ON UNIVERSAL GRAPHS FOR SPANNING TREES

F. R. K. CHUNG AND R. L. GRAHAM

Introduction

A number of papers [1, 2, 3, 4, 6] recently have been concerned with the following question. What is the minimum number $s(n)$ of edges a graph G on n vertices can have so that any tree on n vertices is isomorphic to some spanning tree of G? We call such a graph *universal for spanning trees*. Since K_n, the complete graph on n vertices (see [5] for terminology), has the required property, it is immediate that

$$(1) \qquad s(n) \leqslant \binom{n}{2} \sim \tfrac{1}{2}n^2 .$$

A previous bound of Nebeský [6] asserted that

$$(2) \qquad s(n) \leqslant \left(1 + o(1)\right)\frac{n^2}{5} .$$

If G is merely required to contain all n-vertex trees as subgraphs (not necessarily spanning) then the corresponding minimum number $s^*(n)$ of edges was shown [4] to satisfy

$$(3) \qquad s^*(n) = O\left(n \log n(\log \log n)^2\right) .$$

On the other hand, a degree constraint argument (see [2]) implies at once that

$$(4) \qquad s(n) \geqslant s^*(n) > \tfrac{1}{2}n \log n .$$

In this note we close the gap between the lower bound of (4) and the upper bounds of (2) and (3) considerably. In particular, we prove that

$$(5) \qquad s(n) \leqslant \frac{5}{\log 4} n \log n + O(n) .$$

Preliminaries

We begin with some notation and definitions. By the *binary tree with k levels*, denoted by $B(k)$, we mean the graph defined as follows. The vertex set $V(k)$ of $B(k)$ is

given by

(6) $V(k) = \{\alpha_1\alpha_2 \ldots \alpha_j : \alpha_i = 0 \text{ or } 1, 1 \leqslant i \leqslant j, 0 \leqslant j \leqslant k\}$

where the point denoted by the expression $\alpha_1\alpha_2 \ldots \alpha_j$ for $j = 0$ is called the *root* of $B(k)$ and is denoted by $*$. The only edges in $B(k)$ are the pairs

$$\{\alpha_1\alpha_2 \ldots \alpha_j, \alpha_1\alpha_2 \ldots \alpha_j 0\} \quad \text{and} \quad \{\alpha_1\alpha_2 \ldots \alpha_j, \alpha_1\alpha_2 \ldots \alpha_j 1\}.$$

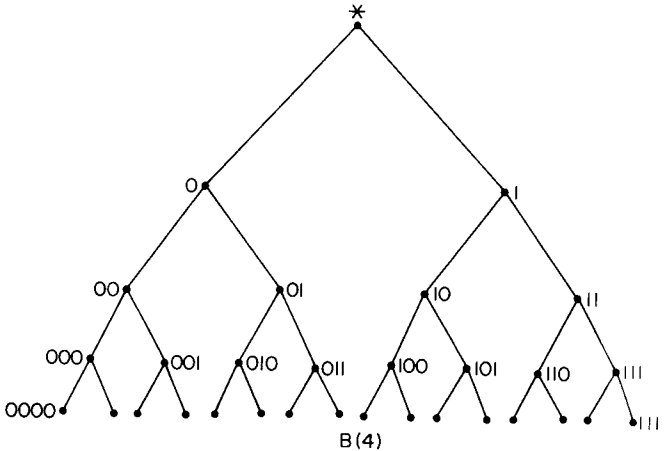In Figure 1 we show an illustration of $B(4)$.

B(4)

FIG. 1

Schematically, we can think of $B(k)$ as being formed from two copies of $B(k-1)$, denoted by $B_0(k-1)$ and $B_1(k-1)$, joined to the root $*$ as shown in Figure 2. Note that $B(k)$ has $2^{k+1} - 1$ vertices.
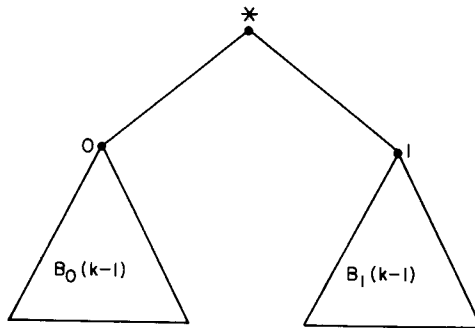

B(k)

FIG. 2

If $\alpha = \alpha_1\alpha_2 \ldots \alpha_j \in V(k)$, $j < k$, the two vertices $\alpha 0 = \alpha_1\alpha_2 \ldots \alpha_j 0$ and $\alpha 1 = \alpha_1\alpha_2 \ldots \alpha_j 1$ are called the *left* and *right* sons of α, respectively. Also, α is called the *father* of $\alpha 0$ and $\alpha 1$, and $\alpha 0$ is called the *left-hand brother* of $\alpha 1$.

