The Number of Maximal Independent Sets in Connected Graphs

Zoltán Füredi
MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCE
BUDAPEST, 1364 P.O.B. 127
HUNGARY

ABSTRACT

Generalizing a theorem of Moon and Moser, we determine the maximum number of maximal independent sets in a connected graph on n vertices for n sufficiently large, e.g., $n > 50$.

1. INTRODUCTION AND EXAMPLES

Let $G = (V, E)$ be a simple graph (i.e., undirected, without loops and multiple edges.) The set of vertices joined to any particular vertex $x \in V$ of G will be denoted by $\Gamma(x)$. The degree of x is $d_G(x) = |\Gamma(x)|$. The maximum degree is denoted by $\Delta(G)$. For a set A the graph $G - A$ is obtained from G by removing the elements of A and their incident edges. For a positive integer n, K_n denotes the complete graph on n vertices, and $G^1 + G^2 + \dots + G^t$ or $\Sigma_{1 \le i \le t} G^i$ denotes the vertex disjoint union of the graphs $G^1, \dots, G^t$. The subset $A \subset V$ is independent if there is no edge $e \in E$, $e \subset A$. The subset A is maximal independent set if it is independent and $A \cup \{x\}$ is not independent for $x \in V - A$. Let $i(G) = |\{A : A \subset V, A \text{ maximal independent set in } G\}|$.

In this paper we deal with the function $i(G)$.

Example 1.1. Let $a_n = i(P_n)$, where P_n denotes the path of length n. Then $a_1 = 1$, $a_2 = 2$, $a_3 = 2$, and $a_n = a_{n-2} + a_{n-3}$ for $n \ge 4$. Hence $a_n \le 2\alpha^{n-2}$, where α is the (unique) real solution of the equation $1 + \alpha = \alpha^3$, $[\alpha = [(1 + \sqrt{23/27})/2]^{1/3} + [(1 - \sqrt{23/27})/2]^{1/3} = 1.32\dots)$.

Example 1.2. Let $b_n = i(C_n)$, where C_n denotes the circuit of length n. Then $b_3 = 3$, $b_4 = 2$, $b_5 = 5$, and $b_n = b_{n-2} + b_{n-3}$ for $n \ge 6$. Hence $b_n \le 3\alpha^{n-3}$. (For α, see Example 1.1).

Journal of Graph Theory, Vol. 11, No. 4, 463–470 (1987)
 CCC 0364-9024/87/040463-08$04.00

Example 1.3. Let $h(1) = 1$ and for $n \geq 2$ define

$$h(n) = \begin{cases} 3^t & \\ 4 \cdot 3^{t-1} & \text{for } n = \\ 2 \cdot 3^t & \end{cases} \begin{matrix} 3t \\ 3t + 1 \\ 3t + 2, \end{matrix}$$

where $t \geq 0$, integer. Define the graphs G_n with n vertices as follows:

$$G_n = \begin{cases} tK_3 \text{ (i.e., vertex disjoint union of } t \text{ triangles)} & \\ & \text{for } n = 3t, \\ \text{either } (t-1)K_3 + 2K_2 & \text{or} \quad (t-1)K_3 + K_4 \\ & \text{for } n = 3t+1 \text{ (i.e., in this} \\ & \text{case } G_n \text{ denotes two graphs)} \\ tK_3 + K_2 & \text{for } n = 3t + 2. \end{cases}$$

Clearly $i(G_n) = h(n)$.

Example 1.4 Let $F_2 = 2K_1$, $F_3 = P_3$ or $K_2 + K_1$, $F_4 = P_4$ or $K_3 + K_1$, $F_5 = C_5$, $F_6 = 3K_2$, $F_7 = C_5 + K_2$ and for $n \geq 8$

$$F_n = \begin{cases} (t-2)K_3 + 3K_2 & \\ (t-3)K_3 + 5K_2 & \text{for } n = \\ (t-2)K_3 + 4K_2 & \end{cases} \begin{matrix} 3t \\ 3t + 1 \\ 3t + 2 \end{matrix}$$

Denote by $h_2(n) = \max\{i(G): |V(G)| = n, \Delta(G) \leq 2, G$ is not isomorphic to G_n (see Example 1.3)$\}$. Our first result is the following easy

Proposition 1.5. $h_2(2) = 1$, $h_2(3) = 2$, $h_2(4) = 3$, $h_2(5) = 5$, $h_2(6) = 8$, $h_2(7) = 10$, and for $n \geq 8$ we have

$$h_2(n) = i(F_n) = \begin{cases} 8 \cdot 3^{t-2} & \\ 32 \cdot 3^{t-3} & \text{for } n = \\ 16 \cdot 3^{t-2} & \end{cases} \begin{matrix} 3t \\ 3t + 1 \\ 3t + 2, \end{matrix}$$

and the extremal graphs are given by Example 1.4.

