

## A NOTE ON INDEPENDENT SETS IN TREES\*

BRUCE E. SAGAN†

**Abstract.** We give a simple graph-theoretical proof that the largest number of maximal independent vertex sets in a tree with  $n$  vertices is given by

$$m(T) = \begin{cases} 2^{k-1} + 1 & \text{if } n = 2k, \\ 2^k & \text{if } n = 2k + 1, \end{cases}$$

a result first proved by Wilf [*SIAM J. Algebraic Discrete Methods*, 7 (1986), pp. 125–130]. We also characterize those trees achieving this maximum value. Finally we investigate some related problems.

**Key words.** independent vertices, trees, extremal graphs

**AMS(MOS) subject classifications.** 05C35, 05C30, 05C70

**1. Introduction.** Herbert Wilf [5] was the first to answer the following question: What is the largest number of maximal independent vertex sets in a tree with  $n$  vertices? His proof had an algebraic flavor and was somewhat complicated. Subsequently Daniel Cohen [1] was able to provide a graph-theoretical proof, but one which was still fairly complex in view of the simplicity of the bound (see Theorem 3 below). The purpose of this note is to give a simple graph-theoretical demonstration of this result which, in addition, completely characterizes all trees achieving the maximum value. J. Griggs and C. Grinstead [2] independently found a straightforward proof which is similar to ours in some respects but differs in others.

**2. Maximizing independent sets.** We begin with some preliminary definitions and lemmas. For any concepts that are not defined, the reader can consult Harary's book [4].

Given a graph,  $G$ , let  $V(G)$  be the vertex set of  $G$  and let  $v(G) = |V(G)|$  where  $| \cdot |$  denotes cardinality. Recall that a vertex  $u \in V(G)$  is called an *endpoint* if  $\deg u = 1$ . We will say that a vertex  $v \in V(G)$  is *penultimate* if  $v$  is not an endpoint and  $v$  is adjacent to (at least)  $\deg v - 1$  endpoints. Note that  $v$  is adjacent to  $\deg v$  endpoints if and only if  $v$  is the center of the star  $K_{1n}$ .

**LEMMA 1.** *Every finite tree  $T$  with  $v(T) \geq 3$  has a penultimate vertex.*

*Proof.* The next-to-last vertex on any diameter must be penultimate.  $\square$

If  $v$  is penultimate in  $T$ , then  $T - v$  consists of  $\deg v - 1$  isolated vertices and one other component called the *penultimate component*  $P$  (if  $v$  is the center of a star, choose any fixed component as the penultimate one). Now let

$$\text{End } v = \{w \notin P \mid w \text{ is adjacent to } v\}$$

so that  $V(T) = V(P) \cup \{v\} \cup \text{End } v$  where  $\cup$  denotes disjoint union.

Call a set  $I \subseteq V(G)$  *independent* if no two vertices of  $I$  are adjacent in  $G$ . Now let  $M(G) = \{I \subseteq V(G) \mid I \text{ is independent and maximal}\}$ , i.e., if  $I \in M(G)$  then there is no independent set  $J$  with  $I \subsetneq J$ . Also set  $m(G) = |M(G)|$ . We wish to find the maximum value of  $m(T)$  over all trees  $T$  with  $v(T) = n$ . First, however, we need an upper bound.

\* Received by the editors April 28, 1986; accepted for publication (in revised form) March 24, 1987. This research was supported in part by a NATO post-doctoral grant administered by the National Science Foundation.

† Department of Mathematics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6395. Present address, Department of Mathematics, Michigan State University, East Lansing, Michigan 48824.

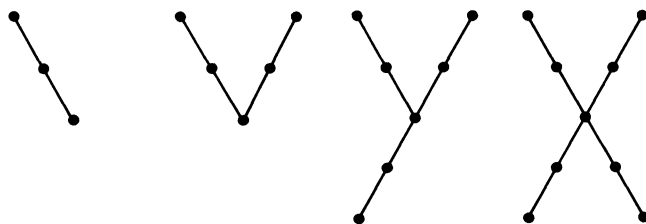


FIG. 1. Batons of length 0.

LEMMA 2. Let  $T$  be a tree and  $v \in V(T)$  be penultimate with corresponding component  $P$ . Then

$$m(T) \leq 2m(P).$$

*Proof.* Let  $I$  be a maximal independent set in  $T$ . Then either  $\text{End } v \subseteq I$  or  $v \in I$  (exclusive or). In the first case,  $I = \text{End } v \cup I_p$  where  $I_p$  is a maximal independent set of  $P$ . In the second,  $I = \{v\} \cup (I_p - \{w\})$  where  $w$  is the unique vertex of  $P$  adjacent to  $v$  ( $w$  need not be in  $I_p$ ).  $\square$

Define a *baton of length  $l$*  as follows. Start with a path  $L$  of length  $l$  and attach any number of paths of length two to  $L$ 's endpoints. Hence the batons of length 0 are just "extended" stars and the first few are displayed in Fig. 1.

Similarly, the batons of length 1 form a family some of whose members are shown in Fig. 2.