We next come to the most important definition in the paper. A subtree A of $B(k)$ is said to be *admissible* if one of the following holds:

(i) A is empty;

(ii)   *A* consists of the single vertex *;

(iii) *A* has the form shown in either Figure 3(a) or 3(b) where *A'* is an admissible subtree of $B(k-1)$ attached at its root to the indicated vertex (that is, 0 in (a) and 1 in (b)).
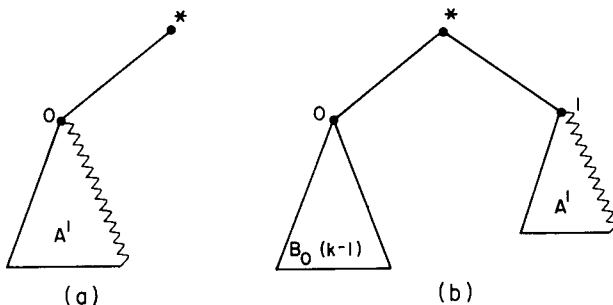

FIG. 3

If β is a vertex of *A*, we let $v_A(\beta)$ denote the total number of desdendants of β in *A*, that is, sons of β, sons of sons of β, and so forth.

We next form the graph $G(k)$ on the vertex set $V(k)$ as follows. For each vertex $\alpha \in V(k)$, there are edges between α and

(i)   every descendant of α,

(ii)   the left-hand brother of α (if α has one) and its descendants,

(iii) the left-hand brother of the father of α, called β (if it exists), and all of the descendants of β.

Finally, a subgraph $G \subseteq G(k)$ is said to be *admissible* if it is an induced subgraph of $G(k)$ on the set of vertices of some admissible subtree of $B(k)$.

What we shall prove is that every admissible graph is universal for spanning trees. In fact, we shall prove by induction on $|V(G)|$ the following stronger statement.

(∗)   *For any tree T with $|V(T)| \leqslant |V(G)|$ and for any $v \in V(T)$ there is an embedding $\lambda: V(T) \to V(G)$ such that the subgraph induced by $V(G) - \lambda(V(T))$ is admissible and*

$$v_G(\lambda(v)) < |V(T)| \leqslant v_G(\lambda(v)^*),$$

where $\lambda(v)^*$ denotes the father of $\lambda(v)$ (if it exists).

Before beginning the proof of (∗) we first mention a lemma which will be needed several times during our proof. Given a tree *T* and a vertex *u* of *T*, by a *u-component* of *T* we mean any component formed by removing *u* from *T*. The following result follows at once from the related theorem in [2].

LEMMA.   *Let T be a tree with at least $k+1$ vertices. Then for some vertex u, there is a set of u-components C_i of T, $1 \leqslant i \leqslant t$, such that the numbers of vertices $|V(C_i)|$ in the C_i satisfy*

(7) $$k \leqslant \sum_{i=1}^{t} |V(C_i)| < 2k.$$

The statement (∗) clearly holds if $|V(G)| = 1$. We therefore assume that G is a fixed admissible graph with $|V(G)| > 1$, and that (∗) holds for all admissible graphs with at most $|V(G)| - 1$ vertices. The assertion also holds when T has one vertex. Thus, assume that T is a given tree with $|V(T)| > 1$, that v is a fixed vertex of T and suppose that (∗) holds for G with all trees having at most $|V(T)| - 1$ vertices. There are now a number of cases to consider.

