 5.5 we have the following inexpressibility result.

Theorem 6.3 $DIV_k \notin bin\Sigma_1^1(1,r)$

Next we consider $(2,2)$-game types.

Lemma 6.4 *Assume m vertex colors and n edge colors. Then $\Lambda(2,2;u) \leq (mn)2^{(mn^2)}$ and $\Lambda(2,2;G) \leq 2^{\Lambda(2,2;u)}$.*

Proof. Given a vertex u, the $(2,2)$-game type of u is determined by: (i) its isomorphism type, (ii) the isomorphism type of any other vertex v, and (iii) the isomorphism type of the edge (u,v). There are (mn) possible vertex and self-edge colors for u, (mn) vertex and self-edge colors for v, and n possible colors for the edge (u,v). Hence there are at most a total of (mn^2) possible combinations of colors for v and (u,v) of which there are at most 2^{mn^2} possible subsets that can be associated with u. Therefore, $\Lambda(2,2;u) \leq (mn)2^{mn^2}$. As mentioned above the game type of any $G \in \mathcal{G}_{m,n}$ is determined by the set of game types of its single vertices, hence $\Lambda(2,2;G) \leq 2^{\Lambda(2,2;u)}$.

Lemma 6.5 *Let $G \in \mathcal{G}_{m,n}$. Assume $|G| > \Lambda(2,2;u)$. Then there exists $G' \in \mathcal{G}_{m,n}$ such that $|G'| = |G| + 1$ and $\zeta_{2,2}(G') = \zeta_{2,2}(G)$.*

Proof. From Lemma 6.4, there must be $u_1, u_2 \in G$ that have the same $(2,2)$-game type. Add to G a new vertex v, color it and its self-edge exactly as u_1. Connect v to every vertex in G. For every $w \in G$ such that $w \neq u_1$, use the color of the edge (u_1,w) to color the edge (v,w). Use the color of the edge (u_1,u_2) to color (v,u_1). Let G' be the new graph. It is easy to check that $\zeta_{2,2}(v;G') = \zeta_{2,2}(u_1;G)$ and for every $w \in G$, $\zeta_{2,2}(w;G) = \zeta_{2,2}(w;G')$. Hence, $\zeta_{2,2}(G') = \zeta_{2,2}(G)$.

As a direct consequence of this lemma we have the following inexpressibility result.

Theorem 6.6 $DIV_k \notin bin\Sigma_1^1(2,2)$

7 $\mathbf{DIV_k \notin bin\Sigma_1^1(2,3)}$

We show that $DIV_k \notin bin\Sigma_1^1(2,3)$ by looking at the $(2,3)$-game types of graphs in $\mathcal{G}$.

Remark 7.1 *Assume $G \in \mathcal{G}$ and let $u \in G$. Then $\zeta_{2,3}(u;G)$ can be characterized by the set of all paths in G of length 2 starting from u. This includes paths of the form uvu (going from u to v then back to u). Actually as we will see below these* *latter kind of paths is the main reason for the inexpressibility in* $bin\Sigma^1_1(2,3)$*. Given one such path uvw (two or all vertices may be identical) we will represent it by the tuple*

$$t = (c_1c_2c_3, d_1d_2d_3, e_1e_2)$$

where the first triple represents the colors of the vertices $u, v,$ *and* w *respectively, the second triple represents the colors of their self-edges, and the last pair represents the colors of the edges* uv *and* vw *respectively. In the following discussion the* $(2,3)$*-game type of a single vertex* u *will be taken to be the collection of all possible such tuples. So we can say things like* $t \in \zeta_{2,3}(u)$*. Sometimes we will need to ignore the vertex and self-edge colors when they do not play any role in the discussion. In such cases we consider* $t = (e_1e_2) \in \zeta_{2,3}(u)$*.*

Lemma 7.2 *Assume* m *vertex colors and* n *edge colors. Then* $\Lambda(2,3;u) \leq (mn)2^{n\Lambda(2,2;u)}$ *and* $\Lambda(2,3;G) \leq 2^{\Lambda(2,3;u)}$*.*

Proof. Given the recursive nature of the definition of game types, the $(2,3)$-game type of a single vertex u is determined by (i) its isomorphism type which is represented by the first multiplicand (mn) and (ii) all possible combinations of the pairs: $\langle$the isomorphism type of (u,v), the $(2,2)$-game type of $v\,\rangle$ for every vertex $v \in G$. There are $n\Lambda(2,2;u)$ such pairs (excluding the isomorphism type of u for it is already counted in (i) and the isomorphism type of v for it is already counted in $\Lambda(2,2;u)$), hence all possible subsets of such pairs is given by the multiplicand $2^{n\Lambda(2,2;u)}$. The upper bound on $\Lambda(2,3;G)$ is clear.

