JOURNAL OF COMBINATORIAL THEORY, Series B 35, 207–223 (1983)

On the Average Number of Nodes in a Subtree of a Tree

ROBERT E. JAMISON*

Department of Mathematical Science,
Clemson University, Clemson, South Carolina 29631

Communicated by the Editors

Received September 14, 1981

For any tree T (labelled, not rooted) of order n, it will be shown that the average number of nodes in a subtree of T is at least $(n+2)/3$, with this minimum achieved iff T is a path. For T rooted, the average number of nodes in a subtree containing the root is at least $(n+1)/2$ and always exceeds the average over all unrooted subtrees. For the maximum mean, examples show that there are arbitrarily large trees in which the average subtree contains essentially all of the nodes. The mean subtree order as a function on trees is also shown to be monotone with respect to inclusion.

1. INTRODUCTION

Throughout this paper T will denote a tree (connected, acyclic, undirected graph) on n nodes. By a *k-subtree* S of T we shall mean a set of k nodes of T which induce a connected subgraph of T. The collection of all subtrees of T forms a meet-distributive lattice (cf. Edelman [3]) in which the k-subtrees are precisely the elements of height (or rank) k. The number $A_k(T)$ of k-subtrees is thus the kth *Whitney number* (or rank number) [4] of the lattice of subtrees.

For any tree, $A_1(T)$ is the number n of nodes in T, $A_2(T)$ the number $n-1$ of edges, $A_{n-1}(T)$ the number of endnodes and $A_n(T)=1$. In general, the other A_k are more difficult to describe although there are simple recursive procedures for calculating them [7].

The average number of nodes in a subtree of T will be denoted M_T and called the (*global*) *mean* of T. (For technical reasons, it is desirable not to count the empty subtree in this average.) If we define the generating function,

$$\Phi_T(x) = \sum_{k=1}^{n} A_k(T)\, x^k, \tag{1.1}$$

* The author was supported for this research from NSF Grant MSC-80-02543 and NSF EPSCoR Grant ISP-80-11451.

0095-8956/83 $3.00

then the mean may be expressed by the logarithmic derivative

$$M_T = \Phi_T'(1)/\Phi_T(1). \tag{1.2}$$

The mean M_T provides a rough measure for shape of the lattice of subtrees of T and is a rather interesting invariant of the tree T. For the sake of comparing trees of different orders, it is convenient to normalize the mean and define the *density* of T to be $\mathrm{Den}(T) = M_T/n$. Thus if $\mathrm{Den}(T)$ is close to 1, the lattice of subtrees is top-heavy, and if $\mathrm{Den}(T)$ is near 0, then the lattice is fuller at the bottom. The density of T may also be interpreted as the probability that a node chosen at random from T will belong to a randomly selected subtree of T. The means and densities of some selected trees of order 15 are shown in Fig. 1.

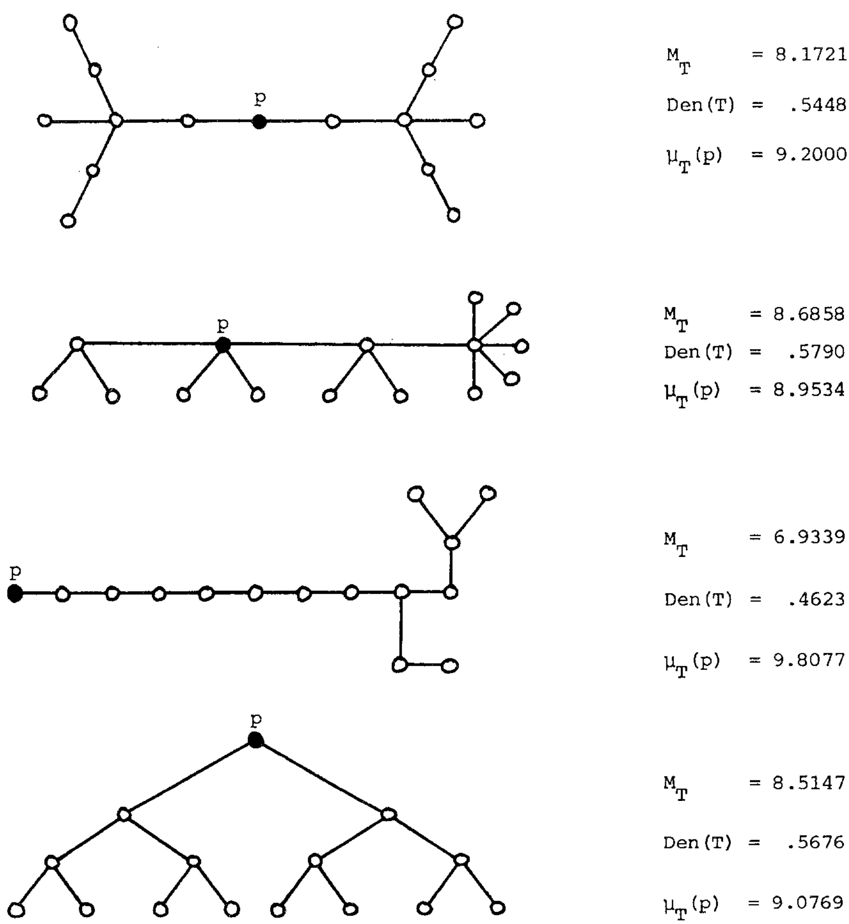

FIG. 1. Means in some selected trees of order 15.

