ON MAXIMAL TRANSITIVE SUBTOURNAMENTS

by J. W. MOON†

(Received 11th March 1971)

1. Introduction

A *tournament* T_n consists of a finite set of *nodes* 1, 2, ..., n such that each pair of distinct nodes i and j is joined by exactly one of the *arcs* ij or ji. If the arc ij is in T_n we say that i *beats* j or j *loses* to i and write $i \to j$. If each node of a subtournament A beats each node of a subtournament B we write $A \to B$ and let $A+B$ denote the tournament determined by the nodes of A and B.

A tournament T_n is *transitive* if its nodes can be labelled in such a way that $i \to j$ if and only if $i > j$; in this case we call node n the *top* node. A transitive subtournament of a tournament T_n is *maximal* if it is not a proper subtournament of any other transitive subtournament of T_n. Let $f(n)$ denote the maximum number of maximal transitive subtournaments a tournament T_n can have; we find by inspection, for example, that $f(1) = f(2) = 1$ and $f(3) = f(4) = 3$. Our object here is to prove the following result.

Theorem. *If* $n = 5m + r$, *where* $m \geqq 1$ *and* $0 \leqq r \leqq 4$, *then*

$$c_r 7^{m-1} \leqq f(n) \leqq (1{\cdot}717)^n$$

where $c_0 = 7$, $c_1 = 9$, $c_2 = 15$, $c_3 = 19$, *and* $c_4 = 31$.

Corollary. *If* $\theta = \lim_{n \to \infty} (f(n))^{1/n}$, *then* θ *exists and* $1{\cdot}4757 \leqq \theta \leqq 1{\cdot}717$.

2. A lower bound for $f(n)$

A tournament T_n is *strong* if it cannot be expressed as $T_n = A+B$ for some nonempty tournaments A and B. If T_n is not strong it has a unique expression of the type $T_n = A+B+\ldots+K$ where the non-empty tournaments $A, B, \ldots, K$ all are strong; if this is the case, then $f(T_n) = f(A)f(B)\ldots f(K)$ where $f(X)$ denotes the number of maximal transitive subtournaments in the tournament X. It follows, therefore, that if $a+b = n$ then

$$f(n) \geqq f(a)f(b). \tag{1}$$

If T_n is any tournament with n nodes, let T_{n+2} denote the tournament obtained by adjoining two nodes p and q to T_n such that $p \to T_n$, $T_n \to q$, and $q \to p$. It is not difficult to see that $f(T_{n+2}) = 2f(T_n)+1$; consequently,

$$f(n+2) \geqq 2f(n)+1. \tag{2}$$

† On leave from the University of Alberta.

https://doi.org/10.1017/S0013091500009639 Published online by Cambridge University Press

Since $f(3) = 3$, it follows that $f(5) \geqq 7, f(7) \geqq 15$, and $f(9) \geqq 31$; furthermore, $f(6) \geqq (f(3))^2 = 9$, by (1), whence $f(8) \geqq 19$ by (2).

Let us now suppose that $n = 5m+r$ where $m \geqq 2$ and $0 \leqq r \leqq 4$. Then

$$f(n) \geqq f(5(m-1))f(5+r) \geqq c_r(f(5))^{m-1} \geqq c_r 7^{m-1},$$

by (1) and the results in the preceding paragraph. We remark that the existence of the limit in the corollary follows from inequality (1) and a well-known result on sub-additive functions (**2**, Problem 98, pp. 17, 171).

3. An upper bound for $f(n)$: a special case

We shall prove that $f(T_n) \leqq \beta^n$, where $\beta = 1{\cdot}717$, by induction on n. The inequality certainly holds when $1 \leqq n \leqq 4$ and we may restrict our attention to strong tournaments T_n in view of the observation made earlier.

The *score* of a node i in a tournament is the number s_i of nodes that i beats. If x is the top node of any maximal transitive subtournament M of a tournament T_n, let N denote the tournament obtained from M by deleting x (M must have at least two nodes when $n \geqq 2$). It is easy to verify that N is a maximal transitive subtournament of the tournament determined by the s_x nodes of T_n that lose to x; thus x is the top node of at most $f(s_x)$ maximal transitive subtournaments of T_n. It follows, therefore, that if $(s_1, \ldots, s_n)$ denotes the score sequence of a tournament T_n, then

$$f(T_n) \leqq \sum_{i=1}^{n} f(s_i). \tag{3}$$

If T_n is strong then $s_i \leqq n-2$ for every node i. In this section we treat the case where there exists a node p in T_n such that $s_p = n-2$. Let q denote the unique node of T_n that beats p. If two nodes of T_n have score $n-2$, then we may take p and q to be these nodes. If three nodes had score $n-2$ then these nodes would beat all the remaining nodes when $n>3$ and T_n would not be strong. Thus we may suppose that $s_i \leqq n-3$ for any node i of T_n other than p or q. Let T_a and T_b denote the subtournaments determined by those nodes of T_n other than p that beat q and lose to q, respectively. Since $s_q \leqq n-2$ it must be that $a \geqq 1$ and $b \leqq n-3$.

