CLIQUE-DIVERGENCE IS NOT FIRST-ORDER EXPRESSIBLE FOR THE CLASS OF FINITE GRAPHS

C. CEDILLO[1,3,4,5] AND M.A. PIZAÑA[2,3]

Abstract. The *clique graph*, $K(G)$, of a graph G is the intersection graph of its (maximal) cliques. The *iterated clique graphs* of G are then defined by: $K^0(G) = G$ and $K^n(G) = K(K^{n-1}(G))$. We say that G is *clique-divergent* if the set of orders of its iterated clique graphs, $\{|K^n(G)| : n \in \mathbb{N}\}$ is unbounded. Clique graphs and iterated clique graphs have been studied extensively, but no characterization for clique-divergence has been found so far.

Recently, it was proved that the clique-divergence is undecidable for the class of (not necessarily finite) automatic graphs [2], which implies that clique-divergence is not first-order expressible for the same class.

Here we strengthened the latter result by proving that the clique-divergence property is not first-order expressible even for the class of finite graphs. Logic expressibility has strong relations with complexity theory and consequently, new avenues of research are opened for clique graph theory.

1. Introduction

Our graphs are finite and simple. Let $\mathcal{G}$ be the class of all graphs. In *graph dynamics* [15] we are interested in the properties of the discrete dynamic system resulting from a given operator $\Phi : \mathcal{G} \to \mathcal{G}$. This setting has applications in certain approaches to loop quantum gravity [16–18] where the quantum spacetime foam is to be obtained as an emergent property from the (hypothetical) underlying discrete spacetime.

Given a graph operator Φ, we can define the corresponding *iterated* graph operators by $\Phi^0(G) = G$ and $\Phi^n(G) = \Phi(\Phi^{n-1}(G))$. One of the central topics of study in graph dynamics is that of Φ-*divergence*: A graph G is said to be Φ-divergent, if the sequence of orders $|\Phi^n(G)|_{n\in\mathbb{N}}$ grows without limit; otherwise, we say that G is Φ-*convergent* (in which case, we necessarily have that $\Phi^n(G) \cong \Phi^m(G)$, for some $n < m$). Φ-divergence have been fully characterized for many graphs operators, including the characterization of convergence for iterated line graphs [14], the characterization for iterated biclique graphs [6] among others [15].

The *clique operator*, K, however is widely considered one of the most complex ones [15] and a characterization of K-divergence (or clique-divergence) has resisted all attempts during the 50 years since the notion of iterated clique graphs was introduced in [7]. A growing consensus among experts is that clique-convergence

[1]Email:mc.cedilloc@gmail.com
[2]Email:mpizana@gmail.com
[3]Universidad Autónoma Metropolitana - Iztapalapa, Mexico City, Mexico.
[4]Centro Universitario UAEM Nezahualcóyotl, Nezahualcoyotl City, Mexico.
[5]Partially supported by CONACyT, scholarship 397900.
Key words and phrases. graph theory, graph dynamics, clique graphs, first-order logic, Ehrenfeucht-Fraïssé game.

might be undecidable although no substantial progress has been made in this direction either. Hence, other measures of difficulty have been pursued: Recently [2], it was shown that clique-divergence is undecidable for the class of (not necessarily finite) automatic graphs, and consequently, that clique-divergence is not first order expressible for the same class.

Here we extend the latter result by proving that clique-divergence is not first-order expressible even for the class of finite graphs. Logic expressibility has many known relations to complexity theory [8, 13] and hence new approaches to understand the difficulty of deciding clique-divergence are opened.

2. Preliminaries

We refer the reader to the standard literature for logic [3], finite model theory [13], graphs [1] and clique graphs [20]. In what follows, we briefly review some of the needed terminology and results.