Several authors [1, 2, 8–10, 13, 14] have investigated the average behavior of various parameters (e.g., height, distance, diameter) in trees. These studies have usually focused on averages over all trees and have often relied on analytic techniques. By contrast, this paper is concerned with mean order as an invariant of a fixed but arbitrary tree, and although generating functions are used, the arguments are purely combinatorial.

2. High Density Trees

Perhaps the first question to pose regarding an invariant concerns its extreme values. For trees, a reasonable first guess is that paths and stars provide the extreme values. This is indeed the case for many simple invariants such as average distance between nodes [2] but may fail for more complex invariants as the number of cycles in the complement [12].

For the path $P = P_{n-1}$ and star $S = K_{1,n-1}$ on n nodes, the Whitney number are easily seen to be

$$A_k(P) = n - k + 1 \qquad \text{and} \qquad A_k(S) = \binom{n-1}{k-1} \qquad \text{for} \quad 2 \leqslant k \leqslant n. \tag{2.1}$$

Applying standard summation formulae, one can then derive that

$$\Phi_P(1) = \binom{n+1}{2}; \qquad \Phi_P'(1) = \binom{n+2}{3}; \qquad M_P = (n+2)/3; \tag{2.2}$$

$$\begin{gathered} \Phi_S(1) = 2^{n-1} + n - 1; \qquad \Phi_S'(1) = (n+1)\, 2^{n-2} + n - 1; \\ M_S = (n+1)/2 - \varepsilon, \end{gathered} \tag{2.3}$$

where $0 < \varepsilon < (n-1)^2/2^n$. (Thus the path and the star have densities of roughly $\frac{1}{3}$ and $\frac{1}{2}$, respectively.)

It is true that paths and stars provide the extremal cases for the number of subtrees and, in fact, for the number A_k of k-subtrees for each k between 3 and $n - 1$. This may be verified by induction on n [16] but will be derived here in a "local" form from the results in Section 5.

It is also true that the path uniquely realizes the minimum mean for each n. Unfortunately, I do not know a simple proof of this and the proof given in Section 5 depends on a number of preliminary results. The idea of the proof, however, is simple: one shows (Theorem 5.9) that if a tree of sufficiently small density is *not* a path, then it can be modified (by moving a single edge) to produce a tree of smaller mean.

The conjecture that stars have the highest mean is true through order 8. For order $n \geqslant 9$, however, stars fail to have the greatest density. Figure 2 shows the trees of greatest known density for each order between 8 and 15.

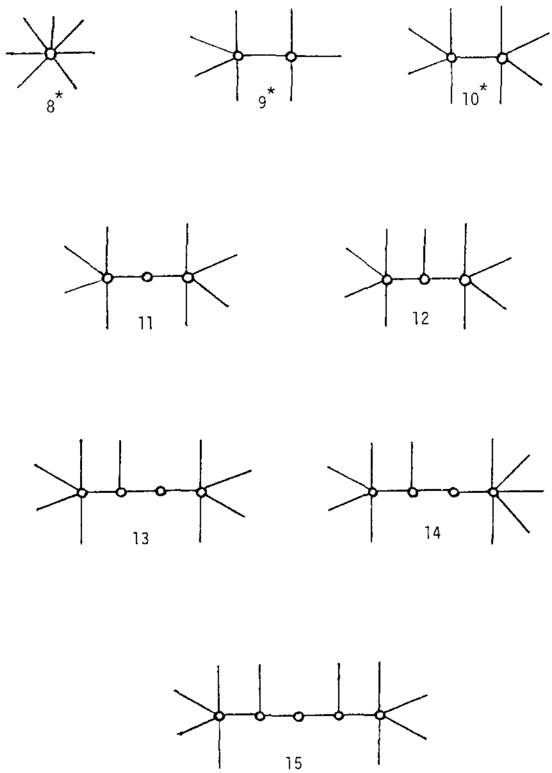

FIG. 2. High density trees of low order (* denotes confirmed maximum).

TABLE I

Means of High Density Trees in Fig. 2.

Order	Mean	Density	Density of Star of Order n
8	4.3185*	0.5398	0.5398
9	4.8993*	0.5443	0.5420
10	5.5405*	0.5540	0.5422
11	6.1550	0.5595	0.5410
12	6.8432	0.5702	0.5392
13	7.5014	0.5770	0.5371
14	8.2141	0.5867	0.5349
15	8.9250	0.5950	0.5329

Note. Asterisk denotes confirmed maximum.

Their densities are compared with the star of the same order in Table I. A computer search by David Whited through Harary's list [5] of trees has determined the exact maxima for orders $n = 1$ through 10.

It is not hard, in fact, to produce trees whose densities are arbitrarily close to 1. By an (s, t)-*baton* we shall mean the tree formed by joining the centers (say, u and v) of two stars $K_{1,s}$ by a path of t nodes, for a total of $n = t + 2s + 2$ nodes. There are then $\binom{t+1}{2} + 2s$ subtrees missing both u and v, $2[(t+1)\,2^s]$ subtrees containing exactly one of u and v, and 2^{2s} subtrees containing both u and v. It is clear that by appropriate choices of s and t (say, $t = s^2$) one can force the third type of subtree to predominate as $s \to \infty$. Hence the global mean will be nearly the average number of nodes in this type of subtree—namely, $t + s + 2 = n - s$.

