# PROBLÈMES COMBINATOIRES
# ET
# THÉORIE DES GRAPHES

*ORSAY*

*9-13 Juillet 1976*

*EXTRAIT*

# INTERSECTION THEOREMS FOR GRAPHS

Miklós SIMONOVITS and Vera T. SÓS

Eötvös Loránd University, Budapest

**Résumé.** — Dans cet article, nous considérons des problèmes analogues à un problème d'Erdös et de De Bruijn [1]. Soient $G_1, \ldots, G_t$ une famille de graphes simples construits sur le même ensemble de $n$ sommets. Nous désignons par $G_i \cap G_j$ le graphe dont les arêtes sont les arêtes communes à $G_i$ et $G_j$ et dont les sommets sont ceux de ces arêtes. Nous montrons que si pour tout $i, j$ ($1 \leq i < j \leq t$) $G_i \cap G_j$ est une chaîne ou est vide alors $t = O(n^5)$; si $G_i \cap G_j$ est une chaîne non vide $t = O(n^4)$; si $G_i \cap G_j$ est un cycle ou est vide $t = O(n^4)$.

**Introduction.** — A well known result of N. G. de Bruijn and P. Erdös asserts that if $A_i$ ($i = 1, \ldots, t$) are subsets of an $n$ element set $V$, then the conditions $|A_i \cap A_j| = 1$ for every $i \neq j$ implies that $t \leq n$. (Here $|E|$ denotes the cardinality of $E$.) There is an extensive literature on problems like this and we are interested in the following generalization of the problem above.

**The graph problem.** — Given a family $\mathcal{A}$ of graphs and for a given $n$ element set $V$ and graphs $G_i$ on the vertex set $V$ we know that $G_i \cap G_j \in \mathcal{A}$, whenever $1 \leq i < j \leq t$. Here the intersection $G_i \cap G_j$ is the graph whose edges are the edges belonging to both $G_i$ and $G_j$ and the vertices are the endvertices of these edges. (Observe that by this definition of $G_i \cap G_j$ it has no isolated vertices !) We would like to know, maximally how large $t$ can be (under the condition $G_i \cap G_j \in \mathcal{A}$) ?

*Remark 1.* — If $\mathcal{A}$ is the set of the graphs of $k$ edges, we get back to the original set intersection problem.

For some families $\mathcal{A}$ we shall publish our results in [2]. Here we restrict our investigations to the following three cases :

*Case 1.* — $\mathcal{A}_0$ is the family of path with vertex set in $V$, where the empty graph is also considered to be a path : $\varnothing \in \mathcal{A}_0$.

*Case 2.* — $\mathcal{A}_1$ is the family of paths with vertex set in $V$, but the empty path is excluded : $\varnothing \notin \mathcal{A}_1$.

*Case 3.* — $\mathcal{A}_2$ is the family of cycles on $V$, where the empty graph is also in $\mathcal{A}_2$ : $\varnothing \in \mathcal{A}_2$.

*Remark 2.* — The interesting feature of the cases mentioned above is, that, as we shall see from the following theorems, the bounds are polynomial in $n$. If we assume that the intersections $G_i \cap G_j$ are all trees on $V$, then $t = 2^{n-1}$ can still occur : let $H$ be the graph on $V = (1, \ldots, n)$ with the edges $(1, 2)$, $(1, 3), \ldots, (1, n)$ and let $G_i$ be an arbitrary spanned

subgraph of $H$ containing 1. Clearly, the intersection $G_i \cap G_j$ is always a tree, moreover, a star, while $t = 2^{n-1}$. Even if we assume that the intersection $G_i \cap G_j$ is a star with at least 2 vertices, a slight modification of the above construction shows that $t$ can be exponentially large.

