Moscow Journal of Combinatorics and Number Theory

2020
vol. 9 no. 3

First-order definitions of subgraph isomorphism through the adjacency and order relations

Oleg Grigoryan, Mikhail Makarov and Maksim Zhukovskii

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 3, 2020
https://doi.org/10.2140/moscow.2020.9.293

First-order definitions of subgraph isomorphism through the adjacency and order relations

Oleg Grigoryan, Mikhail Makarov and Maksim Zhukovskii

We study first-order definitions of graph properties over the vocabulary consisting of the adjacency and order relations. We compare logical complexities of subgraph isomorphism in terms of the minimum quantifier depth in two settings: with and without the order relation. We prove that, for pattern-trees, it is at least (roughly) two times smaller in the former case. We find the minimum quantifier depths of $<$-sentences defining subgraph isomorphism for all pattern graphs with at most 4 vertices.

1. Introduction

We consider graphs with the vertex sets $[n] := \{1, \dots, n\}$, $n \in \mathbb{N}$. A *graph property* is a set of graphs closed under isomorphism (isomorphism preserves the adjacency relation $\sim$). In this paper, we focus on *subgraph isomorphism properties* defined below.

Given a *pattern-graph* F, let $\mathcal{S}(F)$ denote the property of containing a (not necessarily induced) subgraph isomorphic to F (i.e., $\mathcal{S}(F)$ is the set of all graphs containing a subgraph isomorphic to F). Let $\Sigma = \{\sim, =, R_1, \dots, R_s\}$ be a finite set of relational symbols, where R_i, $i \in [s]$, represents a certain predicate on $\mathbb{N}$ of arity a_i. Below, $<$ denotes the usual linear order on $\mathbb{N}$.

We consider the first-order logic over the vocabulary Σ. For a sentence φ in this logic and a graph property $\mathcal{P}$, we say that φ expresses $\mathcal{P}$ if

$$G \in \mathcal{P} \Leftrightarrow G \models \varphi.$$

Clearly, for an arbitrary graph F on $[\ell]$ with the edge set $E(F)$, $\mathcal{S}(F)$ is expressed by the first-order sentence

$$\exists x_1 \dots \exists x_\ell \quad \Big[\bigwedge_{i \neq j} x_i \neq x_j\Big] \wedge \Big[\bigwedge_{\{i,j\} \in E(F)} x_i \sim x_j\Big]. \tag{1}$$

Thus, $\mathcal{S}(F)$ is definable in the most laconic first-order logic (over the vocabulary $\{\sim, =\}$).

Consider the parameters $D_\Sigma(F)$ and $W_\Sigma(F)$ defined, respectively, as the minimum quantifier depth (maximum number of variables in a sequence of nested quantifiers, see formal definition in [Libkin 2004], Definition 3.8) and the minimum variable width (the number of distinct variables, see the definition just before Proposition 6.6 in [Libkin 2004]) of a first-order sentence in the vocabulary Σ expressing $\mathcal{S}(F)$. For $\Sigma = \{\sim, =\}$, we will omit the indices and write, simply, $D(F)$ and $W(F)$. For $\Sigma = \{\sim, <\}$, we will write $D_<(F)$ and $W_<(F)$. Finally, $D_{\text{Arb}}(F)$ and $W_{\text{Arb}}(F)$ denote the minimum of $D_\Sigma(F)$ and W_Σ(F)

This work is supported by the grant N 18-71-00069 of Russian Science Foundation.
MSC2010: 03C13, 68Q19.
Keywords: first-order logic, subgraph isomorphism, order, logical complexity, quantifier depth.

over all finite Σ respectively. Notice that the relation $W_\Sigma(F) \le D_\Sigma(F)$ follows from the obvious fact that any first-order sentence of the quantifier depth d can be rewritten using at most d variables.

Since the sentence (1) has the quantifier depth ℓ, we have $D(F) \le \ell$. On the other hand, note that K_ℓ, the complete graph on ℓ vertices, contains F as a subgraph, while $K_{\ell-1}$ does not. Since ℓ first-order variables are necessary in order to distinguish between K_ℓ and $K_{\ell-1}$ (indeed, any quantifier-free subformula of a sentence with $\ell-1$ variables has at most $\ell-1$ variables, and so, it has the same truth value on both $K_{\ell-1}$ and K_ℓ), we have $W(F) = D(F) = \ell$. (However, the problem of estimating the values of these parameters becomes nontrivial for sufficiently large connected *input graphs*; see [Verbitsky and Zhukovskii 2019b; Verbitsky and Zhukovskii 2019c]).

A straightforward conversion of a first-order sentence in the vocabulary $\{\sim, =\}$ defining $\mathcal{S}(F)$ into an algorithm for deciding is there a copy of F in an input graph on n vertices leads to the trivial time bound $O(n^{D(F)})$ which can actually be improved to $O(n^{W(F)})$; see [Libkin 2004, Proposition 6.6]. The preceding discussion shows that first-order sentences in the vocabulary $\{\sim, =\}$ defining $\mathcal{S}(F)$ can only be considered as a very weak computational model for the corresponding decision problem. Notice that the general time bound for the mentioned decision problem established by Nešetřil and Poljak [1985] is $O(n^{(\omega/3)\ell+2})$, where ω is the exponent of fast matrix multiplication, whose value lies between 2 and 2.373 [Le Gall 2014].