Given a graph G and $x, y \in V(G)$, $A, B \subseteq V(G)$, the *distance* from x to y in G is denoted by $d_G(x,y)$; we also use $d_G(x,A) = \min\{d_G(x,y) : y \in A\}$ and $d_G(A,B) = \min\{d_G(x,y) : x \in A, y \in B\}$. The *closed neighborhood* of a vertex x in a graph G is denoted by $N_G[x]$ or $N[x]$. A vertex x is *dominated* by $y \neq x$ whenever $N[x] \subseteq N[y]$. A graph G is *dismantleable* to H if H can be obtained from G by removing dominated vertices iteratively, for instance, an n-path graph is dismantleable to the one-vertex graph. A *clique* of G, is a maximal complete subgraph G. The *clique graph*, $K(G)$, of a graph G is the intersection graph of all its cliques. Here, K is the *clique operator*. *Iterated clique graphs* are defined inductively by: $K^0(G) = G$ and $K^{n+1}(G) = K(K^n(G))$. A graph G is K-*convergent* (or *clique-convergent*) if $K^n(G) \cong K^m(G)$ for some $n < m$; otherwise it is K-*divergent* (or *clique-divergent*).

Theorem 2.1. [5, Thm. 5] *If G is dismantleable to H, G and H have the same K-behavior. In particular, if x is a dominated vertex of G, G and $G - \{x\}$ have the same K-behavior.*

We shall need the following two families of graphs (see figures 1(a) and 1(b)):

Definition 2.2. *Let $m \geq 2$. The graph X_m has vertex set $V(X_m) = \mathbb{Z}_{6m}$ and adjacencies given by: $x \sim y$ whenever $x \equiv 1 \mod 3$ and $y - x \in \{1,3\}$, or when $x \not\equiv 1 \mod 3$ and $y - x \in \{1,2,3\}$ (or when $y \sim x$ according to the preceding rules). The graph Y_m is obtained from X_m by removing exactly one edge: $\{6m-1, 2\} \in E(X_m)$.*

The graphs just defined belong to the class of *clockwork graphs*. The clique behavior of clockwork graphs is very well understood [4, 10, 11] and there is even a polynomial time algorithm for deciding it [12]. We shall need to know the clique behavior of X_m and Y_m:

Remark 2.3. *For all $m \geq 2$, X_m is K-convergent and Y_m is K-divergent.*

Proof. In the terminology of [12], X_m and Y_m are both clockwork graphs with no *covered vertices* and with zero and one *good segments* respectively. It follows by the algorithm described in [12, Thm 3.6] that $|K^n(X_m)| = |X_m| + 0 \cdot n = 6m$ and that $|K^n(Y_m)| = |Y_m| + 1 \cdot n = 6m + n$. Hence X_m is K-convergent and Y_m is K-divergent. □

In fact, X_m and Y_m are very simple cases of clockwork graphs and the interested reader, would find it not too difficult to produce a stand-alone proof of the previous remark. Indeed, it can be readily verified that X_m is *self-clique* ($K(X_m) \cong X_m$) and with some extra effort, that $K^n(Y_m)$ is always an induced subgraph of $K^{n+1}(Y_m)$, with the latter having exactly one extra vertex.

In model theory, a (relational) *signature* σ is a tuple of relational symbols $(R_1, R_2, \ldots, R_s)$, where each R_i has some associated arity $r_i \in \mathbb{N}$. Given a signature σ, a σ*-structure* (also called a *model*) is a tuple $\mathfrak{A} = (A, R_1^{\mathfrak{A}}, R_2^{\mathfrak{A}}, \ldots, R_s^{\mathfrak{A}})$ consisting of a *domain* A and for each symbol R_i in σ, an r_i-ary relation $R_i^{\mathfrak{A}} \subseteq A^{r_i}$. The first order logic (FO) sentences are then the sentences we can form using these relational symbols, together with the logic symbols ($\exists$, $\forall$, $=$, $\neg$, $\vee$, $\wedge$), variables (x, y, x_1, x_2, $\ldots$) and parenthesis. It is important to note that, in this context, quantifiers refer always exclusively to the domain A.