Our main results are formulated in the following theorems :

**Theorem 1.** — *If $G_i$ ($i = 1, \ldots, t$) is a system of graphs with the common vertex set $V$, $|V| = n$ and $G_i \cap G_j \in \mathcal{A}_0$ whenever $i \neq j$, then $t = O(n^5)$.*

**Theorem 2.** — *If $G_i$ ($i = 1, \ldots, t$) is a system of graphs with the common vertex set $V$, $|V| = n$ and $G_i \cap G_j \in \mathcal{A}_1$ whenever $i \neq j$, then*

$$t = O(n^4) .$$

**Theorem 3.** — *If $G_i$ ($i = 1, 2, \ldots, t$) is a system of graphs with the common vertex set $V$, $|V| = n$ and $G_i \cap G_j$ is always a (possibly empty) cycle unless $i = j$ ($G_i \cap G_j \in \mathcal{A}_2$), then*

$$t = O(n^4) .$$

*Remark 3.* — The exponents in the above theorems are sharp, as one can see from the following constructions :

*Construction 1.* — Let $C_1, \ldots, C_5$ be a partition of $n$ vertices into five classes, each of which contains at least $[n/5]$ vertices. Let the system $G_i$ be the system of graphs on $n$ vertices, each of which contains a cycle $(x_1 x_2 x_3 x_4 x_5)$, $x_i \in C_i$, and isolated vertices. Here we have $\geq [n/5]^5$ graphs any two of which intersect in a path. Hence theorem 1 is sharp. (This construction can easily improved by adding new graphs to the system, for example, graphs containing a cycle of type $(x_1 x_3 x_5 x_2 x_4)$ and isolated vertices.)

*Construction 2.* — Let $a$ and $b$ be two fixed vertices and $n - 2$ further vertices be partitioned into 4 classes

$C_1$, $C_2$, $C_3$, $C_4$ each of which has at least $\left\lceil \dfrac{n-2}{4} \right\rceil$ vertices. If we consider the system of graphs each of which contains one path $(x_1\ x_2\ abx_3\ x_4)$, $x_i \in C_i$, and isolated vertices, then this system contains

$$t \geqslant \left\lceil \frac{n-2}{4} \right\rceil^4$$

members and any two (different) member intersects in a nonempty path. Hence theorem 2 is sharp.

*Construction* 3. — Let $T_1$, $T_2$, ... be a system of triangles any two of which have at most one vertex in common. As it is well known, one can find (approximately) $\dfrac{1}{3}\dbinom{n}{2}$ such triangles on $n$ vertices and if the system $G_i$ consists of the graphs each containing exactly two of these triangles and the further isolated vertices, then any two (different) intersect in a cycle : theorem 3 is sharp.

**Proofs.** — *Proof of theorem 1.* — Let us consider an extremal system for the class $\mathcal{A}_0$, that is, a system $G_1$, ..., $G_t$ for which $G_i \cap G_j \in \mathcal{A}_0$ if $i \neq j$ and $t$ is the maximum possible. Without loss of generality we may assume that

$$(+) \quad \sum_j e(G_j) \text{ is the minimum possible}$$

(here $e(G_j)$ denotes the number of edges of $G_j$).

First we remark, that if $K(1, 3)$ is the graph on 4 vertices $x_1$, $x_2$, $x_3$, $x_4$ where $x_1$ is joined to $x_2$, $x_3$, $x_4$ but $x_2$, $x_3$, $x_4$ are independent, then any $K(1, 3)$ is contained in at most one $G_i$. Thus, if $L_i \subseteq G_i$ is isomorphic to $K(1, 3)$ for some $i$, we can replace $G_i$ by this $L_i$ : this $L_i$ was not contained in the old system unless $L_i = G_i$ and it intersects any $G_j$ in a path. By $(+)$ $G_i = L_i$. Hence *our system consists of graphs whose edges span a $K(1, 3)$ and of graphs in which the maximum valence $\leq 2$.* If a graph $G_i$ contains a cycle $C$, then we can replace $G_i$ by this $C$ and obtain a new extremal family : the number of members does not change, since this $C$ cannot occur in any other $G_j$ and $G_j$ can intersect only one connected component of $G_i$, hence either $G_i \cap G_j = C \cap G_j$ or $C \cap G_j = \varnothing$. Thus the new system is really an extremal one. By $(+)$ $G_i = C$. Hence our system consists of graphs whose edges span

either a $K(1, 3)$
or a cycle
or one or more paths .

Let $H$ consists of two independent edges. Again, if a $G_i$ contains $H$ and the two independent edges of $H$ belong to two different connected components of $G_i$, then we can replace $G_i$ by $H$ and the new system will again be an extremal system : $H$ does not belong to the old system and any $G_j$ intersects only one connected component of $G_i$, hence $H \cap G_j$ has at most one edge. By $(+)$ $H = G_i$.