However, the above time bounds in terms of the quantifier depth and the variable width apply to the vocabulary $\{\sim, <\}$ as well. It is well known [Libkin 2004; Schweikardt 2013] that there are properties of finite structures that can be defined in first-order logics with $<$ but not without. Thus, it is not surprising, that for certain graphs F on ℓ vertices, $\mathcal{S}(F)$ can be defined much more succinctly in the first-order logic over $\{\sim, <\}$. As a simple example, for $F = K_{1,\ell-1}$ (star graph), it is easy to see that $D_<(F) \le \lceil \log_2(\ell-2) \rceil + 3$ and $W_<(F) \le 3$. Indeed, let

$$\varphi_1(x, y, z) = (x < y) \wedge (x \sim z) \wedge (y \sim z),$$
$$\varphi_k(x, y, z) = \exists w \;\; \varphi_{\lfloor k/2 \rfloor}(x, w, z) \wedge \varphi_{\lceil k/2 \rceil}(w, y, z), \quad k \ge 2.$$

Then $\mathcal{S}(F)$ is expressed by the sentence $\exists z \exists x \exists y \quad \varphi_{\ell-2}(x, y, z)$ of the quantifier depth $\lceil \log_2(\ell-2) \rceil + 3$. The same property is expressed by the sentence

$$\exists z \quad \Big(\exists x \;\, [x \sim z] \wedge \Big[\exists y \;\, (x < y) \wedge (y \sim z) \wedge \Big(\exists x \;\, [y < x] \wedge [x \sim z] \wedge [\dots]\Big)\Big]\Big)$$

with 3 variables.

The natural question to ask is what about an arbitrary tree F? Are $D_<(F)$ and $W_<(F)$ significantly less than ℓ? In this paper, we restrict ourselves with estimating the parameter $D_<(F)$ only. In Section 2, we prove that, for an arbitrary tree F, $D_<(F) \le \frac{1}{2}\ell + \lceil \log_2(\ell+2) \rceil - 1$, which is, roughly, a half of $D(F)$. Unfortunately, we can not prove that it is close to optimal at least for paths. Notice that K_ℓ on $[\ell]$ and $K_{\ell-1}$ on $[\ell-1]$ can not be distinguished by a sentence of the quantifier depth smaller than $\log_2(\ell+1) - 2$ (it can be shown, for example, in a usual way via the Ehrenfeucht–Fraïssé game). This implies the trivial lower bound $D_<(F) \ge \log_2(\ell-2)$. On the other hand, $D_{\mathrm{Arb}}(F) \le \mathrm{td}(F) + 2$ [Rossman 2016], where $\mathrm{td}(F)$ is the treedepth of F. In particular, $\mathrm{td}(P_\ell) = \lceil \log_2(\ell+1) \rceil$.

The color coding algorithm of Alon, Yuster and Zwick [Alon et al. 1995] guarantees the time bound $O(n^2)$ for the decision problem for $\mathcal{S}(F)$. Therefore, we can not compete with the best known algorithm.

However, we motivate our interest to the question above exclusively by comparison of the expressive powers of first-order logics with and without the $<$ relation.

In Section 4, we find $D_<(F)$ for all F on at most 4 vertices. For F having 3 vertices, we prove that $D_<(F)=3$ if and only if F is connected. Otherwise, $D_<(F)=2$. For F having 4 vertices, the following is true: $D_<(F)=4$ if and only if F contains a C_4; otherwise, $D_<(F)=3$. In particular, $D_<(K_3)=3$, $D_<(K_4)=4$. It is known that $W_{\mathrm{Arb}}(K_\ell)\geq\frac{1}{4}\ell$ [Rossman 2008]. Moreover, the main result of [He 2015] states that $W_{\mathrm{Arb}}(K_\ell)=\ell$ that, obviously, implies $D_<(K_\ell)=\ell$. Unfortunately, there is no journal version of the latter result. By this reason, we give the proofs for K_3 and K_4 in our paper.

2. Pattern trees of arbitrary size

Theorem 1. *Let F be a tree on ℓ vertices. Then $D_<(F)\leq\frac{1}{2}\ell+\lceil\log_2(\ell+2)\rceil-1$.*

Proof. Consider an arbitrary linear order $<$ on the set of vertices of F: $v_1<\cdots<v_\ell$. Below, we construct a sentence with the desired quantifier depth that defines the existence of a subgraph isomorphic to F and an isomorphism between F and its copy that preserves both $\sim$ and $<$ relations. Having this, it remains to consider a conjunction of all such sentences over all possible linear orders on the vertex set of F.

Let $\mathcal{E}=\{v_2,v_4,\dots\}$ be the set of vertices with even positions and $\mathcal{O}=\{v_1,v_3,\dots\}$ be the set of vertices with odd positions.