Node q is the only node that beats p. It follows that every maximal transitive subtournament of T_n that does not contain q must certainly contain p. It is not difficult to see that there are at most $f(n-2)$ such subtournaments. A maximal transitive subtournament that contains both p and q cannot contain any nodes of T_a, for it would not be transitive otherwise. There are at most $f(b)$ such subtournaments (we adopt the convention that $f(0) = 1$).

We now consider those maximal transitive subtournaments of T_n that contain q but not p. There are certainly no more than $f(n-1)$ such subtournaments in general. We can obtain a sharper bound when $a \leqq 2$ by observing that the top node of such subtournaments must belong to T_a; if this were not the case then they would contain no nodes of T_a and node p could be adjoined without destroying the transitivity properly. Hence, there are at most $af(n-3)$ such subtournaments.

We may suppose that $f(m)<\beta^m$ if $m<n$ and that $f(m)$ is an increasing function. It follows, therefore, that if T_n is strong and has a node of score $n-2$, then

$$f(T_n) \leqq \begin{cases} 2f(n-3)+f(n-2) & \leqq (2+\beta)\beta^{n-3}, \\ f(n-4)+2f(n-3)+f(n-2) & \leqq (1+\beta)^2\beta^{n-4}, \\ f(n-5)+f(n-2)+f(n-1) & \leqq (1+\beta^3+\beta^4)\beta^{n-5}, \end{cases}$$

according as $b = n-3$, $b = n-4$, or $b \leqq n-5$. Each of the last three bounds is smaller than β^n when $\beta = 1{\cdot}717$ and a similar argument shows that $f(T_n) \leqq \beta^n$ when T_n is strong and has a node of score 1. (Notice that the first three bounds imply that $f(T_5) \leqq 7$ when T_5 has a node of score 1 or 3; we shall use this inequality in the next section.)

4. An upper bound for $f(n)$: the general case

Let T_n denote a strong tournament with no nodes of score 1 or $n-2$; if the nodes of T_n are labelled so that the sequence $s = (s_1, \ldots, s_n)$ of scores is non-decreasing, then it follows from our assumptions that

$$2 \leqq s_1 \leqq \ldots \leqq s_n \leqq n-3, \tag{4}$$

$$\sum_{i=1}^{k} s_i \geqq \binom{k}{2}+1, \text{ for } 1 \leqq k \leqq n-1, \text{ and} \tag{5}$$

$$\sum_{i=1}^{n} s_i = \binom{n}{2}. \tag{6}$$

We remark that for any sequence s satisfying these conditions there exists at least one tournament with score sequence s; this is a consequence of a theorem due to Landau (see (**1**; p. 61)).

We first treat the cases where $5 \leqq n \leqq 8$. It follows from inequality (3) that $f(T_n) \leqq nf(n-3)$ for tournaments T_n whose scores satisfy conditions (4)-(6). Consequently, $f(T_5) < 5{\cdot}1$, $f(T_6) \leqq 6{\cdot}3$, $f(T_7) \leqq 7{\cdot}3$, and $f(T_8) \leqq 8{\cdot}7$ for such tournaments when $5 \leqq n \leqq 8$. All these bounds are less than β^n for the appropriate values of n, so we may now suppose that $n \geqq 9$ and that $f(m) \leqq \beta^m$ for $m<n$.

If $n \geqq 9$, let S_n denote the set of all sequences $s = (s_1, \ldots, s_n)$ of n integers that satisfy conditions (4)-(6). Let s^* denote the sequence (2, 2, 2, 2, 4, 6, 6, 6, 6) or (2, 2, 2, 2, 3, 5, 6, ..., $n-6$, $n-4$, $n-3$, $n-3$, $n-3$, $n-3$) according as $n = 9$ or $n \geqq 10$; it is easy to verify that $s^* \in S_n$. If $g(x) = \beta^x$ let

$$G(s) = g(s_1)+\ldots+g(s_n)$$

for any s in S_n. We shall prove the following result in the next section.

Lemma. *If $s \in S_n$ where $n \geqq 9$, then $G(s) \leqq G(s^*)$.*

It follows from inequality (3), the induction hypothesis, and the lemma, that

$$f(T_9) \leqq G(s) \leqq G(s^*) = 4\beta^2 + \beta^4 + 4\beta^6 \leqq \beta^9$$

and, in general, that

$$f(T_n) \leqq G(s) \leqq G(s^*) = 4\beta^2 + \beta^3 + (\beta^{n-5} - \beta^5)/(\beta - 1) + \beta^{n-4} + 4\beta^{n-3}$$
$$\leqq \beta^{n-5}/(\beta - 1) + \beta^{n-4} + 4\beta^{n-3} \leqq \beta^n$$

when $\beta = 1{\cdot}717$. This will suffice to complete the proof of the theorem. (We remark that it is easy to show that if T_n is strong and $n \geqq 3$ then the minimum value $f(T_n)$ can have is 3.)

