MR0186575 (32 #4034) 05.40
Erdős, Paul; Goodman, A. W.; Pósa, Lajos
The representation of a graph by set intersections.
*Canad. J. Math.* 18 1966 106–112

It is shown that for any graph $G$ with $n$ vertices, there exists a set $S$ with $\lceil n^2/4 \rceil$ elements and a
family $S_1, S_2, \ldots$ of $n$ subsets of $S$ (not necessarily distinct) such that two vertices $v_i$ and $v_j$ are
adjacent in $G$ if and only if the sets $S_i$ and $S_j$ intersect. Further, $\lceil n^2/4 \rceil$ is the smallest such number.
Reviewed by F. Harary

© Copyright American Mathematical Society 1966, 2008
THE REPRESENTATION OF A GRAPH BY SET INTERSECTIONS

PAUL ERDŐS, A. W. GOODMAN, AND LOUIS PÓSA

1. Introduction. Geometrically, a graph is a collection of points (or vertices) together with a set of edges (or curves) each of which joins two distinct vertices of the graph, and no two of which have points in common except possibly end points. Two given vertices of the graph may be joined by no edge or one edge, but may not be joined by more than one edge. From an abstract point of view, a graph $G$ is a collection of elements $\{v_1, v_2, \ldots \}$ called points or vertices, together with a second collection $C$ of certain pairs $(v_a, v_b)$ of distinct points of $G$. It is helpful to retain the geometric language, and refer to any pair in $C$ as an edge (or a curve) of $G$ that joins the points $v_a$ and $v_b$.

A family of sets $S_1, S_2, \ldots$ gives a graph in a natural way, if to each set $S_a$ we associate a point $x_a$ and agree that

\[(1) \quad x_a \text{ and } x_b \text{ are joined by an edge of } G \text{ if and only if } x \neq y \text{ and } S_a \cap S_b \neq \emptyset,\]

where $\emptyset$ denotes the empty set. As far as we know it was E. Sperlrajn-Marczewski (2) who first proved that the converse is also true; see also Čulik (1).

**Theorem SM.** Let $G$ be an arbitrary graph. Then there is a set $S$ and a family of subsets $S_1, S_2, \ldots$ of $S$ which can be put into one-to-one correspondence with the vertices of $G$ in such a way that $1)$ holds.

Notice that Theorem SM remains true if we replace $S_a \cap S_b \neq \emptyset$ by $S_a \cap S_b = \emptyset$ in (1), because we can always replace $G$ by its complement.

Our objective in this paper is to determine the minimum number of elements in the set $S$. In fact we shall prove the following theorem.

**Theorem 1.** If $G$ is any graph with $n$ vertices, then there is a set $S$ with $\lceil n^2/4 \rceil$ elements and a family of $n$ subsets of $S$ such that $1)$ holds. Further $\lceil n^2/4 \rceil$ is the smallest such number.

2. Coverings by complete graphs. A graph $G$ is said to be complete if every pair of points of $G$ is joined by an edge of $G$. A complete graph on two points is just a line segment, and a complete graph on three points is just a triangle. We define the sum $G = G_1 + G_2$ of two graphs as follows: (1) $x$ is a vertex of $G$ if it is a vertex of $G_1$ or of $G_2$. (2) $x_a$ and $x_b$ are connected by an edge in $G$ if they are connected by an edge in $G_1$ or in $G_2$. We remark that if they are connected in both $G_1$ and $G_2$, then they are still connected by just a single edge in the sum. If a graph $G$ is the sum of graphs $G_1, G_2, \ldots, G_n$, we

---

Received September 14, 1964.
shall say that these graphs cover \( G \). An isolated point of a graph is a point that does not belong to any edge of the graph. The number of vertices of \( G \) is called the order of \( G \). With these agreements we have the following theorem.

**Theorem 2.** Any graph \( G^{(n)} \) of order \( n \geq 2 \) with no isolated points can be covered by at most \([n^2/4]\) complete graphs. Further, in the covering we need to use only edges and triangles.