Remove from F those edges that have both vertices in $\mathcal{E}$. After that, remove from the obtained forest all isolated vertices. Finally, for every vertex of $\mathcal{E}$ that has degree d bigger than 1 in this graph, cut this vertex into d vertices each with its own unique edge. Denote the trees in the final forest by $F_1,\dots,F_r$. Clearly, if $v\in\mathcal{E}$ belongs to F_j, then v is a leaf. For $j\in[r]$, let F_j have the vertex set $\{v_{k_1^j},\dots,v_{k_{m(j)}^j}\}$ and let $v_{s_1^j},\dots,v_{s_{t(j)}^j}$ be the vertices of F_j that are from $\mathcal{E}$. We will call $v_{s_1^j},\dots,v_{s_{t(j)}^j}$ *distinguished leaves* of F_j.

Let

$$\varphi=\exists x_2\exists x_4\dots\exists x_{2\lfloor\ell/2\rfloor}\left[\bigwedge_{j=1}^{\lfloor\ell/2\rfloor-1}(x_{2j}<x_{2j+2})\right]\wedge\left[\bigwedge_{\{v_{2i},v_{2j}\}\in E(F)}(x_{2i}\sim x_{2j})\right]\left[\bigwedge_{j=1}^{r}\varphi_j(x_{s_1^j},\dots,x_{s_{t(j)}^j})\right],$$

where $\varphi_j(x_{s_1^j},\dots,x_{s_{t(j)}^j})$ states that there exists a copy T_j of F_j on the vertex set $\{x_{k_1^j},\dots,x_{k_{m(j)}^j}\}$ such that

$$v_{i^1}\sim v_{i^2}\text{ in }F_j\Leftrightarrow x_{i^1}\sim x_{i^2}\text{ in }T_j\qquad\text{for }v_{i^1},v_{i^2}\in V(F_j),$$
$$v_{i^1}<v_{i^2}\Leftrightarrow x_{i^1}<x_{i^2}\qquad\text{for }(v_{i^1},v_{i^2})\in[V(F_j)\cup\mathcal{E}]^2\setminus\mathcal{E}^2.$$

As in F_j, the vertices $x_{s_1^j},\dots,x_{s_{t(j)}^j}$ are *distinguished leaves* of T_j. It remains to show that $\varphi_j(x_{s_1^j},\dots,x_{s_{t(j)}^j})$ may be efficiently written (with the quantifier depth not more than $\lceil\log_2(\ell+2)\rceil-1$).

For every i, let $R_i(x)=(x_{2i}<x)\wedge(x<x_{2i+2})$ be the predicate that defines the position of x_{2i+1} in the described copy of F.

Let $j\in[r]$. Clearly, the diameter of F_j is at most $\lceil\ell/2\rceil+1$. Let c_j be a *central vertex* of F_j, i.e., that vertex that minimizes the maximum distance to leaves. Let $c_j=v_{2\gamma-1}$ for a certain $\gamma\in\{0,1,\dots,\lceil\ell/2\rceil\}$. We call a subtree of F_j its *c_j-branch*, if this subtree contains c_j itself, one of its children (here, F_j is rooted in c_j) and all the descendants of this child (a *descendant vertex* u of v is a vertex such that v is in the path from u to the root). Let $F_j^1, \dots, F_j^\mu$ be all the c_j-branches of F_j (μ is the number of children of c_j). Clearly, c_j is a leaf of all these trees. Let us distinguish it. For $q \in [\mu]$, let $v_{k_i^{j,q}}$, $i \in [m(j,q)]$, be all the vertices of F_j^q, and $v_{s_i^{j,q}}$, $i \in [t(j,q)]$, be the distinguished vertices of F_j^q (one of them is $v_{2\gamma+1}$, and all the others are those vertices of F_j^q that are distinguished in F_j).

Then,

$$\varphi_j(x_{s_1^j}, \dots, x_{s_{t(j)}^j}) \;=\; \exists x_{2\gamma+1} \quad R_\gamma(x_{2\gamma+1}) \wedge \Big[\bigwedge_{i=1}^{\mu} \varphi_j^i(x_{s_1^{j,q}}, \dots, x_{s_{t(j,q)}^{j,q}})\Big];$$

here $\varphi_j^i(x_{s_1^{j,q}}, \dots, v_{x_{t(j,q)}^{j,q}})$ states that there exists a copy T_j^i of F_j^i defined on the vertex set $\{x_{k_1^{j,q}}, \dots, x_{k_{m(j,q)}^{j,q}}\}$ such that

$$\begin{aligned} v_{i^1} \sim v_{i^2} \text{ in } F_j^i \Leftrightarrow x_{i^1} \sim x_{i^2} \text{ in } T_j^i \qquad & \text{for } v_{i^1}, v_{i^2} \in V(F_j^i),\\ v_{i^1} < v_{i^2} \Leftrightarrow x_{i^1} < x_{i^2}, \qquad & \text{for } (v_{i^1}, v_{i^2}) \in [V(F_j^i) \cup \mathcal{E}]^2 \setminus [\mathcal{E} \cup \{v_{2\gamma+1}\}]^2. \end{aligned}$$

The vertices $x_{s_1^{j,q}}, \dots, x_{s_{t(j,q)}^{j,q}}$ become the *distinguished leaves* of T_j^i.