In the case of graph theory, the signature is usually simply $\sigma = (\sim)$ where '$\sim$' is the symbol for the adjacency relation. Then a σ-structure is simply a graph $\mathfrak{A} = (V, E)$, with $E \subseteq V^2$ (an unoriented edge $x \sim y$ is represented here by the collection of both ordered pairs $\{(x,y),(y,x)\} \subseteq E$, as it is common practice in the literature). We shall use the standard LaTeX typeface "mathnormal" $(A, B, \ldots)$ for denoting graphs unless the result being considered is more model-theoretic than graph-theoretic in which case, we shall use the fraktur typeface ($\mathfrak{A}$, $\mathfrak{B}$, $\ldots$).

Given a pair of graphs $\mathfrak{A}$, $\mathfrak{B}$, and an integer $k \in \mathbb{N}$, a *k-round Ehrenfeucht-Fraïssé game* [13, 19] on $\mathfrak{A}$ and $\mathfrak{B}$ is played by two players *Duplicator* and *Spoiler*: At each round, Spoiler first selects a vertex x either in $\mathfrak{A}$ or in $\mathfrak{B}$ (his choice) and Duplicator replies by selecting some vertex y on the other graph. The winner is determined as follows: after k-rounds, some vertices $a_1, a_2, \ldots, a_k$ are selected in $\mathfrak{A}$ and some other vertices $b_1, b_2, \ldots, b_k$ are selected in $\mathfrak{B}$ (the subscript indicate the round in which the vertex was selected, repetitions of vertices are allowed, at this point it does not matter which player selected which vertex), then Duplicator wins the game when the mapping given by $a_i \mapsto b_i$ is an isomorphism from the subgraph of $\mathfrak{A}$ induced by $\{a_1, a_2, \ldots, a_k\}$ to the subgraph of $\mathfrak{B}$ induced by $\{b_1, b_2, \ldots, b_k\}$; otherwise, Spoiler wins the game. We say that Duplicator has a *k-round winning strategy* if there is a way in which Duplicator can play to guarantee the victory after k rounds, no matter how Spoiler plays.

Theorem 2.4. (Ehrenfeucht-Fraïssé, [13, Cor. 3.10]) *A property P of finite σ-structures is not expressible in FO if for every $k \in \mathbb{N}$, there exist two finite σ-structures, $\mathfrak{A}_k$ and $\mathfrak{B}_k$, such that:*

(1) *Duplicator has a k-round winning strategy for the Ehrenfeucht-Fraïssé game on $\mathfrak{A}_k$ and $\mathfrak{B}_k$.*
(2) *$\mathfrak{A}_k$ has property P, and $\mathfrak{B}_k$ does not.*

3. Cliques and games

Look at the graphs in Figure 1, (c) and (d), they are not isomorphic, but there are a pair of vertex bijections which are almost isomorphisms: a simple translation works well except in the red regions, while a reflection (followed by a translation) works well except in the green regions. This is the notion we want to capture with the definitions of quasi-isomorphisms and quasi-isomorphic graphs in the following paragraphs.

Given two graphs A, B a *quasi-isomorphism* is a bijection $f : V(A) \to V(B)$ together with two *failure regions* $X \subseteq V(A)$ and $X' \subseteq V(B)$ such that whenever f fails to be an isomorphism at a pair of vertices x, y (i.e. when $x \sim y$ and $f(x) \nsim f(y)$ or when $x \nsim y$ and $f(x) \sim f(y)$), we have that $x, y \in X$ and $f(x), f(y) \in X'$.

A pair of graphs A, B are *quasi-isomorphic with failure distance s* if there are two quasi-isomorphisms $r : V(A) \to V(B)$ with failure regions G and G' and $g : V(A) \to V(B)$ with failure regions R and R' such that $r(G) = G' = g(G)$ and $r(R) = R' = g(R)$ and such that $d_A(R, G) = s$.

