UDC 519.171.1+519.176

B. V. Novikov, L. Yu. Polyakova, G. N. Zholtkevich (Karazin Nat. Univ., Kharkov, Ukraine)

A DECOMPOSITION OF DIRECTED GRAPHS AND THE TURÁN PROBLEM

ОДИН РОЗКЛАД ОРІЄНТОВАНИХ ГРАФІВ І ЗАДАЧА ТУРАНА

We consider vertex decompositions of (di)graphs appearing in the Automata Theory and establish some properties of these decompositions. We apply these decompositions to the problem of forbidden subgraphs.

Розглянуто вершинні декомпозиції (ор)графів, що виникають у теорії автоматів, встановлено деякі їх властивості, а також наведено застосування їх до задачі про заборонені підграфи.

Introduction. This note has arisen from attempts to extend on pre-automata [1] the concepts of regions and intervals used in the translation theory [2]. Unlike the known model, the uniqueness of the header of an interval is an unacceptable condition for pre-automata. So we had to consider a generalized problem; and it was convenient to collect obtained graph-theoretic results in a separate article.

General definitions and results are given in Section 1. Regions and intervals are considered in Section 2 as a special case. Next we study the decompositions of undirected graphs (Section 3) and in particular consider their connection with maximal matchings. In Section 4 we study the main application of decompositions — the problem of forbidden graphs. Note that this problem can be posed also for digraphs. We hope that in this case our construction will be even more useful.

Mainly, we will use the definitions and notations of [3]. Thus by the **directed graph** (or **digraph**) we mean a pair $G = (V, E)$ where $V = V(G)$ is a set whose elements are called **vertices**; $E = E(G) \subset V \times V$ is a binary relation whose elements are called **arcs**.

For every vertex $v \in V$ define the sets $\mathrm{D}^-(v) = \{u \in V \mid (u, v) \in E\}$ and $\mathrm{D}^+(v) = \{u \in V \mid (v, u) \in E\}$ whose elements are called the **inputs** and **outputs** of the vertex v respectively. Note that loops are included both in $\mathrm{D}^-(v)$ and $\mathrm{D}^+(v)$.

The numbers of inputs and outputs are denoted by $\mathrm{d}^-(v)$ and $\mathrm{d}^+(v)$ respectively.

The subgraph of (di)graph G generated by a subset of vertices $U \subset V$ is denoted by $G[U]$. A subset $U \subset V$ is called **connected** if the graph $G[U]$ is connected (i.e., for any two vertices $u, v \in U$ there exists a directed path in $G[U]$ starting at u and ending at v).

The symbol $\bigsqcup$ is used for the union of disjoint sets.

1. Inflation and stability.

Let $G = (V, E)$ be a digraph.

Definition 1.1. *An inflation of a set $U \subset V$ is a set*

$$\operatorname{Inf} U = U \cup \{v \in V \mid \varnothing \neq \mathrm{D}^-(v)\varnothing U\}.$$

We need a property of the inflation:

Proposition 1.1. *For any subsets $X, Y \subset V(G)$*

$$\operatorname{Inf} X \cap \operatorname{Inf} Y = (X \cap \operatorname{Inf} Y) \cup (\operatorname{Inf} X \cap Y) \cup \operatorname{Inf}(X \cap Y).$$

Proof. The inclusion $\supseteq$ is obvious. Prove the converse. The case $v \in X \cup Y$ is also evident. Let $v \in (\operatorname{Inf} X \cap \operatorname{Inf} Y) \setminus (X \cup Y)$. Then by the definition of inflation $\varnothing \neq \mathrm{D}^-(v) \subset X \cap Y$, i.e., $v \in \operatorname{Inf}(X \cap Y)$.

Proposition 1.1 is proved.

We can consider the operator $\operatorname{Inf} : \mathbf{2}^V \to \mathbf{2}^V : U \mapsto \operatorname{Inf} U$ and its iterations Inf^n, $n > 0$. Suppose, in addition, $\operatorname{Inf}^0 U = U$.

Definition 1.2. *A hyperinflation of a subset U is a set*

$$\operatorname{Inf}^\infty U = \bigcup_{n \geq 0} \operatorname{Inf}^n U.$$

Sometimes we say that U is a hyperinflation if $U = \operatorname{Inf}^\infty U'$ for some U'.

The following statement will be used bellow:

Lemma 1.1. *Let $U \subset V$ and $v \in V \setminus \operatorname{Inf}^\infty U$. Then $\mathrm{D}^+(v) \cap \operatorname{Inf}^\infty U \subset U$.*

Proof. The case $\mathrm{D}^+(v) \cap \operatorname{Inf}^\infty U = \varnothing$ is obvious. Let $u \in \mathrm{D}^+(v) \cap \operatorname{Inf}^\infty U \neq \varnothing$, $u \notin U$. Then $u \in \operatorname{Inf}^n U = \operatorname{Inf}(\operatorname{Inf}^{n-1} U)$ for some $n \geq 0$. This contradicts the fact that $\mathrm{D}^-(u) \ni v \notin \operatorname{Inf}^{n-1} U$.

Lemma 1.1 is proved.

We define the notion of a hull which is close to the hyperinflation.

Definition 1.3. *A set $U \subset V$ is called stable if $\operatorname{Inf} U = U$.*

Lemma 1.2. *An intersection of stable sets is stable.*

Proof. Let U_i, $i \in I$, be stable sets, $X = \bigcap_{i \in I} U_i$, and $v \in \operatorname{Inf} X \setminus X$. By the definition of the inflation $\varnothing \neq \mathrm{D}^-(v) \subseteq X \subseteq U_i$ for every $i \in I$. Therefore, $v \in \operatorname{Inf} U_i = U_i$ whence $v \in X$.

Lemma 1.2 is proved.

Since the set V of vertices is stable, we have the following corollary.

***Corollary* 1.1.** *For every vertex set $U \subseteq V$ there exists the smallest stable set $\operatorname{Hull} U$ containing U.*

Definition 1.4. *We say that $\operatorname{Hull} U$ is the hull of a set U.*

It is clear that $\operatorname{Hull} U$ is the intersection of all stable sets containing U.