**Proof.** We use induction, going from index \( n \) to index \( n + 2 \). The theorem is obviously true for \( n = 2 \) and \( n = 3 \). Further, we note that for any positive integer \( n \)

\[
(n + 2)^2/4 = [n^2/4] + n + 1.
\]

Now let \( G^{(n+2)} \) be a graph of order \( n + 2 \) and let \( x_1 \) and \( x_2 \) be any two points that are connected by an edge of \( G^{(n+2)} \). Let \( G^{(n)} \) be the subgraph consisting of the vertex set \( V = \{x_1, x_2, \ldots, x_{n+2}\} \) and those edges of \( G^{(n+2)} \) that connect pairs of points in \( V \). By hypothesis, this graph can be covered by at most \([n^2/4]\) triangles and edges. Consider \( x_3 \in V \). If \( x_3 \) is joined to both \( x_1 \) and \( x_2 \) in \( G^{(n+2)} \), then we introduce a triangle \( x_1 x_2 x_3 \), and call this \( G_n \). If \( x_3 \) is connected to \( x_1 \) or \( x_2 \), but not both, then we introduce for \( G_n \) an edge \( x_1 x_3 \) or \( x_2 x_3 \). If \( x_3 \) is not connected to either \( x_1 \) or \( x_2 \), then there is no need to introduce a line segment or triangle. Hence for \( k = 3, 4, \ldots, n + 2 \) we have at most \( n \) complete graphs \( G_n \). Finally we need \( G_k \), the edge connecting \( x_1 \) and \( x_2 \). Since \( G^{(n)} \) is a sum of at most \([n^2/4]\) edges and triangles, \( G^{(n+2)} \) is the sum of at most \([n^2/4] + (n+1)\) edges and triangles. From (2), this completes the proof of the theorem.

Theorem 2 was also proved independently by L. Lovász (oral communication).

It is easy to prove that the number \([n^2/4]\) that occurs in Theorem 2 cannot be replaced by any smaller number. Let \( n = 2k \) or \( 2k + 1 \) and let \( A \) be a collection of \( k \) points and \( B \) be a collection of the remaining points (either \( k \) or \( k + 1 \) in number). We define \( T^{(n)} \) to be the special graph of order \( n \) in which \( x_a \) and \( x_r \) are joined by an edge if and only if one of the points is in \( A \) and the other point is in \( B \). Clearly \( T^{(n)} \) has no triangles and the number of edges is:

\[
k^2 = [(2k)^2/4] = [n^2/4], \quad \text{if } n \text{ is even},
\]

and

\[
k(k + 1) = [(2k + 1)^2/4] = [n^2/4], \quad \text{if } n \text{ is odd}.
\]

Hence the graph \( T^{(n)} \) will always require \([n^2/4]\) complete graphs for a cover.

We shall give a refinement of Theorem 2 in \( \S \).4.

3. **Proof of the main result.** If the graph \( G \) of Theorem 1 has any isolated points \( x_m \), we can select the empty set for \( S_m \) and for such points condition (1) will be satisfied. Hence in proving Theorem 1 we may assume that \( G \) has no isolated points. We next cover \( G \) with \( N \) complete graphs \( G_1, G_2, \ldots, G_N \),

\[
88 \quad \text{GRAPH THEORY}
\]
where \( N < \lfloor n^2/4 \rfloor \). By Theorem 2, this can be done and in fact in such a way that each \( G_e \) must be either an edge or a triangle. With each graph \( G_e \) we associate an element \( e_e \) and with each point \( x_a \) we associate a set \( S_a \) of elements \( e_a \), where
\[
e_e \in S_e \quad \text{if and only if} \quad x_e \in G_e,
\]
i.e. \( S_a \) is the collection of those elements for which the corresponding complete graphs contain \( x_a \). If we set
\[
S = \bigcup_{a=1} S_a,
\]
then clearly \( S \) contains \( N \) elements. Further, \( S_a \cap S_b \neq \emptyset \) if and only if there is a common element \( e_e \). This means that \( x_a \) and \( x_b \) belong to the same complete graph \( G_e \), and this means that \( x_a \) and \( x_b \) are joined by an edge in \( G \). Conversely, if \( x_a \) and \( x_b \) are joined by an edge in \( G \), this edge will appear in some \( G_e \) in the cover and hence \( e_e \) will be in both \( S_a \) and \( S_b \). Consequently, the sets constructed by condition (3) will satisfy condition (1). This concludes the proof of the first part of Theorem 1.

To see that \( \lfloor n^2/4 \rfloor \) is the smallest number for which Theorem 1 is true, we turn again to the special graph \( T^{(n)} \). Here each edge must give rise to at least one element, for if \( x_a \) and \( x_b \) are joined, then \( S_a \cap S_b \) contains some element \( e_e \). But if this element were present in any other intersection, such as \( S_c \cap S_b \), then the points \( x_a, x_b, x_c \), or \( x_a, x_b, x_d \) would be vertices of a triangle in \( T^{(n)} \). But \( T^{(n)} \) contains no triangles. Hence each edge in \( T^{(n)} \) gives rise to at least one distinct element. Hence any representation of \( T^{(n)} \) by the intersection of sets satisfying the condition (1) must use at least \( \lfloor n^2/4 \rfloor \) elements.