Note that, since any minimal length path from R to G (resp. R' to G') does not use edges contained in G (resp. G'), and since r is an isomorphism save for the edges contained in G and G', we must have $d_B(R', G') = d_A(R, G)$.

Theorem 3.1. *If $\mathfrak{A}$ and $\mathfrak{B}$ are quasi-isomorphic graphs with failure distance greater than 2^k, then Duplicator has a k-round winning strategy for the Ehrenfeucht-Fraïssé game on $\mathfrak{A}$ and $\mathfrak{B}$.*

Proof. Assume that the bijections and failure regions are r, g, G, G', R and R' as above. We describe Duplicator's strategy recursively:

At each round i with $1 \le i \le k$, there will be red and green regions $R_i, G_i \subseteq V(\mathfrak{A})$ and $R'_i, G'_i \subseteq V(\mathfrak{B})$ of previously selected vertices (initially set $R_1 = R$, $R'_1 = R'$, $G_1 = G$, $G'_1 = G'$). Assume first that Spoiler selects a vertex $x \in \mathfrak{A}$. The idea is that Duplicator replies with $r(x)$ or $g(x)$ according to whether x is closer to the red region R_i or it is closer to green region G_i, that is, Duplicator selects a vertex $dup(x) \in \mathfrak{B}$ according to the following rule:

$$dup(x) = \begin{cases} r(x) & \text{if } d_{\mathfrak{A}}(x, R_i) \le d_{\mathfrak{A}}(x, G_i), \\ g(x) & \text{if } d_{\mathfrak{A}}(x, R_i) > d_{\mathfrak{A}}(x, G_i). \end{cases}$$

After that, we compute the new red and green regions as follows: if $dup(x) = r(x)$ then set $R_{i+1} = R_i \cup \{x\}$, $R'_{i+1} = R'_i \cup \{r(x)\}$, $G_{i+1} = G_i$ and $G'_{i+1} = G'_i$; else (when $dup(x) = g(x)$) set $G_{i+1} = G_i \cup \{x\}$, $G'_{i+1} = G'_i \cup \{g(x)\}$, $R_{i+1} = R_i$, $R'_{i+1} = R'_i$.

If Spoiler selected $x \in \mathfrak{B}$ instead, Duplicator replies analogously with $r^{-1}(x)$ or $g^{-1}(x)$ according to whether x is closer to R'_i or G'_i; the new sets R_{i+1} R'_{i+1}, G_{i+1}, G'_{i+1} are also computed analogously in this case: if $dup(x) = r^{-1}(x)$ then set $R_{i+1} = R_i \cup \{r^{-1}(x)\}$, $R'_{i+1} = R'_i \cup \{x\}$, $G_{i+1} = G_i$ and $G'_{i+1} = G'_i$; else (when $dup(x) = g^{-1}(x)$) set $G_{i+1} = G_i \cup \{g^{-1}(x)\}$, $G'_{i+1} = G'_i \cup \{x\}$, $R_{i+1} = R_i$, $R'_{i+1} = R'_i$.