There are even high density trees which have no vertices of high degree. An (s, t)-*bridge* is the tree formed by joining the centers (say, u and v) of two paths P_{2s} by a path P_{t-1}, for a total of $2(2s+1) + t$ nodes. Setting $t = s^{3/2}$ will again yield trees whose densities approach 1 as $s \to \infty$.

These examples illustrate the fact that, contrary to what one might first expect, high density trees do not tend to be "bushy." In fact, it will be shown (Theorem 6.2) that in any high density tree, most nodes are of degree 2.

Moon and Meir [11] have in fact determined that the *average* density over *all* trees of order n tends to $1 - e^{-1} = 0.6321...$ as $n \to \infty$. Thus for large n, most trees have density greater than that of the star. Nonetheless, the star is extremal in one special class of trees. It will be shown in Theorem 5.12 that among all trees with at most one node of degree >2 (the *asters*), the star has largest mean.

It remains an open problem to determine for each $n > 10$ the tree of greatest density of order n. Baton-like trees tend to have larger means than the bridges of the same order since the estimates for the batons are exponential, whereas those for the bridges are merely polynomial. It seems likely that the trees of maximal mean are similar to the batons, and I am willing to conjecture that they are caterpillars (i.e., trees from which the removal of the endnodes leaves a path).

3. The Local Mean at a Node

It is both useful in the development and interesting in its own right to consider a local or "rooted" version of average subtree order. If p is a node of T, the *local* mean $\mu_T(p)$ at p is the average number of nodes in a subtree containing p. Letting $\phi_T(p; x)$ denote the enumerator for the k-subtrees containing p, we may also write the local mean as the logarithmic derivative $\phi_T'(p; 1)/\phi_T(p; 1)$.

Let $v_1, \ldots, v_d$ denote the neighbors of p in T and let C_i be the component (branch) of $T\backslash p$ which contains v_i. Any subtree containing p is formed by gluing to p subtrees from the branches C_i. Since the component subtree in the ith branch either is empty or must contain v_i, we have the following basic recursion:

$$\phi_T(p; x) = x \prod_{i=1}^{d} (1 + \phi_{C_i}(v_i; x)). \tag{3.1}$$

To simplify the notation in the statement of some useful consequences of this recursion, set

$$L = \phi_T(p; 1), \qquad W = \phi_T'(p; 1), \qquad L_i = \phi_{C_i}(v_i; 1), \qquad W_i = \phi_{C_i}'(v_i, 1).$$

Thus L is the number of subtrees through p, W the sum of their orders, etc. Also for simplicity, the limits ($i = 1, \ldots, d$) will be suppressed in the sums and products below. The first relation (3.2a) below was noted by Ruskey (Lemma 1 in [15]) and used by him to study the average number of (rooted) subtrees in a random planar rooted tree.

(3.2) LEMMA. *With notation as above, we have*

(a) $L = \prod (1 + L_i)$,

(b) $\mu_T(p) = 1 + \sum \mu_{C_i}(v_i) - \sum W_i/(L_i(1 + L_i))$,

(c) $W/(L(1 + L)) \leqslant \frac{1}{2}$ with equality iff T is a path and p is an endnode,

(d) $\mu_T(p) \geqslant \mu_{C_i}(v_i) + d/2$ for $i = 1, \ldots, d$.

Proof. Part (a) follows by setting $x = 1$ in (3.1). To obtain (b), differentiate (3.1), set $x = 1$, and then factor to get $W = [\prod (1 + L_i)][1 + \sum W_i/(1 + L_i)]$. From (a) and the definition $\mu_T(p) = W/L$, one then gets

$$\mu_T(p) = 1 + \sum W_i/(1 + L_i). \tag{3.3}$$

Using the relation $W_i/(1 + L_i) = W_i/L_i - W_i/(L_i(1 + L_i))$ and the definition $\mu_{C_i}(v_i) = W_i/L_i$, one then obtains (b).

The argument for (c) goes by induction on the order of T, the result certainly holding for the one point tree. From (3.3) above and the inductive hypothesis, we have

$$W/L = 1 + \sum \left(\frac{W_i}{1 + L_i}\right)\left(\frac{L_i}{L_i}\right) \leqslant 1 + \sum \frac{L_i}{2} = \frac{1}{2}\left(2 + \sum L_i\right). \tag{3.4}$$

But by (a) we have

$$1 + L = 1 + \prod (1 + L_i) \geqslant 1 + \left(1 + \sum L_i\right) \tag{3.5}$$

since each L_i is positive. Combining (3.4) and (3.5) yields $W/L \leqslant \frac{1}{2}(1+L)$ from which the inequality in (c) follows.

Now if p is *not* an endnode of T, then $d \geqslant 2$ so the product in (3.5) includes the term $L_1 \ldots L_d > 0$ which is not accounted for in the right-hand sum in (3.5). Hence the inequality in (3.5) and therefore in (c) is strict in this case.

If p is an endnode but T fails to be a path, then $d=1$ and C_1 also fails to be a path. Thus by induction, the inequality in (3.4) and therefore in (c) is strict as desired.

If T is a path and p is an endnode, then there is exactly one k-subtree through p for each k. Hence $L=n$ and $W=\binom{n+1}{2}$, so $W/(L(1+L))=\frac{1}{2}$ as desired.