The maximum diameter of $F_j^1, \dots, F_j^\mu$ is $\lceil D_j/2 \rceil$, where $D_j \le \lceil \ell/2 \rceil + 1$ is the diameter of F_j. By induction, in at most $\lceil \log_2(\ell+2) \rceil - 1$ steps, we will obtain a forest of edges with all the vertices distinguished. Every such edge is expressed by a formula defining the position of the last distinguished vertices in respect to the vertices of $\mathcal{E}$ and saying that its ends are adjacent. □

Remark. Clearly, two variables are enough to express the "odd" part of F. Therefore, $W_<(F) \le \frac{1}{2}\ell + 2$.

3. The Ehrenfeucht–Fraïssé game

The main tool in our proofs is Ehrenfeucht's theorem. It gives necessary and sufficient conditions of the elementary equivalence of finite structures (in our case graphs) in terms of the so-called Ehrenfeucht–Fraïssé game. We are going to apply this theorem to the vocabulary $\{\sim, <\}$ and, thus, recall the rules of the Ehrenfeucht–Fraïssé game only in these special settings in order to avoid heavy notations.

The *k-round Ehrenfeucht–Fraïssé game* (or, simply, the *k-game*) is played on two graphs G and H whose vertex sets are finite subsets of $\mathbb{N}$. There are two players — *Spoiler* and *Duplicator*. In every round, Spoiler chooses a vertex in G or in H, then Duplicator has to choose a vertex in the other graph. After k rounds, $x_1, \dots, x_k$ are chosen in G and $y_1, \dots, y_k$ are chosen in H. Duplicator wins if and only if, for all distinct $i, j \in [k]$,

$$(x_i \sim x_j) \Leftrightarrow (y_i \sim y_j), \qquad (x_i < x_j) \Leftrightarrow (y_i < y_j).$$

The following fact is a corollary of Ehrenfeucht theorem (see details in, e.g., Section 2 of [Verbitsky and Zhukovskii 2019a]).

Lemma 2. *$D_<(F)$ equals the minimum k such that, for any pair of graphs H, G such that $F \subset G$ but $F \not\subset H$, Spoiler has a winning strategy in the k-round game on G and H.*

4. Small patterns

4.1. ***3-patterns.*** There are 4 non-isomorphic graphs on 3 vertices: I_3 (empty graph on 3 vertices), $I_1 \sqcup K_2$, P_3 and K_3.

It is easy to see that $D_<(I_3) = D_<(I_1 \sqcup K_2) = 2$ since containing I_3 is defined by the sentence

$$\exists x \quad (\exists y \quad [y < x]) \wedge (\exists y \quad [x < y])$$

and containing $I_1 \sqcup K_2$ is defined by the sentence

$$\big[\exists x \quad (\exists y \quad [y < x]) \wedge (\exists y \quad [x < y])\big] \wedge \big[\exists x \exists y \quad (x \sim y)\big].$$

Let us switch to $F = K_3$. Let G be the disjoint union of two isolated vertices and K_3. More precisely, the vertex set of G is $\{1, 2, 3, 4, 5\}$ and the set of edges contains only $\{2, 3\}, \{3, 4\}, \{2, 4\}$. Let H be obtained from G by deleting the edge $\{2, 4\}$. It is clear that Duplicator wins the 2-round game on G, H and, therefore, $D_<(K_3) = 3$.

To prove that $D_<(P_3) = 3$, let us consider the following G and H. Let G have the vertex set $\{1, 2, 3, 4, 5\}$ and contain only two edges $\{2, 3\}$ and $\{2, 4\}$. Let H have the vertex set $\{1, 2, 3, 4\}$ and contain the only edge $\{2, 3\}$. Since Duplicator wins the 2-round game on these G and H as well, $D < (P_3) = 3$.

Summing up, for a connected pattern graph F on 3 vertices, $D_<(F) = D(F)$ while, for its complement, $D_<(F) < D(F)$. However, the situation changes for graphs on 4 vertices — $D_<(F)$ becomes smaller even for certain connected patterns F.

4.2. ***4-patterns without C_4.*** Clearly, for every F on 4 vertices, $D_<(F) \geq 3$ since Duplicator wins the 2-round game on K_4 and K_3. In this section, we prove that $D_<(F) = 3$ for every F that does not contain C_4.

Let us first show that, if F does not contain P_4, then $D_<(F) = 3$.

Let us call two graphs $F, \tilde{F}$ on vertex sets $V(F), V(\tilde{F}) \subset \mathbb{N}$ $(\sim, <)$-*isomorphic*, if there exists a bifection $f : V(F) \to V(\tilde{F})$ such that $u \sim v$ in F if and only if $f(u) \sim f(v)$ in $\tilde{F}$, and $u < v$ in F if and only if $f(u) < f(v)$ in $\tilde{F}$. For a pattern graph F on $[\ell]$, denote $D'_<(F)$ the minimum quantifier depth of a first-order sentence defining the property of containing a subgraph $(\sim, <)$-isomorphic to F.