THEOREM 3. Among all labeled trees  $T$  with  $n$  vertices, the maximum value of  $m(T)$  is

$$m(T) = \begin{cases} 2^{k-1} + 1 & \text{if } n = 2k, \\ 2^k & \text{if } n = 2k + 1. \end{cases}$$

Furthermore this maximum is attained only by the batons of length 0 (when  $n$  is odd) or by the batons of lengths 1 and 3 (when  $n$  is even).

*Proof.* Induct on  $n$ . The theorem can be checked by hand for  $v(T) \leq 10$  using Lemma 2 and Harary's tables [4, pp. 233–234]. (The author has made this calculation and does not recommend it to the reader.) Now let  $T$  be a tree with  $m(T)$  maximum among all trees with  $v(T) = n > 10$ . By Lemma 1 there exists a penultimate vertex  $v \in V(T)$  with corresponding component  $P$ . There are two cases depending upon the parity of  $n$ .

If  $n$  is odd,  $n = 2k + 1$ , then consider the unique baton of length 0 with  $n$  vertices, denoted  $T_n$ . Since  $T_{2k+1}$  contains  $k$  paths of length 2, a simple calculation shows that  $m(T_{2k+1}) = 2^k$ . Hence

$$(1) \quad m(T) \geq 2^k.$$

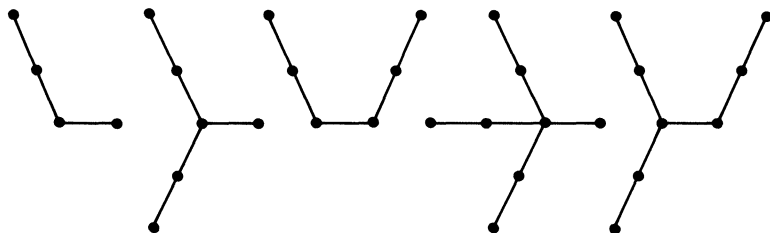


FIG. 2. Batons of length 1.

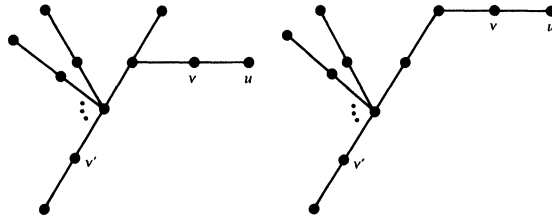


FIG. 3.  $G_1$  and  $G_2$ .

Now  $v(P) \leq n - 2 = 2k - 1$ . However, if  $v(P) < 2k - 1$ , then by Lemma 2 and induction (for  $n$  at least 7) we have

$$(2) \quad m(T) \leq 2m(P) < 2 \cdot 2^{k-1} = 2^k,$$

which contradicts (1). Thus  $v(P) = 2k - 1$  which implies that  $\deg v = 2$ , and  $\text{End } v = \{u\}$  for a single vertex  $u$ . Furthermore  $P = T_{2k-1}$  since if it does not, induction applies which yields  $m(P) < 2^{k-1}$  (this baton is the unique tree attaining the maximum value). But then (2) holds as before, a contradiction unless  $P = T_{2k-1}$ .

Putting all these facts together, we see that  $T$  consists of a tree  $T_{2k-1}$  with a path of length two  $w-v-u$  attached to some  $w \in V(T_{2k-1})$ . This leaves only three possibilities for  $T$ :  $T_{2k+1}$ ,  $G_1$  or  $G_2$ , where  $G_1$  and  $G_2$  are given in Fig. 3.

To eliminate  $G_1$  and  $G_2$  as possibilities, consider a second penultimate vertex  $v'$  as shown. If  $p'$  is the corresponding component, then for  $n \geq 9$  we have  $P' \neq T_{2k-1}$ . Invoking Lemma 2 again we see that

$$m(G_i) \leq 2m(P') < 2 \cdot 2^{k-1} = 2^k$$

for  $i = 1, 2$  and so neither  $G_i$  maximizes  $m(T)$ .

For  $n$  even,  $n = 2k$ , exactly the same line of reasoning as in the odd case can be used. It follows that the only possibilities for  $T$  are those obtained by attaching a path of length 2 to a baton of length 1 or 3 on  $2k - 2$  vertices. Hence  $T$  is either a baton of length 1 or 3 and  $m(T) = 2^{k-1} + 1$  or  $T$  is one of the five graphs  $H_1, \dots, H_5$  shown in Fig. 4.

Note that in  $H_2$  (respectively  $H_3$ ) we require that  $\deg c \geq 3$  (respectively  $\deg d \geq 2$ ) so that the graph does not degenerate into a baton of length 3 (respectively 1).