Now we know that any $G_i$ is either a path or a cycle or two independent edges or a $K(1, 3)$, apart from its isolated vertices. First we give an upper bound on the number of paths, 4-cycles and two independent edges. Let $e$ and $f$ be two edges (not necessarily forming a $G_i$) and observe that at most one $G_i$ of the following property can belong to our system : (a) a 4-cycle in which $e$ and $f$ are not adjacent, (b) a path with the terminal edges $e$ and $f$, (c) $G_i$ having only two edges, namely $e$ and $f$.

Hence the number of graphs listed above is at most $\dbinom{\dbinom{n}{2}}{2}$ . The number of $K(1, 3)$'s is at most $4\dbinom{n}{4}$ .

There exist at most $\dbinom{n}{3}$ triangles, $\dbinom{n}{2}+1$ graphs of at most one edge. Therefore we need only a good bound for the number of cycles of at least 5 edges. This will be obtained in the following way.

We shall consider the pairs $(e, C)$, where $e$ is an edge of the cycle $C = C_i$ and order appropriately to it two adjacent edges of $C$, say $u$ and $v$, so that no other $G_j$ contains all these three edges, among the *cycles* of the extremal system. Since $C$ is not a triangle, at least one of $u$ and $v$ is not adjacent with $e$. Thus either they form a path of 3 edges or both $u$ and $v$ are independent from $e$ but adjacent to each other. Therefore the number of graphs $G_i$ for which $G_i$ is a cycle and for at least one edge $e$ of it we get a path $(e, u, v)$ of 3 edges is at most

$$12\dbinom{n}{4},$$

while the number of cycles $G_i$ for which $(e, u, v)$ is not a path for any edge $e$ of $G_i$ is at most

$$\frac{1}{5} \cdot 10 \cdot 3 \cdot \dbinom{n}{5} = 6\dbinom{n}{5},$$

since each of them is counted at least 5 times. This will complete the proof of theorem 1 as soon as the promised ordering of $(e, u, v)$ to $(e, C)$ is done.

In the following part we shall also use a slight modification of the definition $G_i \cap G_j$ : let $P_1$ and $P_2$ be two paths, then the *p-intersection* of $P_1$ and $P_2$ is their intersection defined earlier, if it is not empty, however, if $P_1$ and $P_2$ have just one endvertex in common, their *p-intersection* is by definition this common endvertex. The *p-intersection* will be denoted by $P_1 \stackrel{.}{\cap} P_2$ or by $\stackrel{.}{\underset{i}{\bigcap}} P_i$.

Given the pair $(e, C)$ the edges $u$ and $v$ will be defined as follows : Let $C_1, \dots, C_s$ be the other graphs in our system containing $e$. Put $I = : C - e$, $I_j = C - C_j$, $(j = 1, \dots, s)$. Now we show that the paths $I_j$ form a Helly system : the *p-intersection* of any two of them is nonempty, hence there exists a vertex contained by all of them. First of all, trivially, $I$ and

$I_j$ are paths. Further, let us assume indirectly, that $I_j \,\dot\cap\, I_k = \varnothing$. Thus $C$ is divided into 4 paths : $I_j$, $I_k$, a path $P$ containing $e$ and a path $Q$ separated from $e$ by $I_j$ and $I_k$. Let $Q = (a_1, \ldots, a_k)$ and $I_j$ be incident to $a_1$. Since the intersections $C \cap C_j$ and $C \cap C_k$ are paths, we obtain that $C_j \supseteq C - I_j$, $C_k \supseteq C - I_k$ and therefore $Q \subseteq C_j \cap C_k$. We show that $Q$ is a connected component of $C_j \cap C_k$. Indeed, at $a_1$ the intersection does not contain any edge but $(a_1, a_2)$ : $C_k$ does contain the edge of $I_j$ incident to $a_1$ but $C_j \cap I_j = \varnothing$, hence $C_j \cap C_k$ really cannot contain any edge but $(a_1, a_2)$ from the edges incident to $a_1$. Similarly, at $a_k$ it has degree 1, thus $Q_1$ is a connected component of it, not containing $e \in C_j \cap C_k$. Thus $C_j \cap C_k$ is not a path. This contradiction proves that $I_j \,\dot\cap\, I_k \ne \varnothing$. By the well known theorem of Helly (which is trivial now, i.e. for intervals), there exists a vertex $c$ contained by each $I_j$. If $u$ and $v$ are the edges incident to $c$ on $C$, then no other cycle $C'$ contains $e$, $u$, and $v$ : from $e \in C'$ follows that $C' = C_j$ for some $j$ and $c \in I_j = C - C_j \subseteq C$ yields that $C_j$ contains at most one of $u$ and $v$. Hence for different $(e, u, v)$ the graphs $G_i$ containing them are different. The reader can easily check that $c$ can be chosen from the vertices of $C$ not incident to $e$. Hence the edges $e$, $u$, $v$ are different (though this is not important here). Trivially, $u$ and $v$ are adjacent. This completes the proof.