Consider connections between the introduced concepts. From $U \subset \operatorname{Inf} U$ it follows $\operatorname{Inf}^\infty U \subset \subset \operatorname{Hull} U$. The reverse inclusion is true if and only if the hyperinflation is stable. The following example shows that, generally speaking, for infinite digraphs this does not hold.

***Example* 1.1.** Consider a digraph $G = (V, E)$ such that

$$V = \{0, 1, 2, \ldots\}, \qquad E = \big\{(n, 0) \mid n \geq 1\big\} \cup \big\{(n, n+1) \mid n \geq 1\big\}.$$

Then $\operatorname{Inf}^n\{1\} = \{1, 2, \ldots, n+1\}$, whence $\operatorname{Inf}^\infty\{1\} = V \setminus \{0\}$. On the other hand, $\operatorname{Hull}\{1\} = = \operatorname{Inf}\big(\operatorname{Inf}^\infty\{1\}\big) = V$.

Definition 1.5. *We call a digraph $G = (V, E)$ locally d^--finite if $\mathrm{d}^-(v) < \infty$ for all $v \in V$.*

Proposition 1.2. *If a digraph $G = (V, E)$ is locally d^--finite, then $\operatorname{Inf}^\infty U = \operatorname{Hull} U$ for every $U \subset V$.*

Proof. Suppose the contrary. Let $v \in \operatorname{Inf}(\operatorname{Inf}^\infty U) \setminus \operatorname{Inf}^\infty U$. As $d^-(v) < \infty$ and $\varnothing \neq \mathrm{D}^-(v) \subseteq \subseteq \operatorname{Inf}^\infty U$, we have $\mathrm{D}^-(v) \subset \operatorname{Inf}^n U$ for some $n \geq 0$. But then $v \in \operatorname{Inf}^{n+1} U \subset \operatorname{Inf}^\infty U$ contrary to assumption.

Proposition 1.2 is proved.

The statement similar to Lemma 1.1 is not true for the hull:

***Example* 1.2.** Add the vertex -1 and the arc $(-1, 0)$ to the digraph G from Example 1.1. Obviously $\{0\} = \mathrm{D}^-(-1) \cap \operatorname{Hull}\{1\} \neq \{1\}$.

Now we introduce the main definition of this article:

Definition 1.6. *A decomposition of a digraph $G = (V, E)$ is a set of its subgraphs $G_i = (V_i, E_i)$, $i \in I$, such that*:

(i) *V_i are hyperinflations*,

(ii) *V is a disjoint union of V_i*,

(iii) *E_i is a restriction of E to V_i*.

***Remark* 1.1.** Our definition differs from the one given, for example, in [4], where a decomposition means a partition of $E(G)$.

Till the end of Section 2 we assume that some locally d^--finite graph $G(V, E)$ is fixed.

Theorem 1.1. *For any subsets $X, Y \subset V$*

$$\operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y = \operatorname{Inf}^{\infty}\big[(X \cap \operatorname{Inf}^{\infty} Y) \cup (\operatorname{Inf}^{\infty} X \cap Y)\big].$$

Proof. The inclusion $\supseteq$ is clear. Indeed,

$$(X \cap \operatorname{Inf}^{\infty} Y) \cup (\operatorname{Inf}^{\infty} X \cap Y) \subset \operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y;$$

and the set $\operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y$ is stable because of the locally d^--finiteness of the original graph and Lemma 1.2.

Let $x \in \operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y$. Then there are integers $m, n \geq 0$ such that $x \in \operatorname{Inf}^{m} X \cap \operatorname{Inf}^{n} Y$, $x \notin \operatorname{Inf}^{m-1} X \cup \operatorname{Inf}^{n-1} Y$, and

$$\mathrm{D}^-(x) \subset \operatorname{Inf}^{m-1} X \cap \operatorname{Inf}^{n-1} Y \subset \operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y.$$

Use the induction on $m + n$ to show that

$$x \in \operatorname{Inf}^{\infty}\big[(X \cap \operatorname{Inf}^{\infty} Y) \cup (\operatorname{Inf}^{\infty} X \cap Y)\big]. \tag{1}$$

If $m = 0$ or $n = 0$, then $x \in X \cup Y$; hence (1) is hold.

If $m = n = 1$, then $\mathrm{D}^-(x) \subset X \cap Y$. Therefore, $x \in \operatorname{Inf}(X \cap Y) \subset \operatorname{Inf}^{\infty}(X \cap Y) \subset$ $\subset \operatorname{Inf}^{\infty}\big[(X \cap \operatorname{Inf}^{\infty} Y) \cup (\operatorname{Inf}^{\infty} X \cap Y)\big]$.

Consider the general case. Since $\mathrm{D}^-(x) \subset \operatorname{Inf}^{m-1} X \cap \operatorname{Inf}^{n-1} Y \subset \operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y$, by the induction assumption

$$\mathrm{D}^-(x) \subset \operatorname{Inf}^{\infty}\big[(X \cap \operatorname{Inf}^{\infty} Y) \cup (\operatorname{Inf}^{\infty} X \cap Y)\big].$$

Consequently (1) is true.

Theorem 1.1 is proved.

***Corollary* 1.2.** $\operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y = \varnothing$ *if and only if*

$$\operatorname{Inf}^{\infty} X \cap Y = X \cap \operatorname{Inf}^{\infty} Y = \varnothing.$$

As we will see below, the hyperinflations of connected subsets are of particular interest.

Theorem 1.2. *Let $X, Y \subset U$, $\operatorname{Inf}^{\infty} X \cap \operatorname{Inf}^{\infty} Y \neq \varnothing$, and X is connected. If* $\operatorname{Inf}^{\infty} X \cap Y = \varnothing$, *then* $\operatorname{Inf}^{\infty} X \subset \operatorname{Inf}^{\infty} Y$.