In the reverse direction of Theorem 2 we can prove:

**Theorem 3.** Let \( G \) be a graph and suppose that for each point \( x_a \in G \) there is a set \( S_a \) such that condition (1) is satisfied. If the set \( S = \bigcup S_a \) contains \( N \) elements, then \( G \) is the sum of \( N \) complete graphs.

**Proof.** For each fixed \( e_e \) in \( S \) we form a complete graph \( G_e \) using those points \( x_a \) for which \( S_a \) contains \( e_e \). Clearly each \( G_e \) is a complete subgraph of \( G \) and \( G = G_1 + G_2 + \ldots + G_N \).

4. A refinement. Let \( G_a \) and \( G_b \) be two of the complete subgraphs constructed in the proof of Theorem 2. It is easy to see that \( G_a \) and \( G_b \) may have an edge in common. With a little more labour we can avoid this.

**Theorem 4.** Any graph \( G^{(a)} \) of order \( n \geq 2 \) with no isolated point can be covered by at most \( \lfloor n^2/4 \rfloor \) complete graphs \( G_1, G_2, \ldots, G_N \), and no two of the graphs \( G_a, G_b \) will have an edge in common. Further, in the covering we need to use only edges and triangles.

**Proof.** We say that a vertex \( x \) has valence \( k \) if \( k \) edges terminate at \( x \).

The theorem is obviously true for \( n = 2 \). We assume that it is true for all
graphs of order less than \( n \) and note that for any positive integer \( n \)

\[
(n^2/4) = [(n - 1)\sqrt[4]{4}] + \lfloor n/2 \rfloor.
\]

Hence in the induction we must show that in going from \( G^{(n-1)} \) to \( G^{(n)} \) we need add at most \( \lfloor n/2 \rfloor \) complete graphs that are pairwise edge disjoint.

Suppose \( G^{(n)} \) has a vertex of valence \( \leq \lfloor n/2 \rfloor \). Call this vertex \( x_1 \) and let \( G^{(n-1)} \) be the subgraph on the points \( \{x_2, x_3, \ldots, x_n \} \). Then in going from \( G^{(n-1)} \) to \( G^{(n)} \) we only need to use the edges joining \( x_1 \) to the other points of \( G^{(n)} \) for our complete graphs, and there are at most \( \lfloor n/2 \rfloor \) of these. In this case the proof is complete.

In the contrary case, every vertex of \( G^{(n)} \) has valence \( > \lfloor n/2 \rfloor \). Let \( x_1 \) be the vertex with the smallest valence \( t \), and set \( t = \lfloor n/2 \rfloor + r \), where by hypothesis \( r > 0 \). Let \( x_1 \) be joined to the vertices \( y_1, y_2, \ldots, y_r \), and let \( G^{(r)} \) be the subgraph of \( G^{(n)} \) spanned by \( y_1, y_2, \ldots, y_r \). Suppose that \( G^{(r)} \) has \( r \) independent edges; that is, no two edges have a common vertex. Call these edges

\[
(y_1, y_2), (y_3, y_4), \ldots, (y_{2r-1}, y_{2r})
\]

and remove them from \( G^{(n-1)} \). Cover the resulting graph with at most \( [(n - 1)^2/4] \) edges or triangles, that are pairwise edge disjoint. Then \( G^{(n)} \) is the sum of these complete graphs together with the triangles

\[
(x_1, y_1, y_2), (x_1, y_3, y_4), \ldots, (x_1, y_{2r-1}, y_{2r})
\]

and the edges \( (x_1, y_k), k = 2r + 1, 2r + 2, \ldots, t \). The number of graphs in the sum is at most

\[
[(n - 1)^2/4] + r + t - 2r = [(n - 1)^2/4] - r + \lfloor n/2 \rfloor + r = [n^2/4].
\]

To complete the proof, we shall show that \( G^{(r)} \) must have \( r \) independent edges. Assume that \( G^{(r)} \) has only \( r - 1 \) independent edges

\[
(y_1, y_2), (y_3, y_4), \ldots, (y_{2r-3}, y_{2r-2}).
\]

By hypothesis, \( y_{2r-1} \) has valence \( \geq \lfloor n/2 \rfloor + r \). It can be joined to at most \( 2r - 2 \) of the points \( y_1, y_2, \ldots, y_{2r-3} \), and to at most \( n - t \) of the points not in \( G^{(r)} \). Hence the valence of \( y_{2r-1} \) is at most

\[
2r - 2 + n - t = 2r - 2 + n - \lfloor n/2 \rfloor - r = n - \lfloor n/2 \rfloor + r - 2 < \lfloor n/2 \rfloor + r.
\]

But this is the minimum valence. Hence \( y_{2r-1} \) is joined to some other point in \( G^{(r)} \) and \( G^{(r)} \) has at least \( r \) independent edges. This completes the proof.