I.   Suppose that G has the form shown in Figure 4, where G_1 is isomorphic to an admissible subgraph of $G(k-1)$ (with 0 playing the role of ∗).
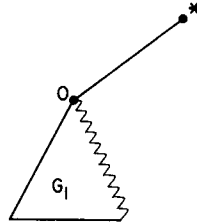

Fig. 4

Case (i), *when* $|V(T)| < |V(G)|$.   Since G_1 is admissible and $|V(T)| \le |V(G_1)|$ it follows by induction that there is an embedding $\lambda : T \to G_1$ such that $G_1 - \lambda(T)$ is admissible and

$$v_{G_1}(\lambda(v)) < |V(T)| \le v_{G_1}(\lambda(v)^*).$$

By the definition of admissibility, $G - \lambda(T)$ is also admissible. Since $v_{G_1}(x) = v_G(x)$ for $x \in V(G_1)$ it follows that λ satisfies all the requirements of (∗).

Case (ii), *when* $|V(T)| = |V(G)|$.   Form the tree T' by removing the fixed vertex v from T and connecting up components with additional edges, if necessary. As before, by induction, there is an embedding $\lambda' : T' \to G_1$. Extend λ' to an embedding of T into G by defining $\lambda'(v) = ∗$. Since

$$v_G(∗) < |V(T)|$$

it follows that λ' satisfies the conditions of (∗).

II.   Suppose that G has the form shown in Figure 5, where G_1 is isomorphic to an admissible subgraph of $G(k-1)$.
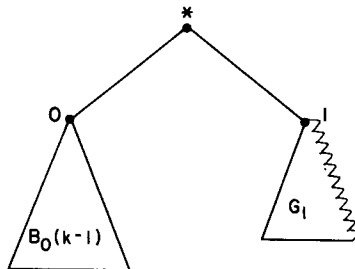

Fig. 5

Case (i), *when* $|V(T)| \le |V(G_1)|$.   In this case an argument identical to that in I shows that the desired embedding exists.

Case (ii), *when* $|V(G_1)| < |V(T)|$.   There are now several subcases to consider.

Subcase (a), *when* G_1 *consists of a single vertex* $\{1\}$.   Form the tree T' by removing v from T and reconnecting the components if necessary. By induction, since $G - G_1 = G - \{1\}$ is admissible, there exists an embedding $\lambda' : T' \to G - \{1\}$ satisfying (∗). Extend λ' to T by defining $\lambda'(v) = 1$. Then it is easily checked that $\lambda' : T \to G$ satisfies (∗).

Subcase (b), *when* G_1 *consists of more than one vertex but* 1 *has only one son* 10 (*see Figure* 6).   There are two cases.
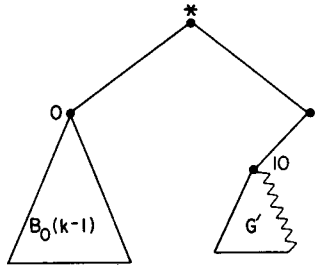

FIG. 6

1.   Suppose that $|V(T)| < |V(G)|$. In the tree T' formed from $T - \{v\}$ by connecting up components if necessary, choose a vertex v_1 such that there exists a collection of v_1-components $C_1, ..., C_t$ having

$$(8) \qquad v_G(10) \leqslant \sum_{i=1}^{t} |V(C_i)| < 2v_G(10).$$

This is possible (by the Lemma) by the assumption of (ii). We can consider G to have the form shown in Figure 7 (by the way the edges in $G(k)$ are defined). The circled
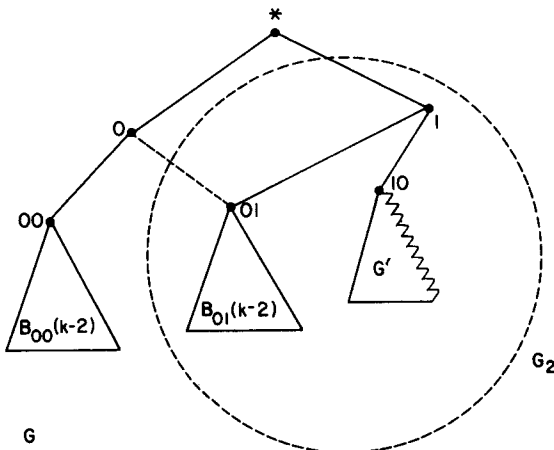