Proof. Corollary 1.2 implies $X \cap \operatorname{Inf}^{\infty} Y \neq \varnothing$. If X is a singleton, then the statement is obvious. Let $|X| > 1$. We choose the smallest n such that $X \cap \operatorname{Inf}^{n} Y \neq \varnothing$ (it follows from the conditions that $n \geq 1$). Let $x \in X \cap \operatorname{Inf}^{n} Y$. Since $x \notin Y$, we have $\mathrm{D}^{-}(x) \subset \operatorname{Inf}^{n-1} Y$. But $\mathrm{D}^{-}(x) \cap X \neq \varnothing$, because X is connected. Therefore, $X \cap \operatorname{Inf}^{n-1} Y \neq \varnothing$; that contradicts the choice of n.

Theorem 1.2 is proved.

Theorem 1.2 allows us to construct (not uniquely) decompositions of finite graphs whose components are hyperinflations of connected sets. Describe the process of constructing in detail.

Let V be a set of vertices of the graph. We take as V_1 an arbitrary connected subset (for example, a vertex). Suppose we have taken components $V_1, \ldots, V_k$ with disjoint hyperinflations. In the complement

$$U = V \setminus \bigsqcup_{i \leq k} \operatorname{Inf}^{\infty} V_i$$

choose an arbitrary connected subset W. If $\operatorname{Inf}^{\infty} W \cap \operatorname{Inf}^{\infty} V_j \neq \varnothing$ for some $1 \leq j \leq k$, then $\operatorname{Inf}^{\infty} V_j \subset \operatorname{Inf}^{\infty} W$ by Theorem 1.2. In this case replace all V_j by W and go on to the choice of the next component. If $\operatorname{Inf}^{\infty} W \cap \bigsqcup_{i \leq k} \operatorname{Inf}^{\infty} V_i = \varnothing$, then we put $V_{k+1} = W$ and continue the process.

2. Regions and intervals. Using the terminology of Computer Science [2] we introduce the following definition.

Definition 2.1. *A subset $U \subset V$ is said to be a region if $U = \operatorname{Inf}^{\infty}\{x\}$ for some $x \in V$. In this case x is called a heading of U. A region is called an interval if it is not contained in any other region.*

Generally speaking, the region can have multiple headings. A sufficient condition for the uniqueness of the heading (this demand is essential for Computer Science) is obtained directly from Lemma 1.1:

Proposition 2.1. *Let $U \subset V$ be a region with a heading x. If there exists $y \in V \setminus U$ such that $\mathrm{D}^{+}(y) \cap U \neq \varnothing$, then x is uniquely defined, i.e., $\mathrm{D}^{+}(y) \cap U = \{x\}$.*

Since a singleton is connected, it follows directly from Theorem 1.2:

Proposition 2.2. *If two regions have a nonempty intersection, then one of them contains the other.*

Now we can state the main result about intervals of finite digraphs:

Theorem 2.1. *Every digraph $G = (U, E)$ with the finite set of vertices has the unique decomposition whose components are intervals.*

Proof. The existence of such a decomposition follows directly from Proposition 2.2; the uniqueness follows from the maximality of each interval.

Consider two extreme cases.

Proposition 2.3. *All components of an interval decomposition of a digraph are singletons if and only if $\mathrm{d}^{-}(v) = 1$ implies $\mathrm{D}^{-}(v) = \big\{(v, v)\big\}$ for any vertex v.*

Proof. Let $G = (V, E)$ be a considered digraph. It is clear that all its components are singletons if and only if $\big|\operatorname{Inf}\{v\}\big| = 1$ for all $v \in V$. If $\mathrm{d}^{-}(v) = 1$ and $\mathrm{D}^{-}(v) \neq \big\{(v, v)\big\}$, then there is a vertex $u \neq v$ such that $(u, v) \in E$ and $v \in \operatorname{Inf}\{u\}$. This implies $\big|\operatorname{Inf}\{u\}\big| > 1$.

Conversely, suppose that the restriction on D^{-} from the proposition conditions is hold and $\big|\operatorname{Inf}\{v\}\big| > 1$ for some v. If $u \in \operatorname{Inf}\{v\} \setminus \{v\}$, then by definition of inflation $\mathrm{d}^{-}(v) = 1$; in addition, the arc from $\mathrm{D}^{-}(v)$ can not be a loop.

Proposition 2.3 is proved.

Now assume that $G = (V, E)$ is finite and its decomposition consists of only one component, i.e., digraph is an interval. Let x be a heading of this interval (in general, not the only one), i.e., $G = \operatorname{Inf}^{\infty}\{x\} = \operatorname{Inf}^{n}\{x\}$ for some $n > 0$.

Definition 2.2. *A finite digraph* $\mathbf{H} = (W, F)$ *with a partition* $W = \bigsqcup_{i=1}^{n} W_i$, $n \in \mathbb{N}$, *is called a **jet** if it satisfies the following conditions*:

(i) *if* $i \leq j$, *then* $(W_j \times W_i) \cap F = \varnothing$;

(ii) *for each* $j \geq 2$ *and every vertex* $x \in W_j$ *there exist* $y_i \in W_i$, $1 \leq i < j$, *forming a directed path*

$$y_1 \to y_2 \to \ldots \to y_{j-1} \to x.$$

Proposition 2.4. *Let* $\mathbf{H} = (W, F)$ *be a jet*, x *be an element not contained in* W, *and* $V = W \cup \{x\}$. *Choose an arbitrary subset*

$$C \subset \bigsqcup_{i>1} \big(W_i \times \{x\}\big) \cup \big\{(x, x)\big\}$$

and put $E = F \cup C \cup \big(\{x\} \times W_1\big)$. *Then in the digraph* $G = (V, E)$ *the subset* V *is an interval with a heading* x.