Proof. Let $\mathbf{G}_2 = \{G: \Delta(G) \leq 2, |V(G)| = n, G \neq G_n\}$. Each $G \in \mathbf{G}_2$ consists of paths, circuits, and isolated vertices. Consider the components of a $G \in \mathbf{G}_2$, $G = G^1 + G^2 + \ldots + G^t$. Then $i(G) = \Pi_{1 \leq j \leq t} i(G^j)$. If the path P_k with $k \geq 7$ is a component of G, then deleting it and replacing it with $C_{k-3} + K_3$ we obtain a graph G', $G' \in G_2$, $i(G') > i(G)$. Similarly, the following operations increase $i(G)$: $C_k \to C_{k-3} + K_3$, $C_6 \to 3K_2$, $P_6 \to 3K_2$, $P_5 \to C_5$, $C_4 \to P_4$.

Suppose G is extremal, i.e., $i(G) = h_2(n)$, $G \in \mathbf{G}_2$. Then we obtained that G consists of some copies of C_5, P_4, K_3, P_3, K_2, and K_1. The following operations show that if C_5 is a component of G then $G = C_5$ or $C_5 + K_2$: $2C_5 \to 5K_2$, $C_5 + P_4 \to 4K_2 + K_1$, $C_5 + K_3 \to 4K_2$, $C_5 + P_3 \to 4K_2$, $C_5 + 2K_2 \to K_3 + 3K_2$, $C_5 + K_1 \to 3K_2$.

From now on we can suppose that G consists of only P_4, K_3, P_3, K_2, and K_1. The following operations imply that if P_4 is a component of G then $G = P_4$: $2P_4 \to 4K_2$, $P_4 + K_3 \to C_5 + K_2$, $P_4 + P_3 \to 3K_2 + K_1$, $P_4 + K_2 \to 3K_2$, $P_4 + K_1 \to 2K_2 + K_1$. In the same way, if P_3 is a component then $G = P_3$ because $2P_3 \to 3K_2$, $P_3 + K_3 \to 3K_2$, $P_3 + K_2 \to C_5$, $P_3 + K_1 \to K_3 + K_1$.

Hence we can suppose that $G = uK_3 + vK_2 + wK_1$. First we prove that $w \geq 1$ implies $n \leq 4$: $3K_1 \to K_2 + K_1$, $K_3 + 2K_1 \to C_5$, $K_2 + 2K_1 \to K_3 + K_1$, $2K_3 + K_1 \to C_5 + K_2$, $2K_2 + K_1 \to C_5$, $K_3 + K_2 + K_1 \to 3K_2$. Finally, for $n \geq 8$ we have that G consists of some copies of K_3 and K_2 only. Using $6K_2 \to 2K_3 + 3K_2$ we get that an extremal G is isomorphic to F_n. ∎

2. GRAPHS WITH MAXIMUM NUMBER OF MAXIMAL INDEPENDENT SETS

Answering a question raised by Erdös and Moser, Erdös, and later Moon and Moser [5] proved that the graphs G_n given by Example 1.3 have the maximum number of maximal independent subsets. Here we prove somewhat more. Define $h_0(n) = h(n - 1)$ for $2 \leq n \leq 6$ and for $n \geq 6$ let

$$h_0(n) = \begin{cases} 8 \cdot 3^{t-2} & \\ 11 \cdot 3^{t-3} & \text{for } n = \begin{cases} 3t \\ 3t + 1 \\ 3t + 2 . \end{cases} \\ 16 \cdot 3^{t-2} & \end{cases}$$

Theorem 2.1. Suppose G is a graph with n vertices. Then either $G \simeq G_n$ [and then $i(G) = h(n)$] or $i(G) \leq h_0(n)$ holds.

This theorem is best possible as the following example shows:

Example 2.2. Let $H_i =: F_i$ for $2 \leq i \leq 5$ and $H_6 = : 2K_3 \cup \{e\}$, i.e., a connected graph with 6 vertices, 7 edges, and $H_7 = K_3 + K_4 \cup \{e\}$. (See Figure 1.) Generally, for $n \geq 6$ let

$$H_n = \begin{cases} H_6 + (t - 2)K_3 & \\ H_7 + (t - 2)K_3 & \text{for } n = \begin{cases} 3t \\ 3t + 1 \\ 3t + 2 . \end{cases} \\ H_6 + K_2 + (t - 2)K_3 & \end{cases}$$

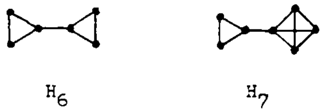

FIGURE 1.

