

The minimum number of spanning paths in a strong tournament

By J. W. MOON (Cape Town)

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

A path is a sequence $P = \{p_1, p_2, \dots, p_k\}$ of distinct nodes of T_n such that $p_i \rightarrow p_{i+1}$ if $1 \leq i < k$; a spanning path contains every node of T_n . Let $h(T_n)$ denote the number of spanning paths in the tournament T_n . RÉDEI [1] showed that $h(T_n)$ is odd for every T_n and SZELE [2] showed that if $H(n)$ denotes the maximum of $h(T_n)$ over all tournaments with n nodes then

$$n!/2^{n-1} \leq H(n) \leq (n+1)!/2^{(3/4)n-3} \quad \text{for } n \geq 1.$$

A tournament T_n is strong if it cannot be expressed as $T_n = A+B$ for some nonempty subtournaments A and B . If a tournament T_n is not strong, or weak, it has a unique expression of the type $T_n = A+B+\dots+K$ where the nonempty subtournaments $A, B, \dots, K$ all are strong and $h(T_n) = h(A) \cdot h(B) \dots h(K)$. Thus in considering lower bounds for $h(T_n)$ we may as well assume T_n is strong. Our object here is to prove the following result.

Theorem. If $h(n)$ denotes the minimum number of spanning paths a strong tournament T_n can have, then $h(1)=1$ and

$$(1) \quad \alpha^{n-1} \leq h(n) \leq \begin{cases} \beta^{n-1} & n \equiv 1 \pmod{3} \\ 9 \cdot 5^{-4/3} \beta^{n-1} & \text{if } n \equiv 2 \pmod{3} \\ 3 \cdot 5^{-2/3} \beta^{n-1} & n \equiv 3 \pmod{3} \end{cases}$$

for $n \geq 3$, where $\alpha = 6^{1/4} = 1.565$ and $\beta = 5^{1/3} = 1.710$.

2. An upper bound for $h(n)$. If $n \geq 3$ let R_n denote the strong tournament in which $i \rightarrow j$ if and only if $i < j$ except that $n \rightarrow 1$. Any spanning path P of R_n that involves the arc $\vec{n1}$ partitions the nodes $(2, 3, \dots, n-1)$ into two subsets, those that come before n and those that come after 1 ; it follows from the definition of R_n that the nodes in these two subsets must occur in natural order in the path P . Conversely, each partition of the nodes $(2, 3, \dots, n-1)$ into two subsets determines a unique spanning path that involves the arc $\vec{n1}$. There is only one spanning path of R_n that doesn't

involve the arc $\vec{n1}$, namely, $\{1, 2, \dots, n\}$. Consequently, $h(R_n) = 1 + 2^{n-2}$ for $n \geq 3$.

More generally, let $\{s_1, s_2, \dots, s_t\}$ denote an increasing subsequence of the nodes $1, 2, \dots, n$ such that $s_1 = 1$, $s_t = n$, and $\alpha_i = s_{i+1} - s_i - 1 \geq 1$ for $1 \leq i < t$. Let S_n denote the strong tournament in which $i < j$ if and only if $i < j$ except that $s_{i+1} \rightarrow s_i$ for $1 \leq i < t$. It is not difficult to show, by an extension of the argument just used for the case $t = 2$, that

$$h(S_n) = \prod_{i=1}^{t-1} (1 + 2^{\alpha_i})$$

for $n \geq 3$. (Notice that $h(S_n)$ is always odd, in accordance with Rédei's theorem.)

To minimize $h(S_n)$ for given values of n , notice that

$$(1 + 2^{\alpha+3}) > (1 + 2^2)(1 + 2^\alpha) \quad \text{for } \alpha \geq 1,$$

$$(1 + 2^3)^2 > (1 + 2^1)(1 + 2^2)^2,$$

and

$$(1 + 2^1)(1 + 2^3) = (1 + 2^1)^3 > (1 + 2^2)^2.$$

Thus if $n = 3k + 1$ the minimum occurs when $t = k + 1$ and all the α 's equal two; if $n = 3k + 2$ it occurs when $t = k + 1$ and one α equals three and the rest equal two or when $t = k + 2$ and two α 's equal one and the rest equal two; if $n = 3k + 3$ it occurs when $t = k + 2$ and one α equals one and the rest equal two. It follows, therefore, that the minimum value of $h(S_n)$ is 5^k , $9 \cdot 5^{k-1}$, or $3 \cdot 5^k$ according as $n = 3k + 1$, $3k + 2$, or $3k + 3$. This implies the upper bound in inequality (1).

3. A lower bound for $h(n)$. Before proceeding we remark that if node y is not contained in a path P and y beats some node x of P then y can be inserted in the path P before x ; in particular, y can be inserted immediately before the first node of P it beats. Similarly, if y loses to x then y can be inserted in P after x .

It is not difficult to verify that $h(1) = 1$, $h(3) = 3$, $h(4) = 5$, and $h(5) = 9$ so the lower bound in inequality (1) certainly holds for small values of n . Suppose the inequality has been established for $3 \leq n \leq m - 1$ and consider a strong tournament T_m such that $h(m) = h(T_m)$.

Let K denote any minimal subset of nodes of T_m whose removal, along with all incident arcs, leaves a weak subtournament R ; such a subset K exists since, for example the removal of the nodes that lose to any given node of T_m leaves a weak subtournament. Since R is weak there exist subtournaments A , B and C where A and C are strong such that $R = A + B + C$; it may be that B is vacuous. Let a , b , c and k denote the number of nodes in A , B , C and K so that $a + b + c + k = m$. It follows from the minimality property of K and the fact that T_m is strong that each node of K loses to at least one node of C and beats at least one node of A . We now construct two families of spanning paths of T_m and thus obtain a lower bound for $h(m)$.