Lemma 7.3 *Let* $G \in \mathcal{G}_{m,n}$*. Assume* $|G| > \Lambda(2,3;u)$*. Then there exists* $G' \in \mathcal{G}_{m,n}$ *such that* $|G'| = |G| + 1$ *and* $\zeta_{2,3}(G') = \zeta_{2,3}(G)$*.*

Proof. By Lemma 7.2 there must be two vertices $u_1, u_2 \in G$ such that $\zeta_{2,3}(u_1) = \zeta_{2,3}(u_2) = \gamma$. Add a new vertex v to G. Color v and its self-edge exactly as u_1's. Connect v to every other vertex in G. For every $w \in G$ such that $w \neq u_1$, use the color of the edge (u_1, w) to color the edge (v, w). Finally, use the color of (u_1, u_2) to color (v, u_1). Let G' be the newly constructed graph. As already mentioned in Remark 7.1, for every edge emanating from u_1 of color e it must be the case that $(ee) \in \zeta_{2,3}(u_1; G)$. This corresponds to putting the first pebble p_1 on u_1, the second p_2 on v, where (u_1, v) has color e, and then removing p_1 and reinserting it onto u_1. Another way through which (ee) can be in $\zeta_{2,3}(u_1; G)$ is that there is a path in G of distinct vertices u_1ww' of color ee. Actually the addition of v as done above will create these latter monochromatic paths starting from v for every color e of an edge emanating from v. Such monochromatic paths of distinct vertices that start from u_1 may not exist, however, u_1 can not be distinguished from v using them since there are only two pebbles, hence $\zeta_{2,3}(u_1; G) = \zeta_{2,3}(u_1; G') = \zeta_{2,3}(v; G')$. It is also obvious that for any other $w \in G$, it is maintained that $\zeta_{2,3}(w; G) = \zeta_{2,3}(w; G')$. Hence, $\zeta_{2,3}(G') = \zeta_{2,3}(G)$.

As a direct consequence of this lemma we have the following inexpressibility result.

Theorem 7.4 *$DIV_k \notin bin\Sigma_1^1(2,3)$*

8 $\mathbf{DIV_k \in bin\Sigma_1^1(3,3)}$

In this section we show that $DIV_k \in bin\Sigma_1^1(3,3)$ by looking at the $(3,3)$-game types of graphs in $\mathcal{G}$. From the proofs one could extract out the actual defining sentence. We will do this in the case of $k=2$.

Remark 8.1 *Assume $G \in \mathcal{G}$ and let $u \in G$. Then $\zeta_{3,3}(u;G)$ can be characterized by the set of all paths in G of length 2 starting from u. Given one such path uvw we will represent it by the tuple*

$$t = (c_1c_2c_3, d_1d_2d_3, e_1e_2, \neg eq(u,w))$$

where the first triple represents the colors of the vertices u, v, w respectively, the second triple represents the colors of their self-edges, e_1e_2 represents the colors of the edges uv, and vw respectively, and finally $\neg eq(u,w)$ represents the truth value of whether u and w are not identical, it is assigned either t for true or f for false. Notice that the existence of three pebbles enables the spoiler to overcome the problem raised in the proof of Lemma 7.3 and caused her to lose the EF game, namely the inability to distinguish between monochromatic paths of the form uvu and monochromatic paths of the form uvw where $u \neq w$. Actually, as we will see below, this distinction is the main reason for successful expressibility of DIV_k in $bin\Sigma_1^1(3,3)$.

Definition 8.2 (Symmetric game types) *Let γ be a $(3,3)$-game type of a vertex $u \in G \in \mathcal{G}_{m,n}$. Let $\mathscr{C}$ be the set of m vertex colors and let $\mathscr{D}$ be the set of n edge colors. Assume $k \leq n$.*

1. *γ is called* k-symmetric *if the following hold:*
   (a) *there exist $c \in \mathscr{C}$ and $d \in \mathscr{D}$ such that if $(c_1c_2c_3, d_1d_2d_3, e_1e_2, *) \in \gamma$, ($*$ means 'do not care') then $c_1 = c_2 = c_3 = c$ and $d_1 = d_2 = d_3 = d$ (so γ is monochromatic with respect to the vertex and self-edge colors)*
   (b) *there exists $D \subseteq \mathscr{D}$ such that $|D| = k$ and for all distinct $e, e' \in \mathscr{D}$, $(ccc, ddd, ee', t), (ccc, ddd, e'e, t) \in \gamma$*
   (c) *if $(ccc, ddd, ee', t), (ccc, ddd, e'e, t) \in \gamma$ and $e \neq e'$ then it must be the case that $e, e' \in D$*
2. *γ is called* fully symmetric *if γ is n-symmetric.*

3. *A graph $G \in \mathcal{G}_{m,n}$ is called k-symmetric if all vertices in G have the same $(3,3)$-game type γ where γ is k-symmetric.*

Notation 8.3 *Most often in the following discussion we will only consider game types $\zeta_{3,3}(u)$ that are monochromatic with respect to the vertex and self-edge colors and/or be concerned only with paths of length 2 of distinct vertices starting from u. For simplicity in such cases, $\zeta_{3,3}(u)$ will be viewed as the collection of pairs (dd') that represent the colors along the path of length 2 starting from u.*

Let G be a graph. Let $\Delta(G)$ denote the maximum degree of G and let $\chi'(G)$ denote its edge chromatic number. The following theorem gives bounds for χ'.