We claim that for $1 \le i \le k+1$, $d_{\mathfrak{A}}(R_i, G_i) = d_{\mathfrak{B}}(R'_i, G'_i) > 2^{k-i+1}$: Indeed for $i = 1$ we have $d_{\mathfrak{A}}(R_1, G_1) > 2^k$ by hypothesis, and $d_{\mathfrak{B}}(R'_1, G'_1) = d_{\mathfrak{A}}(R_1, G_1)$ as noted before. For the induction step, let us assume first that $x \in \mathfrak{A}$ is closer to the red region R_i, then $2^{k-i+1} < d_{\mathfrak{A}}(R_i, G_i) \le d_{\mathfrak{A}}(R_i, x) + d_{\mathfrak{A}}(x, G_i) \le 2d_{\mathfrak{A}}(x, G_i)$. It follows that $d_{\mathfrak{A}}(x, G_i) > 2^{k-i} = 2^{k-(i+1)+1}$. Now, observe that $d_{\mathfrak{A}}(R_{i+1}, G_{i+1}) = \min\{d_{\mathfrak{A}}(R_i, G_i), d_{\mathfrak{A}}(x, G_i)\} > 2^{k-(i+1)+1}$. Also, we must have that $d_{\mathfrak{B}}(R'_{i+1}, G'_{i+1}) = d_{\mathfrak{A}}(R_{i+1}, G_{i+1})$ since any minimal length path from the red to the green region does not use edges from within those regions and since $r : V(\mathfrak{A}) \to V(\mathfrak{B})$ is an isomorphism outside those regions, hence, r maps minimal length paths from R_{i+1} to G_{i+1} onto minimal length paths from R'_{i+1} to G'_{i+1}. It follows that $d_{\mathfrak{B}}(R'_{i+1}, G'_{i+1}) = d_{\mathfrak{A}}(R_{i+1}, G_{i+1}) > 2^{k-(i+1)+1}$ in this case. The other three cases are analogous. This concludes the proof of the claim.

Now, the described strategy is a $k-round$ winning strategy for Duplicator since, after the k-th round, we have $d_{\mathfrak{A}}(R_{k+1}, G_{k+1}) = d_{\mathfrak{B}}(R'_{k+1}, G'_{k+1}) > 2^{k-(k+1)+1} = 1$, which means that the red and the green regions never get in touch (i.e. distance at least 2) and hence there is an isomorphism from the subgraph of $\mathfrak{A}$ induced by the selected vertices to the subgraph of $\mathfrak{B}$ induced by the selected vertices: namely the isomorphism that acts as r on the red region R_{k+1} and acts as g on the green region G_{k+1}. □

We define now the graphs A_m and B_m (see Figure 1 (c) and (d)). To construct A_m take a copy of X_m and a copy of Y_m, rename the vertices of X_m as $\overline{0}, \overline{1}, \dots, \overline{6m-1}$, now identify the vertex $\overline{0} \in X_m$ and the vertex $0 \in Y_m$, finally, remove the edges connecting the set $\{3m-1, 3m, 3m+1\}$ to the set $\{3m+2, 3m+3, 3m+4\}$, the resulting graph is A_m. Now B_m is obtained from A_m by removing the edge $\{\overline{2}, \overline{6m-1}\}$ and adding the edge $\{2, 6m-1\}$.

Theorem 3.2. *Clique-divergence is not first-order expressible.*

Proof. Define $\mathfrak{A}_k = A_{2^k+2}$ and $\mathfrak{B}_k = B_{2^k+2}$. Observe that the failure distance is greater than 2^k in both graphs: for $k = 1$ ($m = 4$) we can readily see that in figures 1(c) and 1(d), and whenever m increases by 1, the gap distance also increases by 1. It follows by Theorem 3.1 that Duplicator has a k-round winning strategy for the Ehrenfeucht-Fraïssé game on $\mathfrak{A}_k$ and $\mathfrak{B}_k$.

Also, observe that A_m dismantles to X_m and that B_m dismantles to Y_m: For instance, in Figure 1(c), the vertices 12 and 13 of A_4 are dominated by the vertex 11; also the vertex 14 is dominated by vertex 15. After removing vertices 12, 13 and 14, the vertices 11, 15 and 16 become dominated. Clearly the process can be continued until only a copy of X_4 remains, and clearly this can be done for all m. Similar considerations apply to B_m. It follows by Theorem 2.1 and Remark 2.3 that every A_m is clique-convergent and every B_m is clique-divergent.

The result now follows from Theorem 2.4. □

Some open problems immediately arise: Is clique-divergence second-order (SO) expressible? If it is, is it also expressible in some known fragment of second-order logic (monadic-SO, FSO [2, 9], etc.)? Is clique-divergence NP-hard? Is clique-divergence computable [2, 12]?