It is easy to verify that because  $n > 10$  we can find in each of these five graphs a second penultimate vertex  $v'$  such that  $P'$  is not a baton of length 1 or 3. It follows from

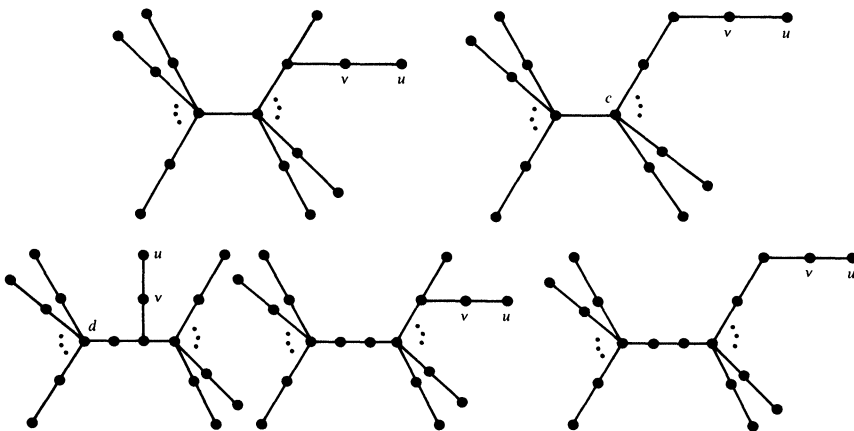


FIG. 4.  $H_1$  through  $H_5$ .

induction and from  $v(P') = n - 2 = 2k - 2$  that we have  $m(P') \leq (2^{k-2} + 1) - 1 = 2^{k-2}$ . Hence by Lemma 2

$$m(H_i) \leq 2m(P') \leq 2 \cdot 2^{k-2} = 2^{k-1}$$

for  $i = 1, \dots, 5$  which is less than the value obtained for the batons. This finishes the proof of the theorem.  $\square$

Call a tree  $T$  *extremal* if  $m(T)$  is a maximum as compared to all other trees with the same number of vertices. Let  $e(n)$  be the number of extremal trees, up to the labeling, on  $n$  vertices.

**COROLLARY 4.** *The number of extremal trees on  $n$  vertices is given by*

$$e(n) = \begin{cases} 1 & \text{if } n = 2k + 1, \\ k - 1 & \text{if } n = 2k. \end{cases}$$

*Proof.* It is a simple matter to count the number of batons of the appropriate sizes.  $\square$

**3. Remarks.** Finding the minimum value of  $m(T)$  is quite easy.

**THEOREM 5.** *The minimum value of  $m(T)$  over all trees with  $n$  vertices,  $n \geq 2$ , is  $m(T) = 2$ . Furthermore the unique tree (up to relabeling) achieving this minimum is the star  $K_{1, n-1}$ .*

*Proof.* If  $v(T) \geq 2$ , then for any edge  $vw$  there is a maximal independent set containing  $v$  and a different one containing  $w$ . Hence  $m(T) \geq 2$  and clearly  $m(K_{1, n-1}) = 2$ .

If  $T \neq K_{1, n-1}$ , then  $T$  contains a path  $u-v-w-x$ . This forces  $m(T) \geq 3$  since a third maximal independent set containing  $u$  and  $x$  also exists.  $\square$

Once one has determined the lower and upper bounds,  $b$  and  $B$ , respectively, for a graphical invariant  $\beta(G)$  one looks for an *interpolation theorem*. Such a result has the following form: For all integers  $z$  satisfying  $b \leq z \leq B$  there is a graph  $G$  with  $\beta(G) = z$ . Unfortunately there is no interpolation theorem for  $m(T)$  since when  $v(T) = 9$  we have  $2 \leq m(T) \leq 16$  but there is no tree with  $m(T) = 15$ .

We should compare our proof of Theorem 3 with that of Griggs and Grinstead mentioned in the Introduction. They also begin by proving Lemmas 1 and 2. Then, however, they use the lemmas to show that the maximum value of  $m(F)$  over all forests  $F$  with  $v(F) = n$  is achieved precisely when  $F$  is a one-factor (possibly with an isolated vertex). By carefully amalgamating the components of the one-factor, they finally obtain the bound and extremal graphs for trees.

Following the dictum that once something is proved for trees it should be extended to all connected graphs, one is led to pose the following question: What is the maximum value of  $m(G)$  over all connected graphs  $G$  with  $v(G) = n$ ? Griggs, Grinstead and Guichard [3] have answered this query. Another proof has been found independently by Füredi.

**Acknowledgments.** I would like to thank Mihály Hujter for pointing out the proof of Lemma 1. I also thank the referee for suggestions that considerably improved the exposition.

#### REFERENCES

- [1] D. COHEN, *Counting stable sets in trees*, in *Seminaire Lotharingien de combinatoire*, 10<sup>ème</sup> session, R. König, ed., Institute de Recherche Mathématique Avancée pub., Strasbourg, France, 1984, pp. 48–52.
- [2] J. GRIGGS AND C. GRINSTEAD, private communication.
- [3] J. GRIGGS, C. GRINSTEAD AND D. GUICHARD, *The number of maximal independent sets in a connected graph*, preprint.
- [4] F. HARARY, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.
- [5] H. WILF, *The number of maximal-independent sets in a tree*, *SIAM J. Algebraic Discrete Methods*, 7 (1986), pp. 125–130.