Theorem 8.4 (Vizing 1964, p.119 in [Diestel 2006])

$$\Delta(G) \leq \chi'(G) \leq \Delta(G) + 1$$

Vizing's theorem divides the finite graphs into two classes based on their edge chromatic number. Those with $\chi' = \Delta$ are called *class I*, and those with $\chi' = \Delta + 1$ are called *class II* [Diestel 2006]. The following lemma applies this classification to complete graphs.

Lemma 8.5 (Theorem 4.1 in [Fiorini et al. 1977]) *Consider the complete graph K_n. If n is even, then it is class I, otherwise it is class II.*

Lemma 8.6 *Let $G \in \mathcal{G}_{m,n}$ be fully symmetric of minimum size k. Then $n+1 \leq k \leq n+2$.*

Proof. Since there are n distinct colors, then $k \geq n+1$. If n is odd, then let $k = n+1$. Since k is even, then by Lemma 8.5 we have $\chi'(K_{n+1}) = n$. If n is even, let $k = n+2$. Again by Lemma 8.5, $\chi'(K_{n+2}) = n+1$. Add a new color c' to the list of given n colors and use the new list to get a proper edge coloring of K_{n+2}. Choose a color c arbitrarily from the original list, and for every edge of color c' change its color to c.

Remark 8.7 *Let $G \in \mathcal{G}_{m,n}$ be fully symmetric. Let $\gamma = \zeta_{3,3}(u)$ for any $u \in G$. Let $d \in \mathscr{D}$ and assume that $(dd) \notin \gamma$. Then it must be the case that $|G| \pmod 2 = 0$. Otherwise either there exists some $u \in G$ with two edges incident on it of color d, hence $(dd) \in \gamma$ which contradicts the assumption or u has no edge incident on it of color d which contradicts the definition of G being fully symmetric.*

Lemma 8.8 *Let k be an even positive integer. Then there exist a pair of positive integers (m,n) and a $(3,3)$-game type Γ for graphs such that for any $G \in \mathcal{G}_{m,n}$ the following holds: $\zeta_{3,3}(G) = \Gamma$ implies that $|G| = bk$ for some integer $b \geq 1$.*

Proof. We will build Γ to be monochromatic with respect to the vertex color and the self-edge color, hence $m = 1$. Let d be the self-edge color. Assume $k = 2j$. Construct symmetric vertex game types $\gamma_0, \ldots, \gamma_{j-1}$ such that $\gamma_i = \{(d_{\gamma_i}, d), (d, d_{\gamma_i}), (d, d)\}$ where $d_{\gamma_i} \neq d$ and is unique for every $i < j$ (each pair in γ_i represents the colors of some path of length 2 starting from the vertex). Let $D = \{d_{i,i+1 \pmod j} : i < j\}$ be a collection of colors such that: (i) if $j = 1$, then $d_{0,0} = d$ and D will just represent the color of self-edges, (ii) if $j = 2$, then $d_{0,1} = d_{1,0}$, and (iii) if $j \geq 2$, then $d \notin D$ and $d_{\gamma_i} \notin D$ for every i.

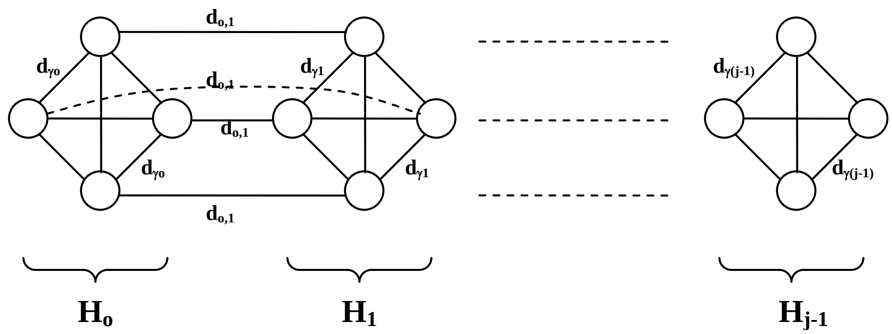

Figure 1: An arbitrary graph with game type Γ must have size multiple of k for k even (all unlabeled edges are colored d)