Let P_1 , P_2 , and P_3 denote any spanning paths of A , B , and C ; there are at least $h(a)h(c)$ spanning paths of R and they all are of the type $P_1 + P_2 + P_3$. Now let X and Y denote any partition of the nodes of K into two subsets. Since each node of X beats some node of A and each node of Y loses to some node of C it follows that the nodes of X and Y can be inserted in the path $P_1 + P_2 + P_3$ so as to form a spanning path Q of T_m that first passes through the nodes of X and A ,

then passes through the nodes of B , and finally passes through the nodes of Y and C . Different paths P_1 and P_3 and different partitions X and Y yield different paths Q so T_m must have at least $2^k h(a)h(c)$ spanning paths of this first type.

Let x denote any given node of K and let u and v denote given nodes of A and C that lose to x and beat x , respectively. For any spanning path P_1 of A let P_{11} denote the (possibly empty) subpath determined by the nodes that come before u in P_1 and let P_{12} denote the subpath determined by u and the nodes that come after u in P_1 ; for any spanning path P_3 of C let P_{31} denote the subpath determined by v and the nodes that come before v in P_3 and let P_{32} denote the (possibly empty) subpath determined by the nodes that come after v in P_3 ; for any partition of the nodes of B into two subsets W and Z , let P_{21} and P_{22} denote spanning paths of the subtournaments determined by W and Z . It is not difficult to see that for each partition W and Z there are at least $h(a)h(c)$ paths of the type $P^* = P_{11} + P_{21} + P_{31} + \{x\} + P_{12} + P_{22} + P_{32}$ that pass through x and the nodes of R . We now estimate the number of ways of inserting the remaining nodes of K into such a path P^* .

If $k \geq 2$, let y denote any node of $K - \{x\}$. Since y beats some node of A it follows that y can be inserted in P^* before the last node of P_{12} ; it can also be inserted after the last node of P_{12} unless it beats that node and all the nodes of P_{22} and P_{32} . Since y loses to some node of C it follows that y can be inserted in P^* after the first node of P_{31} ; it can also be inserted before the first node of P_{31} unless it loses to that node and to all nodes of P_{11} and P_{21} . Thus if y can't be inserted in P^* in at least two ways it must be that $y \rightarrow Z$ and $W \rightarrow y$, among other things. Notice that this last statement also applies when P^* is replaced by any path obtained by inserting other nodes of $K - \{x\}$ in P^* .

Let the partitions of B into two subsets be numbered from $i=1$ to $i=t=2^b$. For the i -th partition, W_i and Z_i , let γ_i denote the number of nodes y of $K - \{x\}$ such that $y \rightarrow Z_i$ and $W_i \rightarrow y$. It follows from the observations in the last paragraph that there are at least $2^{k-1-\gamma_i}$ ways to insert the nodes of $K - \{x\}$ in any path P^* associated with the partition W_i and Z_i so as to form a spanning path of T_m . If we apply this construction for all spanning paths of A and C and for all partitions of B we obtain at least

$$(2) \quad 2^{k-1} (2^{-\gamma_1} + \dots + 2^{-\gamma_t}) h(a)h(c)$$

different spanning paths of T_m . (Some node of C comes before some node of A in each of these paths so they are not the same as any of the paths enumerated earlier.) Since the nodes of B that beat and lose to any node y of $K - \{x\}$ are fixed, each such node y contributes one to just one γ_i and $\gamma_1 + \dots + \gamma_t = k-1$. Therefore, the sum in (2) is minimized when all the γ_i 's equal $(k-1)2^{-b}$ and T_m must have at least $h(a)h(c)2^{b+(k-1)(1-2^{-b})}$ spanning paths of this second type.

It follows from these two constructions that $h(m) \geq (2^k + 2^{b+(k-1)(1-2^{-b})}) h(a)h(c)$. In order to show that this implies that $h(m) \geq \alpha^{m-1} = \alpha^{a+b+c+k-1}$, it suffices to show, in view of the induction hypothesis, that

$$(3) \quad 2^k + 2^{b+(k-1)(1-2^{-b})} \geq \alpha^{b+k+1} = 6^{(b+k+1)/4}$$

for all integers b and k such that $b \geq 0$ and $k \geq 1$.

Since $\log_2 \alpha < \frac{2}{3}$, inequality (3) certainly holds if $k \geq 2(b+k+1)/3$ or $k \geq 2b+2$ and it can easily be verified directly for $1 \leq k \leq 2b+1$ when $0 \leq b \leq 2$. (Notice that equality holds in (3) when $k=1$ and $b=2$.) If $b \geq 3$, then inequality (3) holds if

$$(4) \quad 2^k + 2^{b+7(k-1)/8} \geq \alpha^{b+k+1}$$

for $b \geq 3$ and $k \geq 1$. This inequality certainly holds if $b+7(k-1)/8 \geq 2(b+k+1)/3$ or $5k+8b \geq 37$; if $b \geq 4$ this holds for $k \geq 1$ and if $b=3$ it holds if $5k \geq 13$ or if $k \geq 3$. In the remaining cases, when $b=3$ and $k=1$ or 2 , inequality (4) can be verified directly. This suffices to complete the proof of the lower bound of inequality (1) by induction.

References

- [1] L. RÉDEI, Ein kombinatorischer Satz, *Acta Litt. Szeged*, 7 (1934), 39—43.
- [2] T. SZELE, Kombinatorikai vizsgálatok az irányított teljes gráffal kapcsolatban, *Mat. Fiz. Lapok*, 50, (1943), 223—256. Reprinted as Kombinatorische Untersuchungen über gerichtete vollständige Graphen, *Publ. Math. Debrecen*, 13, (1966), 145—168.

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

(Received September 1, 1970.)