We remark that the given families of graphs A_m and B_m can be easily differentiated using monadic second order logic.

Acknowledgment We thank the anonymous referee for finding some typos in a previous version of this paper.

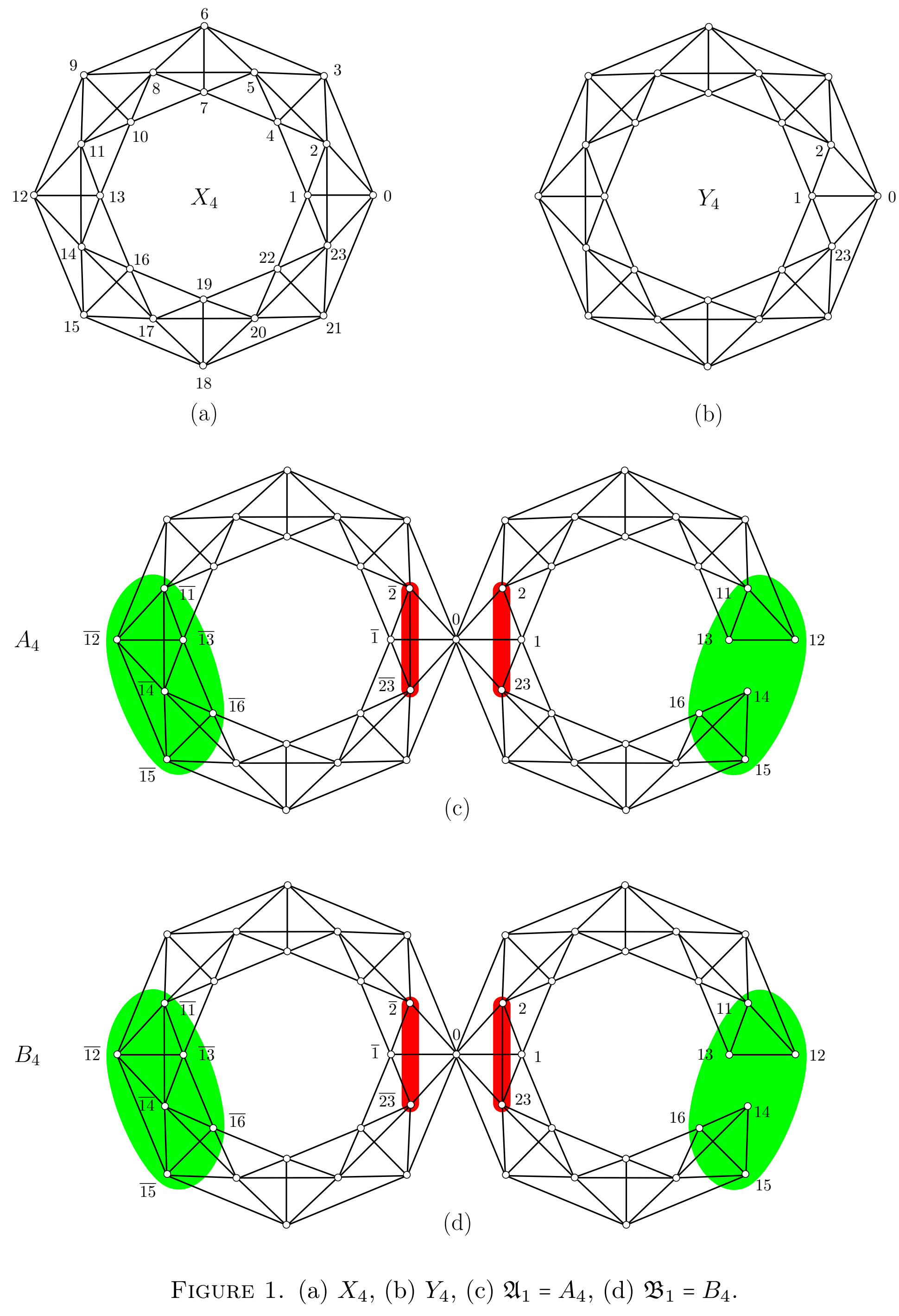

FIGURE 1. (a) X_4, (b) Y_4, (c) $\mathfrak{A}_1 = A_4$, (d) $\mathfrak{B}_1 = B_4$.