*Proof of theorem 2.* — Let $G_1, \ldots, G_t$ be an extremal system. Each $K(1, 3)$ is contained in at most one $G_i$ and therefore the number of $G_i$'s containing $K(1, 3)$ is at most $4\binom{n}{4}$. For a given pair of edges at most one $G_i$ contains these edges in different connected components and at most one $G_i$ can be a path having these two edges as terminal edges. Moreover, for a given pair of edges the two different cases given above cannot occur at the same time. Thus the number of paths and disconnected graphs can be estimated by

$$\frac{1}{2}\binom{n}{2}^2, \text{ more precisely, by } \binom{\binom{n}{2}}{2}.$$ The triangles can

be estimated by $\binom{n}{2}$ and the number of graphs containing just one edge is estimated by $\binom{n}{2}$. The remaining graphs must be connected, with maximum valence 2 and are not paths : they are cycles (of at least 4 vertices). Let $C$ be a cycle. Since it has at most $n$ edges,

we may fix $n$ edges $e_1, \ldots, e_n$ so that any cycle $G_i$ contains at least one $e_j$. We have seen above, in the proof of theorem 1, that the number of cycles containing an edge $e$ is at most $3\binom{n-2}{3} + 4\binom{n-2}{2} \le \frac{n^3}{2}$. Thus the total number of cycles of at least 4 edges is at most $n^4/4$. This completes the proof.

*Remark 4.* — Here the graphs $G_i$ could not be replaced by the corresponding smaller graphs, e.g. by $K(1, 3)$, since that would turn some intersections into empty graph.

*Proof of theorem 3.* — Again, the number of $G_i$'s containing a fixed $K(1, 3)$ is at most one, hence the number of $G_i$'s containing some $K(1, 3)$ is at most $4\binom{n}{4}$. If $e$ is an arbitrary edge, it can be adjacent to a vertex of valence 1 in a graph $G_i$ containing $e$ for at most one value of $i$. Hence all but at most

$$4\binom{n}{4} + \binom{n}{2} + 1$$

graphs are regular graphs of degree 2. (Here 1 stands for the empty $G_i$.) Observe now, that if $C$ is a cycle in any $G_i$ and $C'$ is a cycle in any $G_j$ where $G_i$, $G_j$ are 2-regular, then either $C = C'$ or $CC' = \varnothing$. Hence we define $(C_1, \ldots, C_m)$ as the set of cycles contained by at least one 2-regular $G_i$ and observe that $m \le n^2/6$, since the cycles are edge-independent. If a $G_i$ has more than 2 cycles, we may replace it by two arbitrary cycles of it, the new system remains extremal. Thus we obtained that the number of 2-regular graphs is at most $n^4/36$. This completes the proof.

*Remark 5.* — *a)* The theorems above are not complete in the sence that we do not know the best coefficients of $n^5$, $n^4$ and $n^4$ respectively. Further, we do not know the structure of extremal system either. It would be nice to know for example, that in theorem 2 is it true, that an extremal system $G_1, \ldots, G_t$ always has an edge $e$ belonging to each $G_i$ ?

*b)* We conjecture that the extremal system of theorem 1 consists only of triangles, quadrangles, pentagons and the $4\binom{n}{4}$ $K(1, 3)$, and path of $\le 2$ edges.

*Added in proof.* — The case when $G_i \cap G_j$ is a cycle but the empty graph is excluded is considered in [2].

## References

[1] N. G. DE BRUIJN, P. ERDÖS, On a combinatorial problem, *Nederl. Akad. Wetensh. Proc.* **51** (1948) 1277-1279.
[2] M. SIMONOVITS, V. T. SÓS, Proc. Coll. Math. Soc. J. Bolyai on Combinatorica/theory, Keszthely (1976).