Finally to derive (d), note that by (c) the rightmost sum in (b) is at most $d/2$. Since any local mean is at least 1, (b) can be estimated by $\mu_T(p) \geqslant 1+\mu_{C_i}(v_i)+d-1-d/2$ from which (d) follows. ∎

It is now easy to determine those cases in which the local mean is minimized. A tree T is an *aster* iff T can be formed from a collection of disjoint paths by attaching one endnode of each to a "central" node p. In this case, T is *astral over p*. Evidently T is astral over p iff either T is itself a path or p is the unique node of degree >2 in T.

(3.6) THEOREM. *For any tree T and any node p of T, $\mu_T(p) \geqslant (|T|+1)/2$ with equality iff T is astral over p.*

Proof. For $|T|=1$ the result is obvious. By induction and Lemma 3.2(b),(c).

$$\mu_T(p) \geqslant 1+\sum(|C_i|+1)/2-d/2=(2+\sum|C_i|)/2=(|T|+1)/2. \quad (3.7)$$

By (3.2c) the inequality is strict unless each C_i is a path with v_i as an endnode, in which case T is astral over p. Conversely, if T is astral over p, each C_i is a path with endnode v_i and hence astral over v_i. Thus, equality holds in (3.2c) and in the induction hypothesis and hence in (3.7) as desired. ∎

To conclude this section we show that the local means always exceed the global mean. For this result we shall use the generating function form of the following well-known principle: if a population is partitioned into subpopulations, then its mean is an average of the means over the subpopulations. This will be used so often that it is worth stating formally. For brevity, the logarithmic derivative at $x=1$ of a polynomial $f(x)$ will be denoted by $\mathrm{LD}(f):=f'(1)/f(1)$.

(3.8) LEMMA. *Suppose $f_1,\ldots,f_m$ are nonzero polynomials with nonnegative coefficients. Then* $\mathrm{LD}(f_1+\cdots+f_m)$ *is a convex combination of the logarithmic derivatives* $\mathrm{LD}(f_i)$.

Proof. The coefficient of $\mathrm{LD}(f_i)$ is $f_i(1)/(f_1(1)+\cdots+f_m(1))$. ∎

(3.9) THEOREM. *For any tree T and any node p in T, $M_T \leqslant \mu_T(p)$ with equality iff T is the one point tree K_1.*

Proof. Noting first that when $|T|=1$, both means are trivially 1, we proceed by induction on T using the notation of (3.2). Since any subtree of T either contains p or lies in one of the branches C_i, $\Phi_T(x)=\phi_T(p;x)+\Sigma\,\Phi_{C_i}(x)$. Thus by (3.8), M_T is a convex combination of $\mu_T(p)$ and the means M_{C_i}. By (3.2d) and induction, $\mu_T(p)>\mu_{C_i}(v_i)\geqslant M_{C_i}$. Thus $\mu_T(p)$ is larger than the other means of which M_T is the average so $\mu_T(p)$ must be larger than M_T. ∎

In [6] it is shown that $\mu_T(p)<M_T+(n-1)/2$, but it seems likely that this can be sharpened.

4. LOCAL MEANS AT SUBTREES

If S is any set of nodes of T, the *local mean* $\mu_T(S)$ of T over S is the average number of nodes in a subtree of T containing all the nodes in S. Again, this is given by a logarithmic derivative $\phi_T'(S;1)/\phi_T(S;1)$ where $\phi_T(S;x)$ denotes the enumerator of the k-subtrees containing S. (When the elements of S are explicitly given, say as p, q, $r,\ldots$, we may write $\phi_T(p,q,r,\ldots;x)$ instead of $\phi_T(S;x)$.) Since every subset of T lies in a smallest subtree of T, there is no loss of generality in supposing S is a subtree. We begin the study of local means over subtrees with the following result on the local mean at an edge.

(4.1) LEMMA. *If p and q are adjacent nodes in T, then*

$$\mu_T(p)<\mu(p,q)\leqslant\mu_T(p)+\tfrac{1}{2}.$$

Proof. Removing the edge from p to q breaks T into two components. Let P (resp., Q) be the component containing p (resp., q). Since any subtree in T containing p and q can be realized uniquely as the union of a subtree of P through p with a subtree of Q through q, it follows that

$$\phi_T(p,q;x)=\phi_P(p;x)\,\phi_Q(q;x). \tag{4.2}$$

To simplify notation, set

$$A = \phi_P(p; 1), \qquad B = \phi_Q(q; 1), \qquad C = \phi_T(p, q; 1),$$
$$U = \phi'_P(p; 1), \qquad V = \phi'_Q(q; 1), \qquad W = \phi'_T(p, q; 1).$$

Thus A is the number of subtrees in component P containing node p, U is the sum of their orders, etc. Now by (4.2) and its derivative, we have $C = AB$ and $W = AV + BU$. Using this to calculate the local means yields

$$\mu_T(p, q) = W/C = (AV + BU)/AB = V/B + U/A$$

and

$$\mu_T(p) = (U + W)/(A + C) = (U + AV + BU)/(A + AB) = U/A + V/(1 + B).$$

Subtracting we get $\mu_T(p, q) - \mu_T(p) = V/B - V/(1 + B) = V/(B(1 + B))$ which is $\leqslant \frac{1}{2}$ by Lemma 3.2(c) applied to the node q in the tree Q. This establishes the right inequality in (4.1). The left inequality follows from the observation that since V and B are both at least 1, the above difference is always positive. ∎

On the side, this lemma yields an estimate of the variation in the local mean which may be regarded as a "continuity" result for the local mean μ_T as a function on the nodes of T.