The graph \( T^{(n)} \) shows that the number \( [n^2/4] \), mentioned in the theorem, cannot be replaced by any smaller number.

5. Open questions. These results suggest a number of related problems. For example, suppose that the graph \( G^{(n)} \) has \( [n^2/4] + k \) edges, where \( k \) is a fixed positive integer. Then it is clear that \( G^{(n)} \) can be covered by fewer than \( [n^2/4] \) complete graphs. What then is the new minimum as a function of \( k \)?
Here it may be advantageous to use complete graphs of order greater than 3 if $k$ is large.

In another direction it seems as though every $G^{(r)}$ can be covered by at most $n - 1$ circuits (here a single edge is counted as a circuit) but so far we have not been able to prove this. If we add the side condition that the circuits be pairwise edge disjoint (no two circuits have an edge in common), then $n - 1$ circuits will not suffice as T. Gallai proved in the following way (oral communication). Let the vertices be denoted by $x_1, x_2, x_3, y_1, y_2, y_3, \ldots, y_{n-3}$ and let $G^*(n)$ be the particular graph with the $3(n - 3)$ edges $(x_\alpha, y_\beta), \alpha = 1, 2, 3, \beta = 1, 2, 3, \ldots, n - 3$. Aside from the trivial circuits consisting of single edges, all other circuits have either the form $C_1: (x_1, y_\alpha, x_2, y_\beta, x_3)$, or the form $C_2: (x_1, y_\alpha, x_2, y_\beta, x_3, y_\gamma, x_4)$, or suitable permutations of these. The requirement that the cover be edge disjoint forces the inclusion of the single edge circuits $(x_\alpha, y_\beta)$ and $(x_\gamma, y_\gamma)$ in any cover that includes $C_1$. Hence if all pairs $y_\beta, y_\gamma$ are covered by circuits of type $C_1$, the number of circuits required would be $3(n - 3)/2$ if $n$ is odd, and $3 + 3(n - 4)/2$ if $n$ is even.

If the edge-disjoint cover includes a circuit of type $C_2$, then it must also include the single-edge circuits $(x_1, y_\alpha)$, $(x_2, y_\beta)$, and $(x_3, y_\gamma)$. Suppose that $n \equiv 0 \pmod{3}$ and the $y_\beta$ vertices are grouped in sets of three and that the covering is made up of circuits of type $C_2$ and single-edge vertices. Then the number of circuits is $4(n - 3)/3$. Since this is less than the number of circuits used in the first case, it is clear that for $n \equiv 0 \pmod{3}$, the smallest number of edge-disjoint circuits needed to cover the special graph $G^*(n)$ is $4(n - 3)/3$. The cases $n = 1, 2 \pmod{3}$ lead to a similar result.

Let $f(n)$ denote the smallest integer such that every graph with $n$ vertices can be covered by $f(n)$ or fewer edge-disjoint circuits. The graph $G^*(n)$ proposed by Gallai shows that $\lim \inf f(n)/n \geq 4/3$. It can be shown that

$$f(n) < \frac{1}{2} n \log n + O(n),$$

but it may be true that $f(n) < cn$ for some suitable $c$.

6. Representation of a graph by distinct sets. In the proof of Theorem 1, the sets obtained need not be distinct. Indeed there may be two different vertices $x_\alpha$ and $x_\beta$ for which $S_\alpha = S_\beta$. Both James H. Reed and G. Sabidussi have pointed out that if the $n$ sets corresponding to the vertices of $G$ are required to be all different, then the proof of Theorem 1 is not sufficient. However, if $n > 4$, we obtain the same minimum as in Theorem 1.

**Theorem 5.** Let $d(n)$ be the smallest number of elements in $S$ with the property that for each graph on $n$ vertices, there is a family of $d(n)$ different subsets

$$S_\alpha (\alpha = 1, 2, \ldots, n)$$

of $S$ such that the relation (1) holds. Then $d(2) = 2, d(3) = 3$, and
\[ d(n) = \lfloor n^2/4 \rfloor \]

if \( n \geq 4 \).