Proof. Denote $W_0 = \{x\}$, $Z_j = \bigsqcup_{i=0}^{j} W_i$ and verify that $Z_j = \operatorname{Inf} Z_{j-1}$. The inclusion $\supseteq$ is evident. Conversely, suppose that $y \in W_j$, $j > 0$. By condition (ii) of Definition 2.2 $\mathrm{D}^-(y) \neq \varnothing$. By condition (i) $z \in \mathrm{D}^-(y)$ implies $z \in W_k$ for some $k < j$. It means that $\mathrm{D}^-(y) \subset Z_{j-1}$.

It is easy to see that $\operatorname{Inf}\{x\} = W_1$; and thus $V = \operatorname{Inf}^{n}\{x\} = \operatorname{Inf}^{\infty}\{x\}$.

Proposition 2.4 is proved.

The converse is true. Moreover:

Proposition 2.5. *Let* $\big\{G_j = (V_j, E_j) \mid 1 \leq j \leq N\big\}$ *be an interval decomposition of a finite digraph* $G = (V, E)$ *and* $V_j = \operatorname{Inf}^{\infty}\{x_j\}$. *Then every subgraph* $\big(V_j \setminus \{x_j\}, E_j|_{V_j \setminus \{x_j\}}\big)$ *with the partition*

$$V_j \setminus \{x_j\} = \bigsqcup_{i_j=1}^{\infty} \big(\operatorname{Inf}^{i_j}\{x_j\} \setminus \operatorname{Inf}^{i_j-1}\{x_j\}\big)$$

is a jet.

Proof. Consider an interval V_k and put $W_i = \operatorname{Inf}^{i}\{x_k\} \setminus \operatorname{Inf}^{i-1}\{x_k\}$. If $(u, v) \in (W_j \times W_i) \cap E$ for $1 \leq i \leq j$ then $v \in \mathrm{D}^+(u) \not\subset \operatorname{Inf}^{i-1}\{x_k\}$. Hence for $V_j \setminus \{x_j\}$ condition (i) of Definition 2.2 is hold. Condition (ii) is obvious.

3. Undirected graphs. In this section we assume that $G = (V, E)$ is a finite undirected connected graph without loops. We will use the notations $\mathrm{D}(v)$ and $\mathrm{d}(v)$ instead of $\mathrm{D}^{\pm}(v)$ and $\mathrm{d}^{\pm}(v)$ respectively.

In the undirected case the description of a hyperinflation is simplified:

Proposition 3.1. $\operatorname{Inf}^{\infty} U = \operatorname{Inf} U$ *for every subset* $U \subset V$.

Proof. Let $x \in \operatorname{Inf}^{\infty} U \setminus U$. Then $x \in \operatorname{Inf}^{n} U$ for some $n \geq 1$ and $(y, x) \in E$ for some $y \in \operatorname{Inf}^{n-1} U$. But this is impossible for $n > 1$, otherwise, $x \in \mathrm{D}^+(y) = \mathrm{D}^-(y) \subset \operatorname{Inf}^{n-2} U$. Therefore, $n = 1$ and $x \in \operatorname{Inf} U$.

Proposition 3.1 is proved.

Thus in what follows we may talk about the inflation rather than the hyperinflation and use the appropriate notations.

Consider some variants of decompositions. We will write them in the form of

$$V = \left(\bigsqcup_i \operatorname{Inf} V_i\right) \sqcup U, \tag{2}$$

where $G[V_i]$ are graphs of some (fixed) class and U is a subset of singleton components.

First, in the process described after Theorem 1.2 we can choose nonsingleton connected subsets as V_j until this is possible. Let components $V_1, \ldots, V_k$ be chosen in such way and in $U = V \setminus \bigsqcup_{i \le k} \operatorname{Inf} V_i$ there are no any connected components other than vertices. It means that U is completely disconnected. Moreover $\mathrm{D}(v) \subset V_i$ for every $v \in \operatorname{Inf} V_i \setminus V_i$. This proves the following proposition.

Proposition 3.2. *Each connected graph $G = (V, E)$ with a finite set of vertices has the decomposition of form* (2) *where V_i are nonsingleton connected subsets and U is completely disconnected* (*possibly empty*). *Moreover $\left(\bigsqcup_i \operatorname{Inf} V_i \setminus V_i\right) \cup U$ is completely disconnected subset.*

Another type of a decomposition is obtained if we choose two-element connected subsets, i.e., arcs, as $V_1, \ldots, V_k$. Clearly, the proof will not change, and we get the following corollary.

***Corollary* 3.1.** *Each connected graph $G = (V, E)$ with a finite set of vertices has the decomposition of form* (2) *where V_i are arcs*[1], *U is completely disconnected* (*possibly empty*) *subset as well as $\left(\bigsqcup_i \operatorname{Inf} V_i \setminus V_i\right) \cup U$.*

Recall that a **matching** of a graph is a set of pairwise nonadjacent edges, i.e., the arcs that have no common vertices. A matching is said to be **maximal**, if it is not contained in any other matching of the graph, and is said to be the **greatest**, if it contains the maximum number of arcs.

Decompositions of Corollary 3.1 are characterized in terms of matchings:

Theorem 3.1. *For a finite connected graph $G = (V, E)$ with the decomposition of form* (2) *satisfying the conditions of Corollary* 3.1 *the arcs $V_1, V_2, \ldots$ form the maximal matching. Conversely, if $\{V_1, V_2, \ldots\}$ is a maximal matching, then expression* (2) *is a decomposition satisfying the conditions of Corollary* 3.1.

Proof. By construction different V_i and V_j have no common vertices, therefore $\{V_1, V_2, \ldots\}$ is a matching. The complete disconnectedness of

$$\left(\bigsqcup_i \operatorname{Inf} V_i \setminus V_i\right) \cup U$$

implies its maximality.

Conversely, let $\{V_1, V_2, \ldots\}$ be a maximal matching and $V_i = (x_i, y_i)$. Suppose that $\operatorname{Inf} V_i \cap \operatorname{Inf} V_j \neq \varnothing$, $i \neq j$. According to Corollary 1.2 we can assume that $\operatorname{Inf} V_i \cap V_j \neq \varnothing$. Since $V_i \cap V_j = \varnothing$, either x_j or y_j is contained in $\operatorname{Inf} V_i \setminus V_i$.