(4.3) SCHOLIUM. *For any two nodes p and q of a tree T,*

$$|\mu_T(p) - \mu_T(q)| \leqslant \tfrac{1}{2} d_T(p, q),$$

where $d_T(p, q)$ is the distance from p to q in T.

Proof. If p and q are adjacent, both $\mu_T(p)$ and $\mu_T(q)$ lie between $\mu_T(p, q)$ and $\mu_T(p, q) - \frac{1}{2}$ by Lemma 4.1. Hence they are at most $\frac{1}{2}$ apart as desired. The general case follows by induction on $d_T(p, q)$. ∎

Now consider the *contraction* T/R of a tree T formed by contracting a subtree R of T to a point r. That is, T/R arises by identifying the nodes of R together as a single new node r (discarding any loops that occur). If S is any subtree of T containing R, then S/R is a subtree of T/R through r and $|S| = |S/R| + |R| - 1$. From this and the fact that subtrees of T containing S correspond to subtrees of T/R containing S/R, we get $\phi_T(S; x) = x^{|R|-1} \phi_{T/R}(S/R; x)$. Whence

$$\mu_T(S) = \mu_{T/R}(S/R) + |R| - 1 \tag{4.4}$$

whenever $R \subseteq S \subseteq T$.

(4.5) THEOREM. *If $R \subsetneq S$ are subtrees of T, then*

$$\mu_T(R) < \mu_T(S) \leqslant \mu_T(R) + (|S| - |R|)/2.$$

Proof. Consider first the case that S contains just one point more than R, say, $S = R \cup \{q\}$. Applying Lemma 4.1 to the adjacent nodes $r = R/R$ and q in T/R and then lifting back to T via (4.4) yields the result in this case.

In general we may list the $j = |S| - |R|$ nodes of $S \backslash R$ as $q_1, ..., q_j$ in such a way that, for each i, $R_i = R \cup \{q_1, ..., q_i\}$ is a subtree of T. By applying the first paragraph to each pair $R_i \subseteq R_{i+1}$, the general result follows inductively. ∎

From this, one may obtain the following extension of the bound in Theorem 3.6.

(4.6) THEOREM. *For any nonempty subtree of a tree T,*

$$\mu_T(R) \geqslant (|T| + |R|)/2.$$

Proof. Write the right side of (4.5) in the form $\mu_T(R) \geqslant \mu_T(S) - (|S| - |R|)/2$, set $S = T$, and note that $\mu_T(T) = |T|$. ∎

The left side of (4.5) is a monotonicity result for local means. It implies a similar result in which it is the ambient tree that changes.

(4.7) THEOREM. *For any subtree R of a proper subtree S of T, $\mu_S(R) < \mu_T(R)$.*

Proof. Since any subtree of T can be obtained from T by a sequence of endnode deletions, we may suppose that $S = T \backslash q$ for some endnode q of T, the general case following from this by induction. Letting Q denote the smallest subtree of T containing R and q, we may write $\phi_T(R; x) = \phi_S(R; x) + \phi_T(Q; x)$ since $\phi_S(R)$ enumerates the subtrees of T containing R which miss q, and $\phi_T(Q)$ enumerates the subtrees of T containing both R and q. Thus by Lemma 3.8, $\mu_T(R)$ is a convex combination of $\mu_S(R)$ and $\mu_T(Q)$. Since, by Theorem 4.5, $\mu_T(Q)$ is greater than $\mu_T(R)$, inequality (4.7) follows. ∎

To conclude this section, we prove that the global mean is also inclusion-monotone. It is worth noting by way of contrast that density is *not* monotone. This is illustrated by the density of the stars given in Table I. A further discussion of monotoncity results may be found in [6].

(4.8) THEOREM. *If S is a proper subtree of T, then $M_S < M_T$.*

Proof. Again it suffices to establish the result in the case S is obtained by deleting an endnode p from T. In this case, S is the unique branch of T at p, so

$$\Phi_T(x) = \phi_T(p; x) + \Phi_S(x).$$

Thus M_T is an average of $\mu_T(p)$ and M_S by Lemma 3.8. Because $\mu_T(p) > M_T$ by Theorem 3.9, it follows that M_S must be less than M_T. ∎

5. Paths Have Minimum Mean

Two trees T and T' on the same node set will be called *j-associates* iff there are nodes p and q and j edges $q - v_1, \ldots, q - v_j$ in T such that T' can be formed by deleting these edges and replacing them by the edges $p - v_1, \ldots, p - v_j$. We shall first be interested in a special kind of 1-associate of T. Let p be an endnode of T. If T is not a path, it has a node of degree >2 and hence a unique such node q nearest to p in T. Let v be any neighbor of q other than the one between p and q. The 1-associate T' formed by replacing the edge $q - v$ by $p - v$ will be called a *standard* 1-*associate* of T. (See Fig. 3.)