FIG. 7

portion, denoted by G_2, is (isomorphic to) an admissible graph. Hence, by the

induction hypothesis, there is an embedding

$$\lambda' : \left(\bigcup_{i=1}^{t} C_i \cup \{v_1\} \right) \longrightarrow G_2$$

such that

$$G_3 = G_2 - \lambda' \left(\bigcup_{i=1}^{t} C_i \cup \{v_1\} \right)$$

is admissible and $\lambda'(v_1) = 10$ (by (8)). But $G - G_2 + G_3 - \{1\} = G_4$ can be written as shown in Figure 8. Again by induction, there is an embedding

$$\lambda'' : \left(T - \bigcup_{i=1}^{t} C_i - \{v\} - \{v_1\} \right) \longrightarrow G_4$$
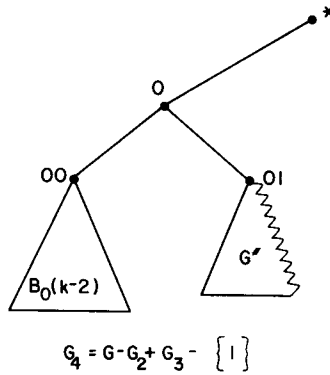

$$G_4 = G - G_2 + G_3 - \{1\}$$

FIG. 8

with $G_4 - \lambda'' \left(T - \bigcup_{i=1}^{t} C_i - \{v\} - \{v_1\} \right)$ admissible. Finally, we form an embedding $\lambda : T \to G$ by defining

$$\lambda(x) = \begin{cases} \lambda'(x) & \text{for } x \in \bigcup_{i=1}^{t} C_i \cup \{v_1\}, \\ \lambda''(x) & \text{for } x \in T - \bigcup_{i=1}^{t} C_i - \{v_1\} - \{v\}, \\ 1 & \text{for } x = v \end{cases}$$

(since 10 is connected to every vertex in $G(k)$). Certainly $G - \lambda(T)$ is admissible. Also v is mapped to the proper place by λ in order to satisfy (∗), so that this subcase is finished.

2.   Suppose that $|V(T)| = |V(G)|$. In this case the desired embedding must map v to ∗. The argument in this subcase is very similar to that in the preceding subcase (since 1 and ∗ are connected to exactly the same sets of vertices in G), and will not be given.

Subcase (c), *when* G_1 *consists of more than one vertex and* 1 *has two sons* (*see Figure* 9).
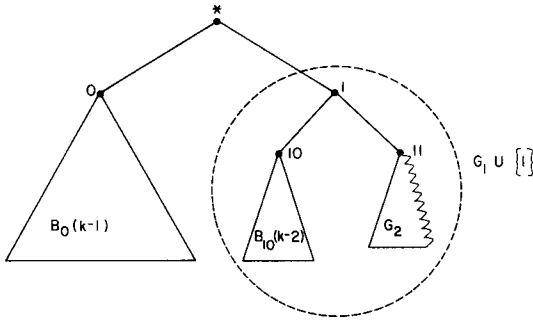

FIG. 9

1.   Suppose that $|V(T)| < |V(G)|$. Then the desired embedding will map v to 1. In $T - \{v\}$, choose v_1 such that for some collection of v_1-components $C_1, ..., C_t$,

$$(9) \qquad v_G(11) \leqslant \sum_{i=1}^{t} |V(C_i)| < 2v_G(11).$$

By induction, $T_1 = \bigcup_{i=1}^{t} C_i \cup \{v_1\}$ can be embedded by some mapping λ_1 into $G_1 \cup \{1\}$ so that $G_1 \cup \{1\} - \lambda_1(T_1)$ is admissible and $\lambda_1(v_1) = 11$ (by (9)). Consider $G - \lambda_1(T_1)$ written as shown in Figure 10. The situation is essentially the same as in
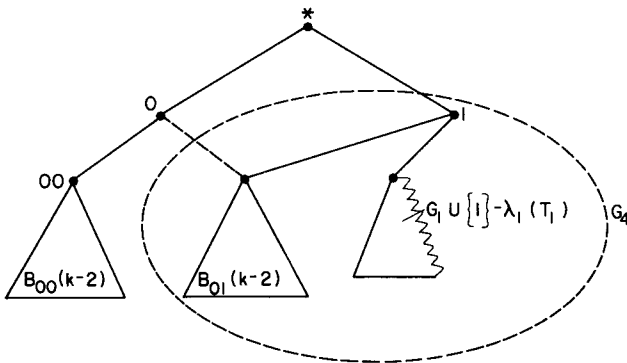

FIG. 10

(b)1, Figure 7. The argument used there now applies here with the eventual conclusion that there is an embedding $\lambda : T \to G$ satisfying (∗) (in particular, with $\lambda(v) = 1$).