If, for example, $x_j \in \operatorname{Inf} V_i \setminus V_i$, then $y_j \in \mathrm{D}^-(x_j) \subset V_i$; that is impossible. Similarly $y_j \in \operatorname{Inf} V_i \setminus V_i$ implies $x_j \in V_i$. Hence $\operatorname{Inf} V_i \cap \operatorname{Inf} V_j = \varnothing$ for all $i \neq j$.

Theorem 3.1 is proved.

We deal with another variant of a decomposition in the next section.

[1]We identify here an arc and the connected set of its vertices.

4. Forbidden subgraphs. In this section we apply a decomposition to the well-known forbidden subgraphs problem. This direction began with Turán's work [5] about the number of edges in the graph that does not contain any clique of given order. A good overview is given in [3]. Among the recent articles we mention also [6].

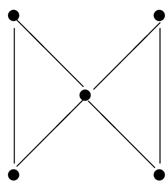
In general, the problem statement is as follows:

Let H be a fixed finite graph (**forbidden graph**). Find the least upper bound $\mathrm{ex}(p, H)$ for the number of arcs of finite graphs with p vertices, not containing H as a subgraph (such graphs are called H-**free**).

We use the number-theoretic functions "floor" $\lfloor x \rfloor$, "ceiling" $\lceil x \rceil$, and fractional part $\{x\}$. Recall that

$$\lfloor x \rfloor = x - \{x\}, \qquad \lceil x \rceil = x + \{-x\}.$$

For K_3 (the complete graph of order 3) $\mathrm{ex}(p, K_3) = \left\lfloor \dfrac{p^2}{4} \right\rfloor$ [3] (Theorem I.2). We obtain a similar evaluation for the graph H of the form



Hereinafter H denotes just this graph. Following [4] we call it a "bowtie".

A sequence of different vertices $U = \{v_1, \ldots, v_n\} \subset V$ of $G(V, E)$ is said to be a **path** if $(v_i, v_{i+1}) \in E$ for all $1 \leq i < n$.

If U_1, U_2 are two disjoint subsets of vertices, then $\mathrm{d}(U_1, U_2)$ denotes the number of arcs connecting vertices of U_1 with those of U_2.

The **volume** of $G = (V, E)$ is a pair $\mathrm{vol}\, G = (p, q)$ where p is a number of vertices and q is a number of arcs in G.

Our main result in this section is the following theorem.

Theorem 4.1. $\mathrm{ex}(p, H) = \left\lfloor \dfrac{p^2}{4} \right\rfloor + 1$ *for* $p > 4$.

To build an H-free graph with exactly $\left\lfloor \dfrac{p^2}{4} \right\rfloor + 1$ arcs, it is sufficient to consider $K_{\lfloor \frac{p}{2} \rfloor, \lceil \frac{p}{2} \rceil}$ $\Big($the complete bipartite graph with partite sets containing $\left\lfloor \dfrac{p}{2} \right\rfloor$ and $\left\lceil \dfrac{p}{2} \right\rceil$ vertices$\Big)$ and to draw one more arc in one of its partite sets.

The example of the graph K_4 shows that for $p = 4$ the statement of Theorem 4.1 is violated.

The remaining part of the article will be devoted to the proof of the proposition which implies, taking into account the facts mentioned above, the theorem.

Proposition 4.1. *Let G be an H-free graph which is not isomorphic to K_4 and* $\mathrm{vol}\, G = (p, q)$. *Then* $q \leq \dfrac{p^2}{4} + 1$.

First, make sure that it is enough to prove this statement for connected graphs.

Lemma 4.1. *Let G be a disconnected graph*, $\mathrm{vol}\, G = (p, q)$ *and* G_j, $j = 1, \ldots, n$, $n \geq 2$, *be all of its connected components with the volumes* $\mathrm{vol}\, G_j = (p_j, q_j)$. *If* $q_j \leq \dfrac{p_j^2}{4} + 1$ *for all* j, *then* $q \leq \dfrac{p^2}{4} + 1$.

Proof. It is clear that we can consider only the case $n = 2$. If $p_1 = p_2 = 1$ then $q_1 = q_2 = 0$, and the lemma is true. Otherwise, $p_1 p_2 \geq 2$ implies $\dfrac{(p_1+p_2)^2}{4} + 1 \geq \left(\dfrac{p_1^2}{4} + 1\right) + \left(\dfrac{p_2^2}{4} + 1\right)$.

Lemma 4.1 is proved.

In what follows we assume that G is a finite and connected graph and do not indicate that specially.

Prove some auxiliary statements.

Lemma 4.2. *Let $G = (V, E)$ be an H-free graph, $U = \{v_1, \ldots, v_l\}$ be a path in G and $x \in V \setminus U$. Then $\mathrm{d}(x, U) \leq \left\lceil \dfrac{l}{2} \right\rceil + 1$. If the equality holds, then:*

(i) *there exist vertices v_j, v_{j+1} adjacent to x;*

(ii) *either v_1 or v_l is adjacent to x, if l is even;*

(iii) *both v_1 and v_l are adjacent to x, if l is odd.*

Proof. Let l be even, $l = 2m$. Suppose that $\mathrm{d}(x, U) \geq \left\lceil \dfrac{l}{2} \right\rceil + 2 = m + 2$. By the pigeonhole principle among m pairs $(v_1, v_2), (v_3, v_4), \ldots, (v_{2m-1}, v_{2m})$ there exist (v_i, v_{i+1}), (v_j, v_{j+1}) such that four arcs outgoing from x end in them. At the same time $i \neq j+1, j \neq i+1$, hence x, v_i, v_{i+1}, v_j, v_{j+1} form a subgraph isomorphic to H.

If there are no v_j, v_{j+1} adjacent to x, then $\mathrm{d}(x, U) \leq m < \left\lceil \dfrac{l}{2} \right\rceil + 1$. This implies (i) for even l.