For every $i < j$, let H_i be a 2-symmetric graph such that (i) for every $u \in H_i$, $\zeta_{3,3}(u) = \gamma_i$, hence $|H_i|$ must be even since $(d_{\gamma_i}, d_{\gamma_i}) \notin \gamma_i$ (see Remark 8.7) and (ii) for every i, i', $|H_i| = |H_{i'}|$. Connect all the graphs H_i's and let H denote the resulting graph. For every i and for every $u \in H_i$ choose a unique $v_u \in H_{(i+1) \pmod j}$ and use $d_{i,(i+1) \pmod j}$ to color the edge (u, v_u) (in case $j = 1$, then $H = H_0$ and $v_u = u$ and this is just coloring the self-edge of u). For all the remaining uncolored edges use d to color them. Hence for any $i < j$, we have $(d_{i,(i+1) \pmod j}, d_{i,(i+1) \pmod j}) \notin \zeta_{3,3}(u; H)$ for any $u \in H$. See Figure 1.

We can easily notice that: (i) for every i, all vertices of the subgraph H_i have the same $(3,3)$-game type inside H, let δ_i denote this type, (ii) δ_i is an extension of γ_i, (iii) for all distinct i, i', we have $\delta_i \neq \delta_{i'}$ $(\delta_i \Delta \delta_{i'} \supseteq \{(d_{\gamma_i}, d), (d_{\gamma_{i'}}, d)\})$, and (iv) each δ_i is 2-symmetric with respect to the two colors d_{γ_i} and d. Let $n = |\{d_{\gamma_i} : i < j\}| + |\{d_{i,(i+1) \pmod j} : i < j\}| + 1 = 2j + 1$ (the last 1 is for the color d). Let $\Gamma = \{\delta_i : i < j\}$.

Let $G \in \mathcal{G}_{m,n}$ such that $\zeta_{3,3}(G) = \Gamma$. Each $\gamma_i \subseteq \delta_i$, which represents the 2-symmetric part of δ_i, must be realized inside G by a subgraph H_i such that $|H_i| \pmod 2 = 0$. Notice that for every $i, i' < j$, $(d_{i,(i+1) \pmod j}, d_{i,(i+1) \pmod j}) \notin$

$\delta_{i'}$, hence all $H_i's$ must have the same size (the edges $d_{i,(i+1) \pmod j}$ may be thought of as creating one-to-one maps between the H_i's so they are forced to have the same size). Therefore, $|G| = 2bj = bk$ for some positive integer b. So $(1, 2j + 1) = (1, k + 1)$ and Γ satisfy the conclusion of the lemma.

From this lemma we can immediately derive the following expressibility result.

Theorem 8.9 *Let k be an even positive integer. Then $DIV_k \in bin\Sigma_1^1(3, 3)$. More specifically, DIV_k can be expressed by a sentence of the following form*

$$\exists R_1 \ldots \exists R_l \varphi$$

where φ is a first-order sentence with 3 first-order variables and quantifier depth 3. Each R_i is a binary second-order variable and $l \leq \lceil \log_4 (k + 1) \rceil$.

Proof. Let Γ be the game type obtained in Lemma 8.8. We will show that $\mathcal{S}$ has a winning strategy in the $bin\Sigma_1^1(3, 3)$ game over the class of structures of cardinalities divisible by k. Assume $\mathcal{D}$ starts the game by choosing a structure $\mathcal{A}$ such that $|A| \pmod k = 0$. Let $\mathcal{S}$ colors $\mathcal{A}$ to get a graph $G \in \mathcal{G}$ such that $\zeta_{3,3}(G) = \Gamma$. $\mathcal{D}$ has then two possible responses: (i) choosing a structure $\mathcal{B}$ and coloring it to obtain $G' \in \mathcal{G}$ such that $\zeta_{3,3}(G') = \Gamma$, but then by Lemma 8.8 it must be the case that $|G'| \pmod k = 0$ and hence $\mathcal{D}$ loses the game at its second-order phase or (ii) choosing a structure $\mathcal{B}$ such that $|B| \pmod k \neq 0$ and color it to obtain $G' \in \mathcal{G}$ with $\zeta_{3,3}(G') = \Gamma'$ but again by Lemma 8.8 it must be the case that $\Gamma \neq \Gamma'$ hence by Proposition 5.5, $\mathcal{D}$ loses the game at its first-order phase. So in any case $\mathcal{S}$ wins the game, hence $DIV_k \in bin\Sigma_1^1(3, 3)$. The upper bound for l is obtained from the value of n derived in the proof of Lemma 8.8 and by realizing that each binary second-order variable contributes exactly 4 new colors.

In the introduction we gave a sentence that defines DIV_2. In the following example we will use the proof of Lemma 8.8 to show how this sentence can be derived systematically.