2.   Suppose that $|V(T)| = |V(G)|$. The only difference in the argument for this case from that in (c)1 is that now $\lambda(v) = *$ is required. However, as in (b)1 and (b)2, there is no essential difference and the subcase can be dealt with without difficulty.

This completes the induction step and (∗) is proved.

The next step in the proof of (5) is to estimate the number of edges in an admissible $G \subseteq G(k)$ with n vertices. First, we count the number of edges in $G(k)$. The i-th level of $G(k)$ has 2^i vertices. It follows from the way that $G(k)$ is defined that the average number of edges going down from an i-th level vertex is at most $\frac{5}{2} \cdot 2^{k+1-i}$. Hence, the total number of edges in $G(k)$ is bounded above by

$$\sum_{i=0}^{k-1} \frac{5}{2} \cdot 2^{k+1-i} \cdot 2^i = 5k \cdot 2^k.$$

For arbitrary n, we form an admissible subgraph G_n of $G(k)$, where

$$2^k - 1 < n \leqslant 2^{k+1} - 1,$$

as shown in Figure 11. More precisely, n is written as a sum
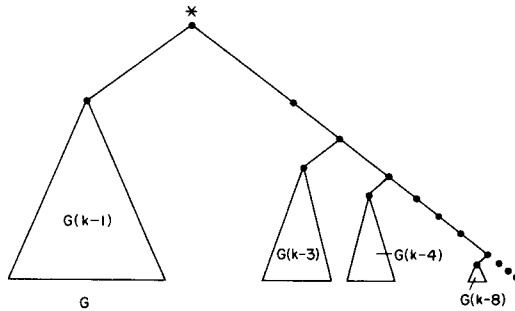
$$n = \sum_{i=0}^{k} d_i$$

FIG. 11

where $d_i = 2^i$ or 1. It is easy to see that the number of edges $e(G_n)$ of G_n satisfies

(10) $$e(G_n) \leqslant \sum_{i=0}^{k} e(G(i))\varepsilon_i + O(n)$$

where $\varepsilon_i = 1$ if $d_i = 2^i$ and 0 otherwise. Thus,

$$e(G_n) \leqslant \sum_{i=0}^{k} 5i \cdot 2^i \varepsilon_i + O(n) \leqslant \frac{5 \log n}{\log 2} \sum_{i=0}^{k} 2^i \varepsilon_i + O(n)$$

$$\leqslant \frac{5}{\log 4} n \log n + O(n).$$

This completes the proof of (5).

Concluding remarks

It may be possible with a little more care to squeeze the coefficient

$5/\log 4 = 3.6067\ldots$ down even further. Conceivably, the right value might even be $\frac{1}{2}$ although at present we certainly do not see how to prove this.

References

1. L. BABAI, F. R. K. CHUNG, P. ERDÖS, R. L. GRAHAM and J. H. SPENCER, 'On graphs which contain all sparse graphs', *Ann. Discrete Math.*, 12 (1982), 21–26.
2. F. R. K. CHUNG and R. L. GRAHAM, 'On graphs which contain all small trees', *J. Combin. Theory Ser. B*, 24 (1978), 14–23.
3. F. R. K. CHUNG and R. L. GRAHAM, 'On universal graphs', *Ann. New York Acad. Sci.*, 319 (1979), 136–140.
4. F. R. K. CHUNG, R. L. GRAHAM and N. PIPPENGER, 'On small graphs which contain all small trees II', *Proceedings, 1976 Hungarian colloquium on combinatorics* (North-Holland, Amsterdam, 1978), pp. 213–223.
5. F. HARARY, *Graph theory* (Addison Wesley, New York, 1969).
6. L. NEBESKÝ, 'On tree-complete graphs', *Časopis Pěst. Mat.*, 100 (1975), 334–338.

Bell Laboratories,
Murray Hill,
New Jersey 07974,
U.S.A.