Proof. If $\Delta(G) \leq 2$ then by Proposition 1.5 we have $i(G) \leq h_2(n) \leq h_0(n)$, and we are ready. From now on we can suppose that $\Delta(G) \geq 3$. We use induction on n. It is easy to check the cases $1 \leq n \leq 4$.

For $n \geq 5$ let x be a vertex with maximum degree. A maximal independent set D is maximal in $G - x$ if $x \notin D$. Similarly, if $x \in D$ then the set $D - x$ is maximal in $G - x - \Gamma(x)$. Hence we have

$$i(G) \leq i(G - x) + i[G - x - \Gamma(x)]. \tag{1}$$

Now (1) implies that

$$i(G) \leq h(n - 1) + h(n - 4).$$

Hence $i(G) \leq 5$ for $n = 5$. For $n \geq 6$ we can use $h(n - 4) = (1/3)h(n - 1)$. We obtain $i(G) \leq (4/3)h(n - 1) = h_0(n)$ for $n \equiv 0$ or $2 \pmod 3$.

From now on it is enough to consider the case $n = 3t + 1$ (≥ 7). If $\Delta(G) \geq 4$ then (1) yields

$$i(G) \leq h(3t) + h(3t - 4) = 3^t + 2 \cdot 3^{t-2} = h_0(3t + 1)$$

and we are ready.

If $G - x \neq G_{3t} = tK_3$ then $i(G - x) \leq 8 \cdot 3^{t-2}$ by the induction hypothesis, hence (1) gives

$$i(G) \leq h_0(3t) + h(3t - 3) = 8 \cdot 3^{t-2} + 3^t = h_0(3t + 1).$$

Finally, the only remaining case is $n = 3t + 1$, $G - x = tK_3$, $\Delta(G) = 3$, $G \neq G_n = K_4 + (t - 1)K_3$. Then either $G = L_{10} + (t - 3)K_3$, or $G = L_7 + (t - 2)K_3$, where $L_7(L_{10})$ is a connected graph on 7(10) vertices consisting of $2K_3(3K_3)$ and an extra vertex of degree 3 (see Figure 2). In both cases $i(G) = 3^t < h_0(n)$. ∎

3. INDEPENDENT SETS IN CONNECTED GRAPHS

H. S. Wilf [10] posed the problem to determine max $i(G)$ over the class of connected graphs. P. Erdös (private communication) conjectured that this maxi-

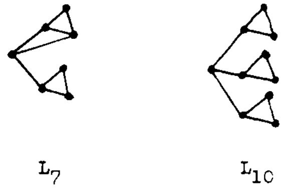

FIGURE 2.

mum is attained in a "connected version of G_n." In this section we prove this conjecture and determine the extremal graphs, at least for n sufficiently large.

Example 3.1. Let $G = tK_3, x \in V(G)$ and join x with one edge to each of the other $(t - 1)$ components. We obtain T_{3t} (see Figure 3).

For $n = 3t + 1$ let $G = K_1 + tK_3$ denote the isolated point by x. Join x with one edge to $(t - 1)$ copies of K_3 and with 3 edges to the tth copy. We obtain T_{3t+1}.

For $n = 3t + 2$ consider $G = K_4 + (t - 1)K_3 + K_1$, and join the isolated point with an edge to each component, and with 3 edges to one K_3.

Clearly, for the above defined graphs T_n $(n \geq 8)$ we have

$$i(T_n) =: c_0(n) = \begin{cases} 2 \cdot 3^{t-1} + 2^{t-1} & 3t \\ 3^t + 2^{t-1} & \text{for } n = 3t + 1 \\ 4 \cdot 3^{t-1} + 3 \cdot 2^{t-2} & 3t + 2\,. \end{cases}$$

Denote by $c(n) =: \max\{i(G): |V(G)| = n, G \text{ is connected}\}$.

Theorem 3.2. For $n > 50$ we have $c(n) = c_0(n)$.

Besides Theorem 2.1 the following lemma is the main tool of the proof:

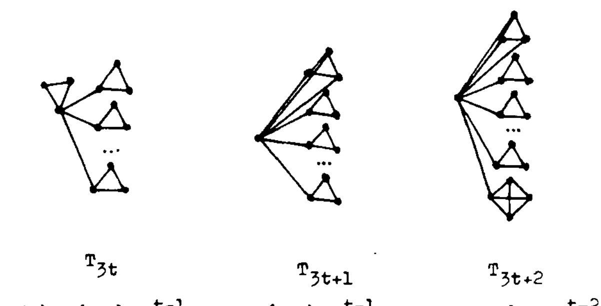

FIGURE 3.