References

[1] J.A. Bondy and U.S.R. Murty. *Graph theory with applications.* American Elsevier Publishing Co., Inc., New York, 1976.

[2] C. Cedillo and M.A. Pizaña. *Clique-convergence is undecidable for automatic graphs.* Journal of Graph Theory **96** (2021) 414–437. `https://doi.org/10.1002/jgt.22622`.

[3] H.B. Enderton. *A mathematical introduction to logic.* Harcourt/Academic Press, Burlington, MA, second edition, 2001.

[4] F. Escalante. *Über iterierte Clique-Graphen.* Abh. Math. Sem. Univ. Hamburg **39** (1973) 58–68.

[5] M.E. Frías-Armenta, V. Neumann-Lara and M.A. Pizaña. *Dismantlings and iterated clique graphs.* Discrete Math. **282** (2004) 263–265.

[6] M. Groshaus, A.L.P. Guedes and L. Montero. *Almost every graph is divergent under the biclique operator.* Discrete Appl. Math. **201** (2016) 130–140. `10.1016/j.dam.2015.07.022`.

[7] S.T. Hedetniemi and P.J. Slater. *Line graphs of triangleless graphs and iterated clique graphs.* Springer Lecture Notes in Math. **303** (1972) 139–147.

[8] N. Immerman. *Upper and lower bounds for first order expressibility.* J. Comput. System Sci. **25** (1982) 76–98. `https://doi.org/10.1016/0022-0000(82)90011-3`.

[9] D. Kuske. Theories of automatic structures and their complexity. In *Algebraic informatics*, volume 5725 of *Lecture Notes in Comput. Sci.*, pages 81–98. Springer, Berlin, 2009. `http://dx.doi.org/10.1007/978-3-642-03564-7_5`.

[10] F. Larrión and V. Neumann-Lara. *A family of clique divergent graphs with linear growth.* Graphs Combin. **13** (1997) 263–266.

[11] F. Larrión and V. Neumann-Lara. *On clique-divergent graphs with linear growth.* Discrete Math. **245** (2002) 139–153.

[12] F. Larrión, V. Neumann-Lara and M.A. Pizaña. *Clique divergent clockwork graphs and partial orders.* Discrete Appl. Math. **141** (2004) 195–207.

[13] L. Libkin. *Elements of finite model theory.* Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2004 `https://doi.org/10.1007/978-3-662-07003-1`.

[14] E. Prisner. *The dynamics of the line and path graph operators.* Graphs and Combinatorics **9** (1993) 335–352.

[15] E. Prisner. *Graph dynamics.* Longman, Harlow, 1995.

[16] M. Requardt. *(Quantum) spacetime as a statistical geometry of lumps in random networks.* Classical Quantum Gravity **17** (2000) 2029–2057.

[17] M. Requardt. Space-time as an order-parameter manifold in random networks and the emergence of physical points. In *Quantum theory and symmetries (Goslar, 1999)*, pages 555–561. World Sci. Publ., River Edge, NJ, 2000.

[18] M. Requardt. *A geometric renormalization group in discrete quantum spacetime.* J. Math. Phys. **44** (2003) 5588–5615.

[19] J. Spencer. *The strange logic of random graphs*, volume 22 of *Algorithms and Combinatorics.* Springer-Verlag, Berlin, 2001.

[20] J.L. Szwarcfiter. A survey on clique graphs. In B.A. Reed and C. Linhares-Sales, editors, *Recent advances in algorithms and combinatorics*, volume 11 of *CMS Books Math./Ouvrages Math. SMC*, pages 109–136. Springer, New York, 2003.