Claim 3. *Let F be a graph on the vertex set $\{1, 2, 3, 4\}$. If either $3 \nsim 1$, or $1 \nsim 4$, or $4 \nsim 2$, then $D'_<(F) \leq 3$.*

Proof. Let F do not contain the edge $\{1, 3\}$. Then the containment of "ordered" F can be defined by a sentence

$$\exists x_2 \exists x_4 \quad (\exists x_1 \quad [x_1 < x_2] \wedge \phi_{1,2,4}) \quad \wedge \quad (\exists x_3 \quad [x_2 < x_3] \wedge [x_3 < x_4] \wedge \phi_{2,3,4}),$$

where subformulas $\phi_{1,2,4}$ and $\phi_{2,3,4}$ have no quantifiers and define adjacencies between vertices x_1, x_2, x_4 and x_2, x_3, x_4 respectively.

In the same way it can be done for two other pairs. □

From the claim, we immediately get the result for all non-connected patterns and the star.

Corollary 4. *If $P_4 \not\subset F$, then $D_<(F) = 3$.*

Now, let us observe that Claim 3 implies that $D_<(F) = 3$ for 4-path and paw graph F as well.

Theorem 5. *If F is either P_4 or the paw graph (i.e., connected 4-vertex graph having a unique triangle), then $D_<(F) = 3$.*

Proof. We start from $F = P_4$. Let $\mathcal{F}$ be the set of all P_4 on $\{1,2,3,4\}$. Let F_0 be the path 3142. For $F \in \mathcal{F}$, let φ_F be a sentence that defines containment of an "ordered" copy of F (i.e., containment of a subgraph $\tilde{F}$ isomorphic to F with the same order of vertices as in F). From Claim 3, it immediately follows, that for $F \neq F_0$, one may choose a sentence with the quantifier depth 3 on the role of φ_F. Let us show that there exists a sentence φ_0 with the quantifier depth 3 which is tautologically equivalent to $\varphi_{F_0} \wedge \big[\bigwedge_{F\in\mathcal{F}:\, F\neq F_0} \neg\varphi_F\big]$. Indeed, since a desired copy of F_0 should not contain an edge between its second and third vertices (otherwise, there is a copy C_4 and, consequently, a "differently ordered" copy of P_4), conditioning on $\bigwedge_{F\in\mathcal{F}:\, F\neq F_0} \neg\varphi_F$ being true, it remains to say

$$\exists x_1 \exists x_2 \quad (x_1 \sim x_2) \wedge \psi_0(x_1, x_2)$$

and

$$\psi_0(x_1,x_2) = \big(\exists x_3\ [x_2 < x_3] \wedge [x_1 \sim x_3] \wedge [x_2 \nsim x_3]\big) \wedge \big(\exists x_4\ [x_2 < x_4] \wedge [x_1 \sim x_4] \wedge [x_2 \sim x_4]\big). \quad (2)$$

Clearly, the sentence $\varphi_0 \vee \big[\bigvee_{F\in\mathcal{F}:\, F\neq F_0} \varphi_F\big]$ defines the P_4-containment and has the quantifier depth 3.

It remains to prove the same but for a paw graph. Let $\mathcal{F}$ be the set of all paw graphs on $\{1,2,3,4\}$. Let $F_0 \in \mathcal{F}$ have edges $\{3,1\}, \{1,4\}, \{4,2\}, \{1,2\}$, F_0' have edges $\{3,1\}, \{1,4\}, \{4,2\}, \{3,4\}$. As above, for $F \in \mathcal{F}$, let φ_F be a sentence that defines containment of an "ordered" copy of F. From Claim 3, for $F \notin \{F_0, F_0'\}$, one may choose a sentence with the quantifier depth 3 on the role of φ_F. It remains to prove that there exist sentences φ_0 and φ_0' with the quantifier depth 3 that are tautologically equivalent to $\varphi_{F_0} \wedge \big[\bigwedge_{F\in\mathcal{F}:\, F\notin\{F_0,F_0'\}} \neg\varphi_F\big]$ and $\varphi_{F_0'} \wedge \big[\bigwedge_{F\in\mathcal{F}:\, F\notin\{F_0,F_0'\}} \neg\varphi_F\big]$ respectively. Let us prove the existence of the first sentence, the proof for the second one is analogous. The desired sentence is