Example 8.10 *Consider the $EVEN$ problem. We need two edge colors d_1 and d_2 and one vertex color c. In the second-order phase of the EF game, $\mathcal{D}$ will first choose G_1 which is just a set of unconnected vertices with $|G_1| \pmod 2 = 0$. $\mathcal{S}$ will then convert G into a complete graph with all self-edges, let G_1' denote the new graph. $\mathcal{S}$ colors G_1' as follows: (i) use c to color all the vertices, (ii) use d_1 to color all the self-edges, (iii) for every distinct pair of vertices $u_i, v_i \in G_1'$, use d_2 to color the edge (u_i, v_i), and (iv) use d_1 to color all the remaining edges. This coloring implies that every $u \in G_1'$ has exactly one edge of color d_2 incident on it, hence d_2 corresponds to d_{γ_i} in the proof of Lemma 8.8. G_1' can be viewed* *as a* 1*-regular graph (a graph with isolated edges) by looking exclusively at the edges of color* d_2*. It can be easily checked that all the vertices in* G'_1 *have the same game type* $\gamma \supseteq \{(d_1d_2, t), (d_2d_1, t), (d_1d_1, t)\}$ *(ignoring the vertex and self-edge colors and considering only paths of length* 2 *with distinct vertices). Next* $\mathcal{D}$ *chooses a set of unconnected vertices* G_2 *with* $|G_2| \pmod 2 = 1$*.* $\mathcal{D}$ *converts* G_2 *into* G'_2*, a complete graph with all self-edges, and then tries to color it so as to have the same* $(3, 3)$*-game type as* G'_1*. Since* $(d_2d_2) \notin \gamma$*, then by Remark 8.7, this is impossible, in other words* G'_2 *can not be converted into a* 1*-regular graph. There must exist some vertex* $u \in G'_2$ *such that either* $(d_2d_2) \in \zeta_{3,3}(u; G'_2)$ *or* $(d_2d_1) \notin \zeta_{3,3}(u; G'_2)$*. Hence,* $\mathcal{S}$ *can win the first-order phase of the game by playing the differentiating path using her* 3 *pebbles. In the following we construct a sentence* $\sigma \in bin\Sigma^1_1(3, 3)$ *that defines EVEN*

$$\varphi_1(R) \overset{def}{=} \forall x \neg R(x, x), \qquad \text{coloring the self-edges of } G'_1 \text{ with } c$$

$$\varphi_2(R) \overset{def}{=} \forall x \forall y \left(R(x, y) \longleftrightarrow R(y, x)\right), \qquad G'_1 \text{ is undirected}$$

$$\varphi_3(R) \overset{def}{=} \forall x \exists y \left(R(x, y) \wedge \forall z \left(R(x, z) \longrightarrow z = y\right)\right), \text{the } d_2 \text{ coloring of edges in } G'_1$$

$$\sigma \overset{def}{=} \exists R \left(\varphi_1(R) \wedge \varphi_2(R) \wedge \varphi_3(R)\right) \tag{8.10.1}$$

Remark 8.11 *From Example 8.10, it is clear that* 1*-regular graphs can be used to characterize divisibility by* 2*. This observation has been used in the proof of Lemma 8.8 to construct the subgraphs* $H'_i s$ *such that each one of them must be of even size. Additionally, the coloring scheme of the interconnections among the* $H'_i s$ *ensures that they have identical sizes. Hence, the total size of the resulting graph must be a multiple of* $2j$ *which is the desired goal.*

Next we turn to expressibility of divisibility by odd numbers.

Lemma 8.12 *Let* $k \neq 1$ *be an odd positive integer. Then there exist a pair of positive integers* (m, n) *and a* $(3, 3)$*-game type* Γ *for graphs such that for any* $G \in \mathcal{G}_{m,n}$ *the following holds:* $\zeta_{3,3}(G) = \Gamma$ *implies that* $|G| = bk$ *for some integer* $b \geq 1$*.*

Proof. We will build Γ to be monochromatic with respect to the vertex color and the self-edge color, hence $m = 1$. Assume $k = 2j + 1$ for $j \geq 1$. Let Γ' be the game type Γ constructed in the proof of Lemma 8.8. Let $H \in \mathcal{G}_{m,n}$ be the graph constructed in the proof of Lemma 8.8 such that H is constructed exactly from the subgraphs $H_0, \ldots, H_{j-1}$ with $|H_i| = 2b$ for some positive integer b.

Let $u_0, \ldots, u_{b-1}$ be new vertices, connect them together and to every vertex in H. Use d to color all the edges between the u_i's. For each $i < j$ choose an arbitrary set of vertices V_i such that (i) $V_i \subseteq H_i$, (ii) $|V_i| = b$, and (iii) for every $w, w' \in V_i$, the edge (w, w') is colored d. For every $i < j$ and for every $i' < b$, choose a unique $w_{u_{i'}} \in V_i$, and use $d_{\gamma_{(i+1) \pmod j}}$ to color the edge $(u_{i'}, w_{u_{i'}})$. Use d to color the remaining uncolored edges from the u_i's to H. Call the new graph H' and notice that $H' \in \mathcal{G}_{m,n}$ where $n = 2j+1$ is the number of colors used to color the edges of H'. See Figure 2.