Since for the path $U' = \{v_2, \ldots, v_{l-1}\}$ consisting of $l - 2$ vertices $\mathrm{d}(x, U') \leq \left\lceil \dfrac{l-2}{2} \right\rceil + 1 = \left\lceil \dfrac{l}{2} \right\rceil$, it follows from $\mathrm{d}(x, U) = \left\lceil \dfrac{l}{2} \right\rceil + 1$ that either v_1 or v_l is adjacent to x; therefore (ii) is hold.

Let $l = 2m+1$. As it was proved above, for $U' = \{v_1, \ldots, v_{2m}\}$ the inequality $\mathrm{d}(x, U') \leq m+1$ holds. Hence $\mathrm{d}(x, U) \leq m + 2 = \left\lceil \dfrac{l}{2} \right\rceil + 1$; and the equality is possible only in the case when the vertex u_{2m+1} is adjacent to x and there exist vertices v_j, v_{j+1} adjacent to x. To complete the proof of (iii) it suffices to consider the path $\{v_2, \ldots, v_{2m+1}\}$ and to deduce that v_1 is adjacent to x.

Lemma 4.2 is proved.

A path $\{v_1, \ldots, v_l\}$ in the graph $G = (V, E)$ is called **premaximal** if there exists a vertex $v_{l+1} \in V \setminus U$ such that the path $\{v_1, \ldots, v_l, v_{l+1}\}$ is maximal, i.e., has the maximum possible length.

Lemma 4.3. *Let $G = (V, E)$ be not completely disconnected and U be a premaximal path. Then $\operatorname{Inf} U \neq U$.*

Proof. Let $U = \{v_1, \ldots, v_l\}$ and $U' = \{v_1, \ldots, v_{l+1}\}$ be a maximal path in G. Then $\mathrm{D}(v_{l+1}) \subset U$. Therefore $v_{l+1} \in \operatorname{Inf} U \neq U$.

Lemma 4.4. *Let $U = \{v_1, \ldots, v_l\}$ be a premaximal path in the H-free graph $G = (V, E)$. If $x \in V \setminus U$ and $\mathrm{d}(x, U) = \left\lceil \dfrac{l}{2} \right\rceil + 1$, then for every vertex $y \in V \setminus U$ such that $y \neq x$, inequality $\mathrm{d}(y, U) \leq \left\lceil \dfrac{l}{2} \right\rceil - 1$ holds.*

Proof. Assume the contrary, let $\mathrm{d}(y, U) \geq \left\lceil \dfrac{l}{2} \right\rceil$. Note that by Lemma 4.2 there is a pair of vertices v_j, v_{j+1} adjacent to x. Then y can not be adjacent to v_1, otherwise the path $\{y, v_1, \ldots, v_j, x,$

$v_{j+1},\dots,v_l\}$ is longer than maximal. Similarly y is not adjacent to v_l. Hence for the path $U' = \{v_2,\dots,v_{l-1}\}$ the inequality $\mathrm{d}(y,U') \geq \left\lceil \frac{l}{2} \right\rceil = \left\lceil \frac{l-2}{2} \right\rceil + 1$ holds. Therefore by Lemma 4.2 we can find vertices v_i, v_{i+1} adjacent to y. Without loss of generality, we can also suppose, in view of statements (ii), (iii) of Lemma 4.2, that x is adjacent to v_1. Then the path $\{x, v_1,\dots,v_i, y, v_{i+1},\dots v_l\}$ is longer than maximal path, a contradiction.

Lemma 4.4 is proved.

***Corollary* 4.1.** *Let $U = \{v_1,\dots,v_l\}$ be a premaximal path in the H-free graph $G = (V,E)$ and $\left|\mathrm{Inf}\, U\right| = p$. If $p - l \geq 2$, then $\mathrm{d}(\mathrm{Inf}\, U \setminus U, U) \leq (p-l)\left\lceil \frac{l}{2} \right\rceil$.*

Let $U = \{v_1,\dots,v_l\}$ be a path. We put $U^{(j)} = U \setminus \{v_j\}$.

Lemma 4.5. *Let $l \geq 3$ and $U = \{v_1,\dots,v_l\}$ be a path in the H- and K_4-free graph $G = (V,E)$. Let $x \in V \setminus U$ and $\mathrm{d}(x,U) = \left\lceil \frac{l}{2} \right\rceil + 1$. If the number of arcs of the subgraph $G[U]$ does not exceed $\frac{l^2}{4} + 1$, then there exists a vertex v_j for which $\mathrm{d}\big(v_j, U^{(j)} \cup \{x\}\big) \leq \left\lceil \frac{l}{2} \right\rceil$.*

Proof. Assume the contrary: for every vertex v_j the inequality

$$\mathrm{d}\big(v_j, U^{(j)} \cup \{x\}\big) \geq \left\lceil \frac{l}{2} \right\rceil + 1$$

holds. By hypothesis $G[U \cup \{x\}]$ contains no more than $\frac{l^2}{4} + \left\lceil \frac{l}{2} \right\rceil + 2$ arcs. The assumption implies:

$$\frac{1}{2}(l+1)\left(\left\lceil \frac{l}{2} \right\rceil + 1\right) \leq \frac{l^2}{4} + \left\lceil \frac{l}{2} \right\rceil + 2.$$

Hence for even l we have $l \leq 6$ and for odd l we have $l \leq 3$. Therefore it is sufficient to consider the cases $l = 3, 4, 6$.

Let A be a set of vertices of U which are adjacent to x, and $\overline{A} = U \setminus A$.

If $l = 3$, then $\mathrm{d}(x,U) = \left\lceil \frac{3}{2} \right\rceil + 1 = 3$; hence $A = U = \{v_1, v_2, v_3\}$. Then the assumption implies that the graph G is isomorphic to K_4; this contradicts the condition.