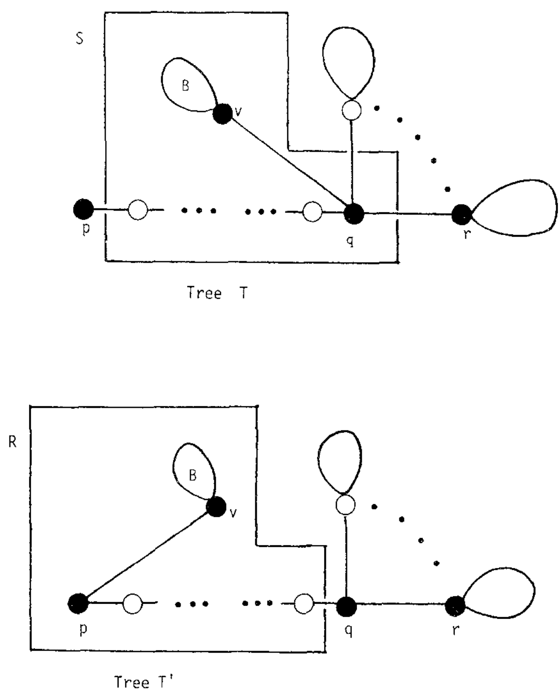

Fig. 3. 1-Associate diagram for Lemma 5.1.

The next result implies that paths uniquely minimize each of the nontrivial Whitney numbers A_k. Our interest, however, lies more with the technique of proof than the result itself since the reductions involved will be used in the proof the main result (5.11) below.

(5.1) LEMMA. *If T' is any standard 1-associate of T, then $A_k(T') < A_k(T)$ for all k such that $2 < k < n$.*

Proof. Our goal is to show that the difference $\delta(x) = \Phi_T(x) - \Phi_{T'}(x)$ has positive coefficients for $2 < k < n$. Removing the edge $q - v$ from T results in two components—say, A containing p and q, and B containing v—which are rejoined to form T' when $p - v$ is inserted. Evidently we have

$$\Phi_T(x) = \phi_T(q, v; x) + \Phi_A(x) + \Phi_B(x) \tag{5.2}$$

since any subtree of T not containing the edge $q - v$ must lie in either A or B. Subtracting the analogous formula for $\Phi_{T'}$ from (5.2), one arrives at

$$\delta(x) = \phi_T(q, v; x) - \phi_{T'}(p, v; x). \tag{5.3}$$

Because any subtree of T does or does not contain the endnode p of T, we have

$$\phi_T(q, v; x) = \phi_T(p, q, v; x) + \phi_{T\backslash p}(q, v; x). \tag{5.4}$$

Now let P be that portion of the tree T which lies between p and q. The choice of q ensures that P is a path. (It could be that $P = \{p, q\}$.) Thus $P\backslash p$ and $P\backslash q$ are both paths with $d(p, q)$ nodes, so the subtree $S = B \cup (P\backslash p)$ of T is isomorphic to the subtree $R = B \cup (P\backslash q)$ of T' by an isomorphism fixing v (in fact, all of B) and exchanging p and q. Therefore,

$$\phi_R(p, v; x) = \phi_S(q, v; x). \tag{5.5}$$

Notice that any subtree of T' which does not contain q but does contain the edge $p - v$ *must* lie in R, so

$$\phi_{T'}(p, v; x) = \phi_{T'}(p, q, v; x) + \phi_R(p, v; x). \tag{5.6}$$

Furthermore, a set of nodes containing p, q, and v forms a subtree (i.e., is connected) in T iff it forms a subtree in T', so

$$\phi_{T'}(p, q, v; x) = \phi_T(p, q, v; x). \tag{5.7}$$

Substituting (5.7) and (5.5) into (5.6), then subtracting from (5.4) and recalling (5.3), one deduces that

$$\delta(x) = \phi_{T\backslash p}(q, v; x) - \phi_S(q, v; x). \tag{5.8}$$

Thus to complete the proof, it need only be shown for $2 < k < n$, that $\alpha_k - \beta_k > 0$ where α_k (resp., β_k) is the number of k-subtrees of $T \setminus p$ (resp., S) which contain both q and v. Because S is a subtree of $T \setminus p$, clearly $\alpha_k \geqslant \beta_k$, so we need only exhibit for each k with $2 < k < n$ some k-subtree of $T \setminus p$ that does not lie in S.

Until now only two neighbors of q have been mentioned in the proof: v and the neighbor between q and p. By choice, $\deg_T(q) \geqslant 3$, so there is at least one other neighbor r of q in T. The set $\{r, q, v\}$ is a 3-subtree of $T \setminus p$ which can be enlarged in at least one way to a k-subtree of $T \setminus p$ for each k from 3 to $n-1$. Since $r \notin S$, none of these trees lie in S, and the proof is complete. ∎

The preceding result does not tell us, however, what happens to the mean in passing to a standard 1-associate. In general, it may go either up or down. This may be seen by considering the two types of standard 1-associate for an (s, s^2)-baton, s large. For trees of low density, however, the mean must in fact decrease. Although it seems likely that every tree other than a path has *some* (standard) 1-associate of lower mean, I have been unable to show this in general.