Lemma 3.3. Let G be a connected graph with n vertices, $\Delta(G) \le 6$. Then $i(G) \le 3^{n/3}\, 1.009^{-n+3}$ [=: $z(n)$].

Proof of 3.3. It is easy to check that $h(n) \le z(n)$ for $1 \le n \le 4$ and $h_0(n) \le z(n)$ holds for $5 \le n \le 9$. We use induction on n. Let x be a vertex of G with maximum degree Δ. Consider the components of $G - x = G^1 + G^2 + \dots + G^u$ $(u \le \Delta)$ and $G - x - \Gamma(x) = H^1 + H^2 + \dots + H^v$ $(v \le \Delta(\Delta - 1))$. Apply the induction hypothesis for G^a and H^b. The inequality (1) yields

$$\begin{aligned} i(G) &\le i(G - x) + i[G - x - \Gamma(x)] \\ &= \Pi_{1\le t\le u} i(G^t) + \Pi_{1\le t\le v} i(H^t) \\ &\le \Pi_{1\le t\le u} z[|V(G^t)|] + \Pi_{1\le t\le v} z[|V(H^t)|] \\ &= 3^{(n-1)/3}\cdot 1.009^{-n+1+3u} + 3^{(n-1-\Delta)/3}\cdot 1.009^{-n+1+\Delta+3v} \\ &\le 3^{n/3}\cdot 1.009^{-n+3}\left[\frac{1}{3^{1/3}}1.009^{3\Delta-2} + \frac{1}{3^{(\Delta+1)/3}}1.009^{3\Delta(\Delta-1)+\Delta-2}\right] \end{aligned}$$

Here the sum in the parentheses is less than 1 for $3 \le \Delta \le 6$. The case $\Delta = 2$ follows from the fact $a_n, b_n \le 3\alpha^{n-3} \le z(n)$. (See Example 1.2). ∎

The proof of Theorem 3.2. Let G be a connected graph on n vertices with $i(G) = c(n) \ge 4 \cdot 3^{-5/3} \cdot 3^{n/3}$. For $n > 50$ we have $z(n) < c(n)$, hence by Lemma 3.3 we can suppose that $\Delta \ge 7$. Let x be a vertex with maximum degree. If $G - x \not\approx G_{n-1}$ given by Example 1.3, then (1) and Theorem 2.1 yields

$$\begin{aligned} h(n-1) < c(n) = i(G) &\le i(G - x) + i[G - x - \Gamma(x)] \\ &\le h_0(n-1) + h(n-8) \le \frac{11}{12}h(n-1) + \frac{1}{12}h(n-1) = h(n-1), \end{aligned}$$

a contradiction. Hence we obtain that $G - x \approx G_{n-1}$. Now, knowing the structure of G it is easy to calculate $i(G)$ and to show that $i(G)$ is maximal when $G \approx T_n$ (given by Example 3.1.). ∎

The first few values of our functions $a_n, b_n, c(n), \dots$ can be seen in Table 1.

4. REMARKS AND PROBLEMS

Let $\mathbf{F}$ be a family of finite sets. We say $\mathbf{F}$ has property Δ if for all distinct $A, B, C \in \mathbf{F}$ we have that the symmetric difference $A \Delta B \not\subset C$. Denote by

$$\begin{aligned} g(n) &= \max\{|\mathbf{F}| : \mathbf{F} \subset 2^V, |V| = n, F \text{ has property } \Delta\}, \quad \text{and} \\ g_k(n) &= \max\{|\mathbf{F}| : \mathbf{F} \subset 2^V, |V| = n, \\ &\qquad \mathbf{F} \text{ has property } \Delta \text{ and for all } F \in \mathbf{F} \text{ we have } |F| = k\}. \end{aligned}$$

Erdös and Katona posed the problem to determine g and g_k. They have

Conjecture 4.1 [4].

$$g(n) = h(n) \le 3^{n/3} \tag{2}$$

and

$$g_k(n) = \Pi_{0\le i\le k-1}\lfloor n + i/k\rfloor \le (n/k)^k. \tag{3}$$

Simple constructions show that the right-hand sides of (2) and (3) are lower bounds. The best upper bound for $g(n), g(n) \le 1.5^n$ is due to Frankl and the author [3]. For $k = 2$, $g_2(n) = \lfloor n^2/4\rfloor$ is a reformulation of Turán's theorem [7]. Bollobás [1] proved the case $k = 3$ and recently Sidorenko [6] proved the case $k = 4$. Some further results about $g_3(n)$ can be found in [2, 8]. Clearly, for a graph $G = (V, E)$ the maximal independent sets form a family with property Δ. Hence (2) would be a generalization of the Moon–Moser theorem.