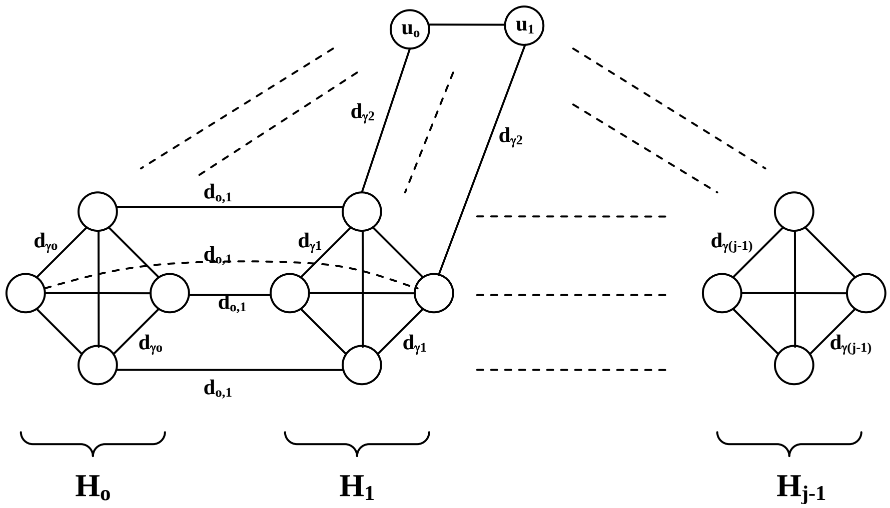

Figure 2: An arbitrary graph with game type Γ must have size multiple of k for k odd (all unlabeled edges are colored d)

Notice the following: (i) for each color d_{γ_i}, u_i has an edge of that color incident on it, (ii) all the u_i's have the same $(3,3)$-game type inside H', let ρ denote that game type, (iii) $(d_{\gamma_i}, d_{\gamma_i}) \notin \rho$ for every $i < j$, however, $(d_{\gamma_i}, d_{\gamma_{(i+1) \pmod j}}) \in \rho$, and (iv) $(d,d) \in \rho$. Now look at the new emerging game types inside H'. For every $i < j$, δ_i no longer exists, but is broken into two new game types: (i) δ_i^0 which is the game type of every vertex in V_i and (ii) δ_i^1 which is the game type of every vertex in $H_i \backslash V_i$. Each vertex u_i has the new game type ρ. An important observation is that for every $u \in H'$, $(d_{\gamma_i}, d_{\gamma_i}) \notin \zeta_{3,3}(u)$ for every $i < j$. Let

$$\Gamma = \{\delta_i^0 : i < j\} \cup \{\delta_i^1 : i < j\} \cup \{\rho\}$$

Let $G \in \mathcal{G}_{m,n}$ be such that $\zeta_{3,3}(G) = \Gamma$. Notice that for every vertex $v \in G$ with $\zeta_{3,3}(v) = \delta_i^0$, there must exist exactly one vertex w_v such that $\zeta_{3,3}(w_v) = \delta_i^1$ and the edge (v, w_v) is colored d_{γ_i}. The converse also holds for vertices of game type δ_i^1. Hence there is a one-to-one correspondence between $\{u \in G : \zeta_{3,3}(u) = \delta_i^0\}$ and $\{u \in G : \zeta_{3,3}(u) = \delta_i^1\}$, therefore $\left|\{u \in G : \zeta_{3,3}(u) = \delta_i^0 \textit{ or } \zeta_{3,3}(u) = \delta_i^1\}\right| =$

$2b'$ for some positive integer $b' \geq 1$. Let W_i denote this last set of vertices. Similarly, we can show that (see also the proof of Lemma 8.8) there is a one-to-one correspondence between W_i and $W_{i'}$ for all $i, i' < j$. Hence, $|\bigcup\{W_i\colon i < j\}| = 2b'j$. Then any $u \in G \backslash \bigcup\{W_i\colon i < j\}$ must be of game type ρ.

Let $T_i = \{u \in G\colon \zeta_{3,3}(u) = \delta_i^0\}$. Note that all the T_i's must have the same size. Let $P = \{u \in G\colon \zeta_{3,3}(u) = \rho\}$. From the construction of H' it must be the case that every $u \in T_i$ uniquely determines a distinct $v_u \in P$ such that (u, v_u) is colored $d_{\gamma_{(i+1) \pmod j}}$ (since $(d_{\gamma_{(i+1) \pmod j}}, d_{\gamma_{(i+1) \pmod j}}) \notin \delta_i^0$). Hence $|T_i| \leq |P|$. Similarly, every $v \in P$ uniquely determines a vertex $w_v \in T_i$ such that (v, w_v) is colored $d_{\gamma_{(i+1) \pmod j}}$ (since $(d_{\gamma_{(i+1) \pmod j}}, d_{\gamma_{(i+1) \pmod j}}) \notin \rho$). Hence $|P| \leq |T_i|$. Therefore, $|P| = |T_i| = b'$. Now we count the number of vertices in G. $|G| = |\bigcup\{W_i\colon i < j\}| + |P| = 2b'j + b' = b'(2j+1) = b'k$. Hence, $(1, 2j+1) = (1, k)$ and Γ are as desired.