(5.9) Lemma. *Suppose T is a tree on n nodes with $M_T < (n+1)/2$. Then $M_{T'} < M_T$ for any standard 1-associate T' of T.*

Proof. Defining $\delta(x) = \Phi_T(x) - \Phi_{T'}(x)$ as above, write (5.8) as

$$\varphi_{T \setminus p}(q, v; x) = \phi_S(q, v; x) + \delta(x). \tag{5.10}$$

Because S is a subtree of $T \setminus p$, we may infer from Theorem 4.7 that the local mean at $\{q, v\}$ is less in S that in T/p. Since δ is nonzero with non-negative coefficients by Lemma 5.1, we can apply Lemma 3.8 to (5.10). This together with the above observation on local means yields $LD(\delta) \geqslant \mu_{T \setminus p}(q, v)$. But by (4.6),

$$\mu_{T \setminus p}(q, v) \geqslant (|T \setminus p| + 2)/2 = (n+1)/2.$$

whence from the hypothesis on M_T it follows that $\mathrm{LD}(\delta) > M_T$.

An application of Lemma 3.8 to $\Phi_T = \Phi_{T'} + \delta$ shows that M_T is an average of $M_{T'}$ and $\mathrm{LD}(\delta)$. Therefore, we may conclude from $\mathrm{LD}(\delta) > M_T$ that $M_{T'} < M_T$ as desired. ∎

The fact that paths have minimal means now follows as an immediate consequence.

(5.11) Main Theorem. *For any tree T on n nodes, $M_T \geqslant (n+2)/3$ with equality iff T is a path.*

Proof. Since $(n+2)/3 < (n+1)/2$ for $n \geqslant 2$ and since any tree other than a path has a standard 1-associate, it follows from Lemma 5.9 that the tree of order n with smallest mean subtree order M_T must be the path. ∎

Lemma 5.9 may also be applied to establish the following maximum property of stars.

(5.12) THEOREM. *Among all asters (trees with at most one node of* deg > 2) *on n nodes, the star $K_{1,n-1}$ uniquely achieves the largest mean.*

Proof. Let A be an aster on n nodes with maximum mean M_A. If A is not a star, there is an endnode v of A not adjacent to the central node q. Let p be the node of A adjacent to v. Remove the edge $p-v$ and insert an edge $q-v$ to form a new aster T. Note that p, q, and v in T conform to the selection procedure in the definition of standard 1-associates, so $A = T'$ is a standard 1-associate of T. By Theorems 3.9 and 3.6 we have $M_T < \mu_T(q) = (n+1)/2$ since T is astral over q. But then $M_A < M_T$ by Lemma 5.9, contrary to the choice of A with maximum mean. ∎

To round out the picture, we close with an analogue of (5.1) which implies that stars uniquely maximize each of the nontrivial Whitney numbers A_k. It is no longer possible to get by with moving just one edge as in (5.1). Indeed, if T has no nodes of degree 2, then moving only one edge could not create a new endnode and hence would not increase A_{n-1}. However, the edges to be moved need only be switched among adjacent nodes. A j-associate T' of T will be called *simple* if the nodes p and q of the definition of j-associate are adjacent.

(5.13) THEOREM. *If T is a tree on n nodes but T is not a star, then there is, for some j, a simple j-associate T' of T such that $A_k(T') > A_k(T)$ for all k such that $2 < k < n$.*

Proof. Let p be an endnode of T and q its unique neighbor. Since T is not a star, not every node of T is joined to q, so there are nodes r and s such that $p-q-r-s$ forms a path in T. Let U consist of all neighbors of q except r. Form T' by removing the edges from U to q and connecting the nodes in U with r instead.

There is a natural one-to-one, cardinality preserving map of the subtrees of T into the set of subtrees of T'. First, the subtrees of T which miss q are unaffected by the edge interchange and may be mapped to themselves as they are also subtrees of T'. Secondly, those subtrees of T which contain both q and r also remain subtrees (i.e., connected node sets of T') even after the edges are swapped although their adjacency structure may be altered. Finally we must account for those subtrees of T which contain q but not r. Associate q with itself. Any other such subtree S of T may be mapped to the subtree S'

of T' formed from S by replacing q by r. Even though this S' misses q, it is not of the kind first considered above since it must contain at least one node in U and hence is not connected in T. Thus our mapping is one-to-one.

Since $p-r-s$ is a path in T' and q is an endnode of T', it is possible to expand $\{p, r, s\}$ to a k-subtree of $T' \backslash q$ for each k from 3 to $n-1$. Notice that such a subtree is not associated with any subtree of T under the map described above. Whence the desired inequality in (5.13) follows. ∎

6. Some Properties of High Density Trees

The results on local means may also be used to obtain some rough information on high density trees. Somewhat surprisingly, they are more path-like than star-like in the sense that they have relatively few endnodes, contain many nodes of degree 2, and have large diameters. Below $V_d(T)$ denotes the number of nodes in T of degree d. The standard notions of radius and diameter are defined in [5].

(6.1) Lemma. *If T is a tree with $|T|>2$, then*

$$M_T < |T| - V_1(T)/2.$$

Proof. Let S be the subtree obtained by removing all the endnodes of T. Since the endnodes of T may be added independently to S to form subtrees containing S, $\mu_T(S) = |S| + V_1(T)/2 = |T| - V_1(T)/2$. Letting p be any point in S, by (4.5) and (3.9) we have $\mu_T(S) \geqslant \mu_T(p) > M_T$. ∎

The batons and stars show this is asymptotically best possible.

(6.2) Theorem. *If T_n is a sequence of trees such that* $\mathrm{Den}(T_n) \to 1$, *then* $V_2(T_n)/|T_n| \to 1$.