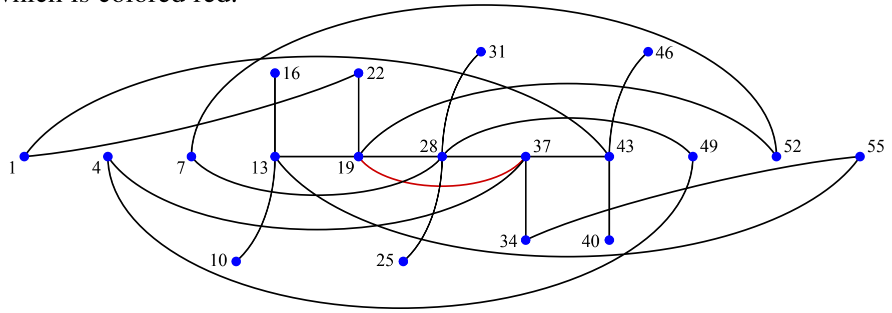
$$\exists x_1 \exists x_2 \quad (x_1 < x_2) \wedge (x_1 \sim x_2) \wedge \psi_0(x_1, x_2),$$

where ψ_0 is defined in (2). Clearly, the sentence $\varphi_0 \vee \varphi_0' \vee \big[\bigvee_{F\in\mathcal{F}:\, F\notin\{F_0,F_0'\}} \varphi_F\big]$ defines the paw-containment and has the quantifier depth 3. □

4.3. *4-patterns with C_4.* In this section, we prove that for all 4-vertex patterns F containing C_4 (i.e., for C_4, the diamond graph $K_4 \setminus e$ and K_4), $D_<(F) = 4$.

Theorem 6. *If $F \supset C_4$ is a 4-vertex graph, then $D_<(F) = 4$.*

Proof. Let us first prove the result for K_4. Consider graphs $G \supset K_4$ and $H \not\supset K_4$ on the vertex set $[55]$. Let V_i be the set of vertices $v \equiv i \mod 3$, $i \in \{0,1,2\}$. All the vertices of V_0 are isolated. All the vertices of V_2 are adjacent to all the vertices of V_1, and there are no adjacencies between the vertices of V_2. The adjacencies between the vertices of V_1 are represented in the figure; G has one more edge than H, $\{19, 37\}$, which is colored red.



By Lemma 2, it is sufficient to show that Duplicator wins the 3-game on G and H. Let us describe the winning strategy. In the first round, Duplicator just copies the Spoiler's move by choosing the same vertex: $x_1 = y_1$.

Assume that, after the Spoiler's choice in the second round, two chosen vertices (in the graph where this choice is made) do not belong to $\{19, 37\}$. Duplicator uses the same strategy as before: $x_2 = y_2$. Clearly, Duplicator wins in the third round by doing the same thing: $x_3 = y_3$. Indeed, if at least one of two vertices of $[55]$ does not belong to $\{19, 37\}$, then the adjacencies between them in G and H are the same.

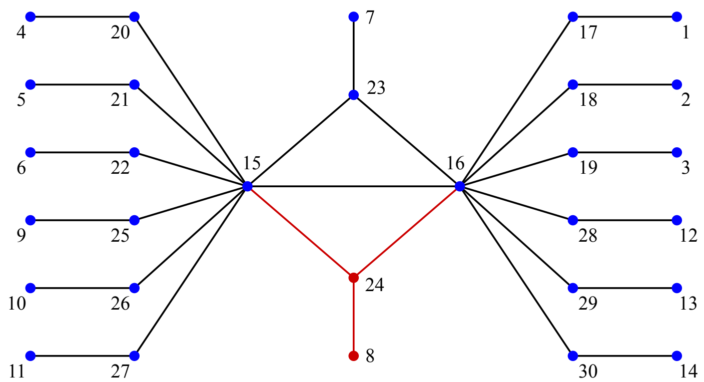
Finally, assume that exactly one of the two vertices chosen by Spoiler belongs to $\{19, 37\}$. Since the graphs are symmetric, we may assume, that this is the vertex 19. Duplicator chooses the same vertex: $x_2 = y_2$.

Clearly, if the vertex chosen by Spoiler in third round does not equal to 37, then Duplicator wins by copying the move of Spoiler. Let Spoiler choose 37. Without loss of generality, $x_2 = 19$.

Assume that Spoiler's move in the third round is in G. If $x_1 \in V_0$, then Duplicator chooses y_3 from V_2 such that, for $j \in \{1, 2\}$, $y_3 > x_j$ if and only if $37 > x_j$. Clearly, there is such a vertex in V_2. If $x_1 \in V_2$, then three situations are possible. If $x_1 < 19$, then Duplicator chooses $y_3 = 28$. If $19 < x_1 < 37$, then $y_3 = 52$. If $x_1 > 37$, then $y_3 = 28$. Assume that $x_1 \in V_1$. If $x_3 \sim x_1$, then Duplicator may find a desired vertex in V_2. If $x_3 \nsim x_1$, then two situations are possible. If $x_1 < 37$, then Duplicator chooses $y_3 = 52$. If $x_1 > 37$, then Duplicator chooses $y_3 = 22$.

It remains to assume that the last Spoiler's move is in H. If $x_1 \in V_0$, then Duplicator chooses y_3 from V_0 such that, for $j \in \{1, 2\}$, $y_3 > x_j$ if and only if $37 > x_j$. If $x_1 \in V_2$, then two situations are possible. If $x_1 < 37$, then Duplicator chooses $y_3 = 43$. If $x_1 > 37$, then $y_3 = 31$. Assume that $x_1 \in V_1$. If $x_3 \nsim x_1$, then Duplicator may find a desired vertex in V_0. If $x_3 \sim x_1$, then four situations are possible. If $x_1 = 4$, then Duplicator chooses 49 If $x_1 = 28$, then $y_3 = 49$. If $x_1 = 34$, then $y_3 = 55$. If $x_1 = 43$, then $y_3 = 40$.