Let $l = 4$. Without loss of generality, we can assume that $\overline{A} = \{v_3\}$ or $\overline{A} = \{v_4\}$. Then $v_1, v_2 \in A$. Hence v_2 and v_4 are not adjacent, otherwise, G contains a “bowtie”. By assumption $\mathrm{d}\big(v_4, U^{(4)} \cup \{x\}\big) \geq 3$, i.e., v_4 is adjacent to v_1 and x, and $\overline{A} = \{v_3\}$. Then $(v_1, v_3) \notin E$, otherwise the vertices of $U \cup \{x\}$ form a “bowtie”. Therefore, $\mathrm{d}\big(v_3, U^{(3)} \cup \{x\}\big) < 3$ contrary to assumption.

Let $l = 6$. Then $\mid \overline{A} \mid = 2$. Without loss of generality, by Lemma 4.2 we can suppose that x is adjacent to v_1.

First, assume that $v_6 \notin A$. Since G is H-free, it follows that either $A = \{v_1, v_2, v_3, v_5\}$ or $A = \{v_1, v_3, v_4, v_5\}$. According to the assumption $\mathrm{d}\big(v_6, U^{(6)}\big) = \mathrm{d}\big(v_6, U^{(6)} \cup \{x\}\big) = 4$. Therefore, as well as for x, there are two variants: the set of vertices adjacent to v_6 equals either $\{v_1, v_2, v_3, v_5\}$ or $\{v_1, v_3, v_4, v_5\}$. Checking straightforwardly four cases, we get a contradiction. So $v_6 \in A$.

Note that the cases $\overline{A} = \{v_3, v_4\}$, $\overline{A} = \{v_3, v_6\}$ and $\overline{A} = \{v_5, v_6\}$ are impossible, since G is H-free. We will obtain the contradiction in every of the remaining variants:

Let $\overline{A} = \{v_2, v_3\}$. Note that $(v_2, v_4), (v_3, v_5) \notin E$, since G is H-free. Then assumption implies that v_2 and v_3 are adjacent to v_6, hence the vertices $\{v_2, v_3, v_6, v_5, x\}$ form a “bowtie”. The case $\overline{A} = \{v_4, v_5\}$ is similar.

Let $\overline{A} = \{v_2, v_4\}$. Note that $(v_3, v_5), (v_1, v_3) \notin E$. The assumption implies that v_3 is adjacent to v_6, hence, $(v_4, v_6), (v_4, v_2) \notin E$ and $\mathrm{d}(v_4, U^{(4)} \cup \{x\}) < 4$ contrary to the assumption. The case $\overline{A} = \{v_3, v_5\}$ is similar.

Let $\overline{A} = \{v_2, v_5\}$. Then $(v_4, v_6), (v_1, v_6) \notin E$ and the assumption implies that v_2 and v_3 are adjacent to v_6. Therefore, the vertices $\{v_2, v_3, v_6, v_4, x\}$ form a "bowtie".

Lemma 4.5 is proved.

Proposition 4.2. *Let G be a connected graph,* $\operatorname{vol} G = (p, q)$, *and the length of the maximal path in G does not exceed* 2. *Then* $q \le \dfrac{p^2}{4} + 1$.

Proof. Note that the graph G satisfying the condition is isomorphic to K_p for $p \le 3$ or to $K_{1,p-1}$. The inequality can be proved by immediate check.

Proposition 4.2 is proved.

Now we are ready to prove Proposition 4.1.

Proof. First, let $G = (V, E)$ be a H- and K_4-free graph, $\operatorname{vol} G = (p, q)$, and the length of the maximal path in G is greater than 2. Construct a decomposition choosing the pathes without self-intersections as V_i and taking a premaximal path as the first component V_1.

Let $\operatorname{vol} G[V_i] = (l_i, m_i)$, $\operatorname{vol} G[\operatorname{Inf} V_i] = (p_i, q_i)$, $i = 1, \ldots, n$. Then $p_1 > l_1$ by Lemma 4.3 and $l_1 \ge 3$ by assumption.

Use an induction on p.

We consider separately the case $n = 1$. We omit the indices in the notations, thus, $p = p_1$, $q = q_1$, $l = l_1$, $m = m_1$.

Let $p > l + 1$. By the induction assumption $m \le \left\lfloor \dfrac{l^2}{4} \right\rfloor + 1$, and by Corollary 4.1 $q - m \le$ $\le (p - l) \left\lceil \dfrac{l}{2} \right\rceil$. Then

$$\frac{p^2}{4} + 1 - q = \frac{p^2}{4} + 1 - m - (q - m) \ge \frac{p^2}{4} + 1 - \left\lfloor \frac{l^2}{4} \right\rfloor - 1 - (p - l) \left\lceil \frac{l}{2} \right\rceil =$$

$$= \frac{p^2}{4} - \frac{l^2}{4} + \left\{ \frac{l^2}{4} \right\} - (p - l)\frac{l}{2} - (p - l) \left\{ -\frac{l}{2} \right\} =$$

$$= \frac{(p - l)^2}{4} + \left\{ \frac{l^2}{4} \right\} - (p - l) \left\{ \frac{l}{2} \right\} = \left(\frac{p - l}{2} - \left\{ \frac{l}{2} \right\} \right)^2 \ge 0. \tag{3}$$

Let $p = l + 1$. Note that there is a vertex $x \in \operatorname{Inf} V_1$ such that $\mathrm{d}\big(x, \operatorname{Inf} V_1 \setminus \{x\}\big) \le \left\lceil \dfrac{l}{2} \right\rceil$. Indeed, $q - m \le \left\lceil \dfrac{l}{2} \right\rceil + 1$ by Lemma 4.2. Therefore the only vertex of the set $\operatorname{Inf} V_1 \setminus V_1$ can be taken as x or, by Lemma 4.5, x can be chosen in V_1.