Proof. By the lemma, $V_1(T_n)/|T_n| < 2(1-\mathrm{Den}(T_n))$ so that this ratio goes to 0. It is easy to see, by induction or the standard edge-counting formula, that in any tree T, the number $h(T)$ of nodes of degree >2 is less than the number of endnodes. Thus $h(T_n)/|T_n|$ also goes to 0. Hence the proportion of the nodes of degree 2 must approach 1. ∎

(6.3) Theorem. *If T is a tree of radius $r>1$, then*

$$\mathrm{Den}(T) < 1 - \frac{1}{2r}.$$

Proof. Let p be a point such that each node of T is within distance r of p. The paths from p to the endnodes of T cover T, and each such path has at most r nodes other than p. If these paths have only p in common, then T is astral over p and the required inequality follows from (3.6) and (3.9). Thus we may assume some two of the paths share some node other than p, so $n = |T| \leqslant rV_1(T)$. The required inequality now follows from Lemma 6.1. ∎

(6.4) COROLLARY. *If* (T_n) *is a sequence of trees such that* $\mathrm{Den}(T_n) \to 1$, *then* $\mathrm{diam}\, T_n \to \infty$. ∎

Based on this corollary and the behavior of batons and bridges, one might be tempted to conjecture that $(\mathrm{diam}\, T_n)/|T_n| \to 1$. This is not so, and in fact, we close with an example showing how to choose T_n so that $\mathrm{Den}(T_n) \to 1$ but $(\mathrm{diam}\, T_n)/|T_n| \to 0$.

An (s, t)-*wand* is formed by attaching one endnode of a path on t nodes to the center of a star $K_{1,s}$. An (r, s, t)-*sparkler* is formed by joining r (s, t)-wands together at their handles. Fixing r, taking $t = s^2$, and letting $s \to \infty$, it is not hard to show that the $(2^s)^r$ subtrees containing all r radial t-paths eventually account for nearly all subtrees of the sparkler. Hence the densities of such sparklers go to 1, but the diameter of any such sparkler is only $2t + 5 \leqslant 2n/r$.

7. SOME OPEN PROBLEMS

It seems appropriate to close with some open questions on mean subtree order.

(7.1) Is the tree of maximum density of each order a caterpillar? (Cf. Section 2.)

(7.2) Does a tree with no nodes of degree 2 (i.e., homomorphically irreducible) necessarily have density $\geqslant \frac{1}{2}$?

(7.3) Do nonisomorphic trees of the same order always have different densities? (Let *PnRiCj* denote the *j*th tree in the *i*th row of trees of order n in Harary's list [5]. Trees of *different* orders can have the same density: P6R1C6, P7R1C6, P8R2C2 all have density $\frac{1}{2}$. Also trees of different orders can have the same mean: P10R2C6 and P9R7C7 both have mean 4.75 and there are several trees whose means coincide with those of paths of higher order.)

(7.4) For any tree T and node p of T does $\mu_T(p) \leqslant 2M_T$ necessarily hold? (Cf. Section 3.)

(7.5) For any tree T, is the largest local mean of T always taken on at an endnode?

(7.6) Does every tree other than a path always have a 1-associate with smaller mean subtree order? (Cf. Section 5.)

ACKNOWLEDGMENTS

It is a pleasure to acknowledge David Whited for his computation of the means of all trees of order $\leqslant 10$ and to thank M. S. Irani, J. W. Moon, D. Shier, and J. Straight for their encouragement and helpful conversations.

REFERENCES

1. DE BRUIJN, N. G., KNUTH, D. E., AND RICE, S. O., The average height of planted plane trees, *in* "Graph Theory and Computing," pp. 15–22, Academic Press, New York, 1972.
2. DOYLE, J. K., AND GRAVER, J. E., Mean distance in a graph, *Discrete Math.* **17** (1977), 147–154.
3. EDELMAN, P. H. Meet-distributive lattices and the anti-exchange closure, *Algebra Universalis* **10** (1980), 290–299.
4. GREENE, C., AND KLEITMAN, D. J., Proof techniques in the theory of finite sets, "Studies in Combinatorics" (G. C. Rota, Ed.), MAA Studies 17, 1978.
5. HARARY, F. "Graph Theory," Addison–Wesley Ps, London, 1969.
6. JAMISON, R. E., Monotonicity of mean order in trees, in preparation.
7. JAMISON, R. E., On the Whitney numbers of trees, in preparation.
8. MEIR, A., AND MOON, J. W., The distance between points in random trees, *J. Combin. Theory* **8** (1970), 99–103.
9. MEIR, A., AND MOON, J. W., On the altitude of nodes in random trees, *Canad. J. Math.* **30** (1978), 997–1015.
10. MEIR, A., MOON, J. W., AND POUNDER, J. R., On the order of random channel networks, *SIAM J. Alg. Disc. Meth.* **1** (1980), 25–33.
11. MOON, J. W., private communication.
12. REID, K. B., Cycles in the complement of a tree, *Discrete Math.* **15** (1976), 163–174.
13. RENYI, A., AND SZEKERES, G., On the height of trees, *Austral. J. Math.* **7** (1967), 497–507.
14. ROBINSON, R. W., AND SCHWENK, A. J., The distribution of degrees in a large random tree, *Discrete Math.* **12** (1975), 359–372.
15. RUSKEY, F., Listing and counting subtrees of a tree. *SIAM J. Comput.* **10** (1981), 141–150.
16. STRAIGHT, J., Private communication.