From Lemmas 8.8 and 8.12 we can derive the following general result.

Lemma 8.13 *Fix a positive integer $k \neq 1$. Let $m = 1$ and $n = k+1$. Then there exits a $(3,3)$-game type Γ for graphs in $\mathcal{G}_{m,n}$ such that for any $G \in \mathcal{G}_{m,n}$ the following holds: $\zeta_{3,3}(G) = \Gamma$ implies that $|G| = bk$ for some integer $b \geq 1$.*

This directly implies the following expressibility result.

Theorem 8.14 *Let $k \neq 1$ be a positive integer. Then $DIV_k \in bin\Sigma_1^1(3,3)$. More specifically, DIV_k can be expressed by a sentence of the following form*

$$\exists R_1 \ldots \exists R_l \varphi$$

where φ is a first-order sentence with 3 first-order variables and quantifier depth 3. Each R_i is a binary second-order variable and $l \leq \lceil \log_4 (k+1) \rceil$.

Proof. Similar to the proof of Theorem 8.9.

Corollary 8.15 $mon\Sigma_1^1 \subset bin\Sigma_1^1$

Proof. This follows directly from the inexpressibility result in Theorem 4.1 and the expressibility result in Theorem 8.14.

Lemma 8.16 *Let l_1, l_2 be two non-negative integers. Define $\Theta \subseteq bin\Sigma_1^1(3,3)$ that consists exactly of sentences that have at most l_1, l_2 unary and binary second-order variables respectively. Then $DIV_k \in \Theta$ for only finitely many k.*

Proof. Let m, n be the corresponding vertex and edge colors respectively. There are at most finitely many $(3,3)$-game types for graphs in $\mathcal{G}_{m,n}$. Assume the conclusion does not hold, then there are two distinct positive integers k_1, k_2 that can be distinguished by the same $(3,3)$-game type. But this implies that $\mathcal{D}$ can win the $bin\Sigma_1^1(3,3)$ game by choosing a structure of cardinality k such that exactly one of k_1 and k_2 is a factor of k. This is a contradiction.

Theorem 8.14 and Lemma 8.16 imply that DIV_k creates a proper hierarchy into $bin\Sigma_1^1(3,3)$.

9 Bounding the Binary Relation Variables

The following theorem gives an inexpressibility result for DIV_k in $bin\Sigma_1^1$ when the sizes of the interpretations of the binary relation variables are bounded.

Theorem 9.1 *Let $\sigma \in bin\Sigma_1^1$ be of the following form*

$$\exists R_1^{\leq f(l)} \ldots \exists R_t^{\leq f(l)} \exists S_1 \ldots \exists S_s \varphi$$

where $f(l) < \frac{l}{2t} - \frac{r2^s}{2t}$, where l is the size of any structure that models this sentence and r is the quantifier depth of φ. Then DIV_k can not be expressed by σ.

Proof. We show $\mathcal{D}$ has a winning strategy in the second-order EF game with r rounds in the first-order phase (assume the number of pebbles $p = r$). $\mathcal{D}$ starts by choosing a complete uncolored graph G with all self-edges such that

$$|G| \pmod{k} = 0 \tag{9.1.1}$$

$$|G| > r2^s + 2tf(|G|) \tag{9.1.2}$$

There are a total of $m = 2^s$ vertex colors. For the edges it is easier to directly handle each R_i separately than to consider the colors resulting from their combinations. $\mathcal{S}$ does the following with G: (i) color the vertices using the given m colors and (ii) construct the edge sets $E_1, \ldots, E_t$ among the vertices of G such that $|E_i| \leq f(|G|)$ for each i. From 9.1.2, there must be at least $r2^s$ vertices with degree 0, that is there is no edge from any of the E_i's that is incident on any of these vertices. Then by the pigeonhole principle there must be at least r of those vertices that are monochromatic, let their color be c. Let Γ be the collection of vertices in G that are colored c and with degree 0, then $|\Gamma| \geq r$. In order for the inequality in 9.1.2 to make sense it must be the case that $f(|G|) < \frac{|G|}{2t} - \frac{r2^s}{2t}$ as given in the theorem hypothesis. $\mathcal{D}$ then chooses a graph $G' = (G \cup \{w\})$ with a new vertex w and does the following: (i) color the vertices of $G \subseteq G'$ exactly as $\mathcal{S}$ did, (ii) color w with c, (iii) construct the edge sets $E_1, \ldots, E_t$ among the vertices of $G \subseteq G'$ exactly as $\mathcal{S}$ did, and (iv) leave the vertex w unconnected to any other vertex. In the first-order phase of the game $\mathcal{D}$ can win by following a similar strategy to that described in the proof of Theorem 4.1.