Now, let us switch to C_4 and the diamond graph. Consider the graphs G and H given as follows:



Black edges and blue vertices belong to both graphs, while red vertices (8 and 24) and the edges connected to them belong only to G. Notice that $G \supset K_4 \setminus e \supset C_4$ while $H \not\supset C_4$. By Lemma 2, to finish the proof of Theorem 6, it is sufficient to show that Duplicator wins the 3-game on G and H.

Let us call the vertices that are at most 14 *children*. The vertices adjacent to *children* are called *parents* (if u is a parent of a child v, then $u = v + 16$). Let us call the vertices 23, 24, 7, 8 *central*.

Let us observe the following straightforward properties of G and H.

Claim 7. *If $x_1 = y_1$, $x_2 = y_2$, $x_3 = y_3$, then Duplicator wins.*

Claim 8. *The vertices 23, 24 and 7, 8 in the graph G are indistinguishable with respect to the non-central vertices, i.e., for any non-central vertex u, either both 23 and 24 (or 7 and 8) are adjacent to u, or both are non-adjacent to u, and either both are larger than u, or both are less than u.*

Define $f\colon \mathbb{N} \to \mathbb{N}$ in the following way:

$$f(x) = x, \text{ if } x \notin \{8, 24\}, \quad f(24) = 23, \quad f(8) = 7.$$

From Claim 7 and Claim 8, we get:

Claim 9. *If x, y, z are chosen in graph G, and $f(x)$, $f(y)$, $f(z)$ respectively in graph G and there is not more than one central vertex among x, y, z, then Duplicator wins.*

Since a parent and its child differs by the constant number, and every child is less than any parent, we get the following.

Claim 10. *Assume that (x_1, x_2, x_3), (y_1, y_2, y_3) is a winning configuration for Duplicator such that x_i is a parent if and only if y_i is a parent. If we replace a parent x_i such that its child u_i is not among x_1, x_2, x_3 with u_i, and do the same thing with y_i, then the new configuration is also winning for Duplicator.*

Now, let us change rules of the game in the following way. Spoiler is restricted to choose a parent before its child. If Spoiler chooses a parent, then Duplicator chooses a parent as well. If Spoiler chooses a son (say, x_i) of previously chosen parent (say, x_j), then Duplicator must choose the son of y_j.

According to Claim 10, if Duplicator wins the modified game in 3 rounds, then he wins the original 3-game as well. Below, we describe a winning strategy of Duplicator in the modified game.

In the first round, Spoiler chooses either a vertex x_1 in G, or a vertex y_1 in H. Duplicator's response is $f(x_1)$ (or $f(y_1)$ respectively).

If, after the second move of Spoiler, in the respective graph, both vertices are non-central, then Duplicator chooses $f(x_2)$ (or $f(y_2)$). Then, in the third round, Duplicator chooses the vertex $f(x_3)$ (or $f(y_3)$) and, by Claim 9, wins.

Assume that at least one central vertex is selected. Let in the second round Spoiler choose a vertex v in a graph where a vertex u is already chosen.

(1) If either $v \in \{15, 16\}$, or $u \in \{15, 16\}$, then Duplicator chooses $f(v)$. Without loss of generality, we assume that $v \in \{15, 16\}$. If the last vertex chosen by Spoiler x_3 (y_3) is not central, then Duplicator chooses $f(x_3)$ (resp. $f(y_3)$) and wins by Claim 9. If x_3 (y_3) is the child of $u = x_1$ (resp. $u = y_1$), then Duplicator chooses the child of y_1 (resp. x_1) in the third round and wins. Finally, if Spoiler chooses the second central parent x_3, then Duplicator chooses a neighbor of y_2, which is less than y_1, if $x_3 < x_1$, and a neighbor of y_2, larger than y_1, otherwise.

(2) If v is the child of the vertex chosen in the first round, then Duplicator chooses the respective child in the other graph and then his further winning strategy is obvious by Claim 8.

(3) Finally, assume that both vertices u, v (where u is the vertex chosen in the first round in the graph where Spoiler moves in the second one) are parents, and at least one of them is central. For a parent z, let $z\downarrow$ be the maximum parent less than z (does not exist for 17) and $z\uparrow$ be the minimum parent

larger than z (does not exist for 30). Let us assume that the vertex chosen x_1 be a central parent. Below, without loss of generality, we assume that v is in G. If $x_2 = x_1\uparrow$, then Duplicator chooses $y_2 = y_1\uparrow$. If $x_2 = x_1\downarrow$, then Duplicator chooses $y_2 = y_1\downarrow$. If $x_2 = 17$, then $y_2 = 17$. If $x_2 = 30$, then $y_2 = 30$. Otherwise, Duplicator chooses $y_1\uparrow\uparrow$ (y_1 is less than 30 in this case), if $x_2 > x_1$, and $y_1\downarrow\downarrow$ (y_1 is larger than 17 in this case), if $x_2 < x_1$.