**Proof.** The exceptional cases \( n = 2 \) and \( n = 3 \) are trivial. We proceed by induction from \( n \) to \( n + 2 \), and we first prove that if \( d(n) \leq \lfloor n^2/4 \rfloor \) for \( n = 4 \) and \( n = 5 \), then the same inequality holds for all \( n \geq 4 \).

If the graph \( G^{(n+2)} \) on \( n + 2 \) points has no edges, then we set \( S = \{ x_n \} \) for \( \alpha = 1, 2, \ldots, n + 2 \). This selection is satisfactory because \( n < \lfloor n^2/4 \rfloor \), for \( n \geq 4 \). Suppose that in the induction from \( n \) to \( n + 2 \) the graph \( G^{(n+2)} \) has an edge \((x_{n+1}, x_{n+2})\) and that neither of these two points are terminal points of any other edge. Let \( G^{(n)} \) be the graph on the \( n \) points \( \{x_1, x_2, \ldots, x_n\} \). Then we add two new elements \( e_1^* \) and \( e_2^* \) to the set for \( G^{(n)} \) and take for our new sets \( S_{n+1} = \{e_1^*, e_2^*\}, S_{n+2} = \{e_1^*, e_2^*\} \) while leaving the sets for \( G^{(n)} \) unchanged. In this case the induction is complete.

In all other cases the graph will contain an edge \((x_{n+1}, x_{n+2})\) that is connected with at least one other point of the graph. By equation (2), we have \( n + 1 \) new elements at our disposal. Call them \( e_1^*, e_2^*, e_3^*, \ldots, e_{n+1}^* \). We form the set \( S_{n+1}^* \) by putting in \( e_n^* \) \((\alpha = 1, 2, \ldots, n)\) if and only if \( x_n \) is connected to \( x_{n+1} \) by an edge. Similarly, \( S_{n+2}^* \) is the set of all \( e_n^* \) for which \((x_n, x_{n+2})\) is an edge in \( G^{(n+2)} \).

It may happen that one of the two sets \( S_{n+1}^* \) and \( S_{n+2}^* \) is empty. In this case we form the sets \( S_{n+1}^* \) and \( S_{n+2}^* \) by adding the element \( e_{n+1}^* \) to both \( S_{n+1}^* \) and \( S_{n+2}^* \). Then \( S_{n+1}^* \neq S_{n+2}^* \). If \( S_{n+1}^* \cap S_{n+2}^* \) is empty, we also adjoin the element \( e_{n+1}^* \) to both sets. In any other case we can set \( S_{n+1}^* = S_{n+1} \) and \( S_{n+2}^* = S_{n+2} \cup \{e_{n+1}^*\} \).

Let \( S_1, S_2, \ldots, S_n \) be the sets that satisfy Theorem 5 for \( G^{(n)} \). For \( \alpha = 1, 2, \ldots, n \), we form the new sets \( S^*_\alpha \) by adding \( e_\alpha^* \) to \( S_\alpha \) if \( x_\alpha \) is connected to either \( x_{\alpha+1} \) or \( x_{\alpha+2} \) by an edge. Then the sets \( S_1^*, S_2^*, \ldots, S_n^* \) satisfy the requirements of Theorem 5.

If \( n = 4 \), it is a simple matter to draw pictures of the 11 different graphs on four points, and to construct the necessary sets with at most \( \lfloor 4^2/4 \rfloor = 4 \) elements. The same technique can be used if \( n = 5 \), but in this case the number of different graphs is sufficiently large to make a short cut desirable. Since \( \lfloor 5^2/4 \rfloor - \lfloor 4^2/4 \rfloor = 6 - 4 = 2 \), we have available two new elements for \( S \) in passing from \( n = 4 \) to \( n = 5 \). If \( G^{(5)} \) has one vertex with valence 2 or less, it is a simple matter to proceed from \( G^{(4)} \) to \( G^{(5)} \) by deleting this vertex and its edges and then to go back to \( G^{(3)} \) using the two new elements. Hence one needs to consider only those graphs on five points for which each vertex has valence greater than or equal to 3. But there are only three such graphs and these are easy to discuss.

This proves that \( d(n) \leq \lfloor n^2/4 \rfloor \). But the same special graph \( T^{(n)} \) used in Theorem 1 also proves that \( d(n) \geq \lfloor n^2/4 \rfloor \).

**References**


92 GRAPH THEORY


*University of Alberta, Edmonton,*  
*The University of South Florida, and*  
*Michael Fazekas High School, Budapest*