Let $\operatorname{vol} G\big[\operatorname{Inf} V_1 \setminus \{x\}\big] = (l, s)$. Then $q - s \le \left\lceil \dfrac{l}{2} \right\rceil$. Again by the induction assumption $s \le \left\lfloor \dfrac{l^2}{4} \right\rfloor + 1$. So we have

$$\frac{p^2}{4} + 1 - q = \frac{(l + 1)^2}{4} + 1 - s - (q - s) \ge \frac{(l + 1)^2}{4} + 1 - \left\lfloor \frac{l^2}{4} \right\rfloor - 1 - \left\lceil \frac{l}{2} \right\rceil. \tag{4}$$

Note that expression (4) is obtained from (3) by substitution $p = l + 1$, hence, it is nonnegative.

Thus for $n = 1$ the statement is proved.

Let $n \geq 2$. Put

$$p' = p - p_1 = \sum_{i=2}^{n} p_i, \qquad q' = q - q_1 = \sum_{i=2}^{n} q_i + \sum_{1 \leq i < j \leq n} \mathrm{d}(V_i, V_j).$$

Applying the induction assumption to the subgraph $G\Big[V_1 \sqcup \bigsqcup_{j=2}^{n} \operatorname{Inf} V_j\Big]$, we have $m_1 + q' \leq \leq \left\lfloor \dfrac{(l_1 + p')^2}{4} \right\rfloor + 1$. Moreover $q_1 - m_1 \leq (p_1 - l_1)\left(\left\lceil \dfrac{l_1}{2} \right\rceil + 1\right)$ by Lemma 4.2. Then

$$\frac{p^2}{4} + 1 - q = \frac{(p_1 + p')^2}{4} + 1 - (m_1 + q') - (q_1 - m_1) \geq$$

$$\geq \frac{(p_1 + p')^2}{4} + 1 - \frac{(l_1 + p')^2}{4} + \left\{ \frac{(l_1 + p')^2}{4} \right\} - 1 - (q_1 - m_1) \geq$$

$$\geq \frac{p_1^2}{4} + \frac{(p_1 - l_1)p'}{2} - \frac{l_1^2}{4} + \left\{ \frac{(l_1 + p')^2}{4} \right\} - (p_1 - l_1)\left(\frac{l_1}{2} + \left\{ -\frac{l_1}{2} \right\} + 1 \right) =$$

$$= \frac{(p_1 - l_1)^2}{4} + (p_1 - l_1)\left(\frac{p'}{2} - \left\{ -\frac{l_1}{2} \right\} - 1 \right) + \left\{ \frac{(l_1 + p')^2}{4} \right\}. \tag{5}$$

Since $\left\{ -\dfrac{l_1}{2} \right\} + 1 \leq \dfrac{3}{2}$ it follows that (5) is nonnegative for $p' \geq 3$.

Note that $p' \neq 1$. Otherwise $n = 2$ and the only vertex of $\operatorname{Inf} V_2$ is contained, in view of connectedness of G, in $\operatorname{Inf} V_1$; this is impossible. If $p' = 2$, then (5) becomes

$$\frac{(p_1 - l_1)^2}{4} - (p_1 - l_1)\left\{ -\frac{l_1}{2} \right\} + \left\{ \frac{(l_1 + 2)^2}{4} \right\} =$$

$$= \frac{(p_1 - l_1)^2}{4} - (p_1 - l_1)\left\{ \frac{l_1}{2} \right\} + \left\{ \frac{l_1^2}{4} \right\} = \left(\frac{p_1 - l_1}{2} - \left\{ \frac{l_1}{2} \right\} \right)^2 \geq 0.$$

Thus we have completed the proof for K_4-free graphs.

Now let G with $\operatorname{vol} G = (p, q)$ be an arbitrary H-free graph, which is not isomorphic to K_4.

Use the induction on the number of subgraphs of G isomorphic to K_4. If there are no such subgraphs, then the statement is already proved; therefore, the basis step is verified.

Let F be a subgraph of G isomorphic to K_4 and $U = \{v_1, v_2, v_3, v_4\}$ be the set of its vertices. Let $\overline{F} = G[V \setminus U]$ and $\operatorname{vol} \overline{F} = (l, m)$. Note that $\mathrm{d}(x, U) \leq 1$ for every vertex $x \in V \setminus U$, otherwise G is not H-free.

If $l = 1$, then $p = 5$, $q = 7$; therefore $q \leq \dfrac{p^2}{4} + 1$.

If $\overline{F}$ is isomorphic to K_4, then $\mathrm{d}(U, V \setminus U) \leq 4$, because each of the vertices of F is adjacent to not more than one vertex of $\overline{F}$. In this case $p = 8$, $q \leq 16 < \left\lfloor \dfrac{64}{4} \right\rfloor + 1$.

If $\overline{F}$ is not isomorphic to K_4 and $l \geq 2$, then applying the induction assumption to $\overline{F}$, we have $m \leq \dfrac{l^2}{4} + 1$. Since $p = l + 4$ and $q \leq m + l + 6$, it follows that

$$\frac{p^2}{4} + 1 - q \geq \frac{(l+4)^2}{4} + 1 - m - l - 6 = \left(\frac{l^2}{4} + 1 - m\right) + (l - 2) \geq 0,$$

as required.

Proposition 4.2 is proved.

1. *Dokuchaev M., Novikov B., Zholtkevych G.* Partial actions and automata // Algebra and Discrete Math. – 2011. – **11**, № 2. – P. 51 – 63.
2. *Aho A. V., Ullman J. D.* The theory of parsing, translation, and compiling. Vol. 2. Compiling. – Prentice-Hall, 1973.
3. *Bollobas B.* Modern graph theory // Grad. Texts Math. – 1998. – **184**.
4. *West D. B.* Introduction to graph theory. – Pearson Education, 2001.
5. *Turán P.* On an extremal problem in graph theory (in Hungarian) // Mat. Fiz. Lapok. – 1941. – **46**. – P. 436 – 452.
6. *Eschen E. M., Hoang Ch. T., Spinrad J. P., Sritharan R.* On graphs without a C_4 or a diamond // Discrete Appl. Math. – 2011. – **159**, № 7. – P. 581 – 587.

Received 11.06.13,
after revision — 06.10.13