Now, let us assume that u is not central, while v is central. The strategy of Duplicator for the second round is described below. If Spoiler's second move is in G, then the following situations are possible. If $x_1 = 25$, $x_2 = 23$, then $y_2 = 22$. If $x_1 = 25$, $x_2 = 24$, then $y_2 = 23$. If $x_1 = 22$, $x_2 = 23$, then $y_2 = 23$. If $x_1 = 22$, $x_2 = 24$, then $y_2 = 25$. Otherwise, $y_2 = f(x_2)$. If Spoiler's second move is in H, then the following situations are possible. If $y_1 = 25$, $y_2 = 23$, then $x_2 = 24$. If $y_1 = 22$, $y_2 = 23$, then $x_2 = 23$. Otherwise, $x_2 = f(y_2)$.

It remains to prove that, in both cases, Duplicator has a winning move in the last round.

Let us note that, in both graphs, the selected vertices have individual neighbors smaller than any parent, and any neighbor of any parent is less than selected vertices. Any pair of parents, one of which is central, have a common neighbor in both graphs. Also, for any pair of parents $z_1 < z_2$, such that $z_2 \neq z_1\uparrow$, $z_1 \neq 17$, $z_2 \neq 30$, among the vertices non-adjacent to both z_1, z_2 there are vertices of all 3 types: larger than both, smaller than both and between them. There is no common non-neighbor of the first type only if $z_2 = 30$. There is no common non-neighbor of the second type only if $z_1 = 17$. There is no common non-neighbor of the third type only if $z_2 = z_1\uparrow$. The winning strategy of Duplicator follows. $\square$

References

[Alon et al. 1995] N. Alon, R. Yuster, and U. Zwick, "Color-coding", *J. Assoc. Comput. Mach.* **42**:4 (1995), 844–856. MR

[He 2015] Y. He, "k variables are needed to define k-clique in first-order logic", preprint, 2015. arXiv

[Le Gall 2014] F. Le Gall, "Powers of tensors and fast matrix multiplication", pp. 296–303 in *ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ACM, New York, 2014. MR

[Libkin 2004] L. Libkin, *Elements of finite model theory*, Springer, 2004. MR

[Nešetřil and Poljak 1985] J. Nešetřil and S. Poljak, "On the complexity of the subgraph problem", *Comment. Math. Univ. Carolin.* **26**:2 (1985), 415–419. MR

[Rossman 2008] B. Rossman, "On the constant-depth complexity of k-clique", pp. 721–730 in *STOC'08: Proceedings of the 40th Annnual ACM Symposium on Theory of Computing*, ACM, New York, 2008. MR

[Rossman 2016] B. Rossman, "Lower bounds for subgraph isomorphism and consequences in first order logic", Presentation at the Workshop on Symmetry, Logic, Computation at the Simons Institute, Berkeley, November 2016, Available at https://simons.berkeley.edu/talks/benjamin-rossman-11-08-2016.

[Schweikardt 2013] N. Schweikardt, "A short tutorial on order-invariant first-order logic", pp. 112–126 in *Computer science—theory and applications*, Lecture Notes in Comput. Sci. **7913**, Springer, 2013. MR

[Verbitsky and Zhukovskii 2019a] O. Verbitsky and M. Zhukovskii, "On the first-order complexity of induced subgraph isomorphism", *Log. Methods Comput. Sci.* **15**:1 (2019), art. id˙ 25. MR

[Verbitsky and Zhukovskii 2019b] O. Verbitsky and M. Zhukovskii, "Tight bounds on the asymptotic descriptive complexity of subgraph isomorphism", *ACM Trans. Comput. Log.* **20**:2 (2019), art. id. 9. MR

[Verbitsky and Zhukovskii 2019c] O. Verbitsky and M. Zhukovskii, "The descriptive complexity of subgraph isomorphism without numerics", *Theory Comput. Syst.* **63**:4 (2019), 902–921. MR

Received 3 Dec 2019. Revised 6 Feb 2020.

OLEG GRIGORYAN:
olegrg@gmail.com
Higher School of Economics, Moscow, Russia

MIKHAIL MAKAROV:
vbif-98@mail.ru
Moscow Institute of Physics and Technology, Dolgoprudny, Russia

MAKSIM ZHUKOVSKII:
zhukmax@gmail.com
Laboratory of Advanced Combinatorics and Network Applications, Moscow Institute of Physics and Technology, Dolgoprudny, Russia

Moscow Journal of Combinatorics and Number Theory

msp.org/moscow

See inside back cover or msp.org/moscow for submission instructions.

The subscription price for 2020 is US $310/year for the electronic version, and $365/year (+$20, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Moscow Journal of Combinatorics and Number Theory (ISSN 2640-7361 electronic, 2220-5438 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

MJCNT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/

Moscow Journal of Combinatorics and Number Theory
vol. 9 no. 3 2020

Thematic Issue on
Complexity