5. Proof of the lemma

Let s denote any sequence in S_n such that $G(s) = \max\{G(t)\colon t \in S_n\}$; such a sequence certainly exists. We shall prove the lemma by establishing a series of assertions (the only property of the function $g(x)$ that we use is that it is strictly convex).

Assertion 1. *If two consecutive elements of s are equal, they must equal* 2 *or* $n-3$.

Suppose there exist integers u and v, where $1 \leqq u < v \leqq n$, such that

$$2 < s_u = \ldots = s_v < n-3;$$

we may suppose that $s_{u-1} < s_u$ if $u > 1$ and $s_v < s_{v+1}$ if $v < n$. Let $r = (r_1, \ldots, r_n)$ denote the sequence that differs from s only in that $r_u = s_u - 1$ and $r_v = s_v + 1$. The sequence r certainly satisfies conditions (4) and (6); it satisfies condition (5) as well unless there exists an integer k, where $u \leqq k < v$, such that

$$\sum_{i=1}^{k} s_i = \binom{k}{2} + 1. \tag{7}$$

If $k = 1$ then $s_1 = 1$ and if $k = n-1$ then $s_n = n-2$ by (6); both these alternatives contradict condition (4) so we may suppose $2 \leqq k \leqq n-2$.

It follows from (7) and condition (5) that

$$s_k = \sum_{i=1}^{k} s_i - \sum_{i=1}^{k-1} s_i \leqq \binom{k}{2} + 1 - \binom{k-1}{2} - 1 = k-1.$$

Furthermore, $s_{k+1} = s_k$ so

$$\sum_{i=1}^{k+1} s_i \leqq \binom{k}{2} + 1 + (k-1) = \binom{k+1}{2}.$$

This contradicts condition (5) so it must be that the assumption that equation (7) holds is incorrect. Consequently, if assertion 1 does not hold then the sequence r is in S_n. Since g is a strictly convex function it follows that

$$G(r) - G(s) = (g(x+1) - g(x)) - (g(x) - g(x-1)) > 0$$

where $x = s_u = s_v$. This contradicts the definition of s so it must be that assertion 1 does in fact hold.

Assertion 2. *The sequence s must start with between two and four 2's and end with between two and four $(n-3)$'s.*

There are at most $(n-4)-2 = n-6$ s_i's between 3 and $n-4$, inclusive, by Assertion 1; thus there are at least six s_i's equal to 2 or $n-3$. It is easy to verify that s would not satisfy condition (5) if it started with more than four 2's or ended with more than four $(n-3)$'s. Assertion 2 now follows.

Assertion 3. *The sequence s must start with four 2's and end with four $(n-3)$'s.*

If s does not start with four 2's then $2 = s_1 = \ldots = s_{a-1} < s_a$ where $a = 3$ or 4, by Assertion 2. Now $a+1 < n-3$, since $n \geqq 9$, so $s_a < s_{a+1} < s_{a+2}$, by Assertions 1 and 2. Let $r = (r_1, \ldots, r_n)$ denote the sequence that differs from s only in that $r_a = s_a - 1$ and $r_{a+1} = s_{a+1}+1$. It is easy to verify that r is in S_n in this case. However,

$$G(r)-G(s) = (g(y+1)-g(y))-(g(x)-g(x-1))>0$$

where $x = s_a < s_{a+1} = y$. This contradiction implies that s must start with four 2's and the last part of the assertion can be proved in a similar way.

Assertion 4. $s = s^*$.

It follows from Assertions 1 and 3 that the middle $n-8$ elements of s consist of $n-8$ of the $n-6$ numbers 3, 4, ..., $n-4$. Condition (6) implies that if h is one of the missing numbers then $n-1-h$ is the other, where we may suppose that $3 \leqq h < \frac{1}{2}(n-1)$. If $h \geqq 5$ then s does not satisfy condition (5) when $k = 6$ so $h = 3$ or 4. When $n = 9$ the only possibility is that $h = 3$; when $n \geqq 10$ both values are possible but the function G has the larger value for the sequence corresponding to $h = 4$. It follows, therefore, that $s = s^*$ and the lemma is proved.

The preparation of this paper was assisted by a grant from the National Research Council of Canada.

REFERENCES

(1) J. W. Moon, *Topics on Tournaments* (Holt, New York, 1968).

(2) G. Pólya u. G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, I (Springer, Berlin, 1925).

Mathematical Institutf
Oxford