Let $c_d(n) = \max\{i(G) : |V(G)| = n, G \text{ connected}, \Delta(G) \le d\}$.

Example 4.2. Let $V = \bigcup_{1\le i\le t(d-1)} V(K_3^i) \cup X$ where $|V(K_3^1)| = \cdots = |V[K_3^{t(d-1)}]| = 3, X = \{x_1, \dots, x_t\}$. Consider $G = t(d-1)K_3 = K_1^1 + K_1^2 + \cdots + K_1^{t(d-1)}$. Join the vertex x_i to triangles K_3^j with an edge for $(i-1)\cdot(d-1) + 1 \le j \le i(d-1) + 1$ (see Figure 4). We obtain the connected graph E_n for $n = (3d-2)t, \Delta(E_n) = d, i(E_n) = 3^{t(d-1)}$. This example shows that the coefficient 1.009 in Lemma 3.3 can be improved to at most $3^{1/39} = 1.028\dots$ We have the following:

TABLE 1.

n	P_n a_n	C_n b_n	G_n $h(n)$	F_n $h_2(n)$	H_n $h_0(n)$	T_n $c_0(n)$	$c(n)$	$z(n)$
1	1		1				1	1.46...
2	2		2	1	1		2	2.09
3	2	3	3	2	2	3	3	3
4	3	2	4	3	3	4	4	4.28
5	4	5	6	5	5	5	5	6.12
6	5	5	9	8	8	8	8	8.76
7	7	7	12	10	11	11	11	12.52
8	9	10	18	16	16	15	15	17.9
9	12	12	27	24	24	22		25.58
10	16	17	36	32	33	31		36.57
11	21	22	54	48	48	42		52.27
12	28	29	81	72	72	62		74.72

FIGURE 4.

Conjecture 4.3. For fixed $d \geq 3$, $\lim_{n\to\infty} \sqrt[n]{c_d(n)} = 3^{(d-1)/(3d-2)}$. It would be interesting to determine max $i(G)$ and the extremal graphs for other classes of graphs.

A more exact calculation, and a version of Lemma 3.3 about $c_5(n)$, yield that Theorem 3.2 holds for $n \geq 48$ and for $n = 40, 43, 45, 46$.

Conjecture 4.4. Theorem 3.2 holds for all n.

Note added in proof. Wilf's problem (and our Conjecture 4.4) was proved independently by J. Griggs, C. Grinstead and D. Guichard [9].

ACKNOWLEDGMENT

The author is indebted to A. Balog (Budapest, Hungary) and H. S. Wilf (Philadelphia, PA) for their helpful comments.

References

[1] B. Bollobás, Three-graphs without two triples whose symmetric difference is contained in a third. *Discrete Math.* **8** (1974) 21–24.
[2] P. Frankl and Z. Füredi, A new generalization of the Erdös–Ko–Rado theorem. *Combinatorica* **3** (1983) 341–349.
[3] P. Frankl and Z. Füredi, Union-free hypergraphs and probability theory. *Eur. J. Combinat.* **5** (1984) 127–131.
[4] G. O. H. Katona, Extremal problems for hypergraphs. In *Combinatorics* M. Hall, Jr., and J. H. Van Lint, Eds., Mathematics Centre Tracts 56, Amsterdam (1974), part 2, 13–42.
[5] J. W. Moon and L. Moser, On cliques in graphs. *Isr. J. Math.* **3** (1965) 23–28.
[6] A. F. Sidorenko, A combinatorial-analytical approach to extremal problems [in Russian]. *U.D.K.* **519** 1–517.
[7] P. Turán, An extremal problem in graph theory. *Mat. Fiz. Lapok* **48** (1941) 436–452 [in Hungarian].
[8] D. de Caen, Uniform hypergraphs with no block containing the symmetric difference of any two other blocks. *Congressus Numerantium* **47** (1985) 249–253.
[9] J. R. Griggs, C. M. Grinstead, and D. Guichard, The maximum number of maximal independent sets in a connected graph, *Discrete Math.*, to appear.
[10] H. S. Wilf, The number of maximal independent sets in a tree, *SIAM J. Alg. Disc. Meth.* **7** (1986) 125–130.