10 Conclusion and Future Work

In this paper we have provided a partial framework for the study of expressibility in Σ_1^1. This framework uses interesting combinatorics based on second-order EF-games and the notion of game types. We have studied the expressibility of DIV_k

in different sublogics of Σ_1^1 getting inexpressibility results until expressibility is obtained inside $bin\Sigma_1^1(3,3)$. Based on k, DIV_k creates a proper hierarchy inside this sublogic. In the future I plan to pursue research in the following points:

1. Finding tight lower/upper bounds for the DIV_k hierarchy in $bin\Sigma_1^1(3,3)$. This is mainly a combinatorial problem and helps understanding game types specially for future plans when using second-order variables with higher arities.

2. Study the expressibility of $\overline{DIV_k}$ in $bin\Sigma_1^1(3,3)$.

3. Study natural extensions of $bin\Sigma_1^1(3,3)$ inside Σ_1^1 within the framework developed above. The parameters (logical resources) used in the abovementioned research, and hence in future extensions, are the following: (i) the arity of the second-order variables, (ii) the second-order quantifier depth, (iii) the number of first-order variables, and (iv) the first-order quantifier depth. Other parameters may also be studied such as the number of alternations of first-order quantifiers and also parameters that arise from the interleaving of first- and second-order quantifiers such as depth and alternation, however, this may require a dramatic change in the rules of the EF games. I plan to use number-theoretic properties for the study of expressibility such as primeness, number and sizes of equivalence classes of a definable equivalence relation, whether two definable subsets of a structure form an amicable number, etc. The main goals of this study are: (i) create proper hierarchies into sublogics of Σ_1^1 and into Σ_1^1 itself, hence giving more insight into NP and (ii) the study of expressibility of some interesting number-theoretic properties for its own sake.

4. Extending the above to Π_1^1 and the whole of second-order logic, hence essentially looking into the whole polynomial hierarchy.

References

[Ajtai et al. 1990] Ajtai M., Fagin R.: Reachability is Harder for Directed than for Undirected Finite Graphs. The Journal of Symbolic Logic, 55(1), 1990.

[Diestel 2006] Diestel R.: Graph Theory. Springer, 2006.

[Ehrenfeucht 1961] Ehrenfeucht A.: An Application of Games to the Completeness Problem for Formalized Theories. Fundamenta Mathematicae, 49:129-141, 1961.

[Fagin 1974] Fagin .: Generalized First-Order Spectra and Polynomial-Time Recognizable Sets. In R. Karp, editor, Complexity of Computation, volume 7, pages 43-73. SIAM-AMS, 1974.

[Fagin 1975] Fagin R.: Monadic Generalized Spectra. Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 21:89-96, 1975.

[Fiorini et al. 1977] Fiorini S., Wilson R. J.: Edge-Colourings of Graphs. Pitman, 1977.

[Fraïssé 1954] Fraïssé R.: Sur quelques classifications des systèmes de relations. Université d'Alger, Publications Scientifiques, Série A, 1:35-182, 1954.

[Immerman 1998] Immerman N.: Descriptive Complexity. Springer, 1998.

[Koucký et al. 2006] Koucký M., Lautemann C., Poloczek S., Thérien D.: Circuit Lower Bounds via Ehrenfeucht-Fraïssé Games. IEEE Conference on Computational Complexity, p.190-201, 2006.

[Kreidler et al. 1997] Kreidler M., Seese D.: Monadic NP and Built-in Trees. Lecture Notes in Computer Science, 1258:260-274, 1997.

[Kreidler et al. 1998] Kreidler M., Seese D.: Monadic NP and Graph Minors. Lecture Notes in Computer Science, 1584:126-141, 1998.

[Libkin 2004] Libkin L.: Elements of Finite Model Theory. Springer, 2004.

[Schwentick 1995] Schwentick T.: Graph Connectivity, Monadic NP and Built-in Relations of Moderate Degree. In Proc. of the 22nd International Colloq. on Automata, Languages, and Programming, pages 405-416, 1995.

[Schwentick 1996] Schwentick T.: On Winning Ehrenfeucht Games and Monadic NP. Annals of Pure and Applied Logic, 79:61-92, 1996.

[Schwentick 1997] Schwentick T.: Padding and the Expressive Power of Existential Second-Order Logics. In 11th Annual Conference of the EACSL, CSL'97, pages 461-477, 1997.

[Stockmeyer 1977] Stockmeyer L.: The Polynomial-Time Hierarchy. Theoretical Computer Science, 3:1-22, 1977.