ON SETS OF CONSISTENT ARCS IN A TOURNAMENT

P. Erdős and J. W. Moon

(received August 30, 1964)

1. Introduction. A (round-robin) tournament T_n consists of n nodes $u_1, u_2, \ldots, u_n$ such that each pair of distinct nodes u_i and u_j is joined by one of the (oriented) arcs $\overrightarrow{u_i u_j}$ or $\overrightarrow{u_j u_i}$. The arcs in some set S are said to be consistent if it is possible to relabel the nodes of the tournament in such a way that if the arc $\overrightarrow{u_i u_j}$ is in S then $i > j$. (This is easily seen to be equivalent to requiring that the tournament contains no oriented cycles composed entirely of arcs of S.) Sets of consistent arcs are of interest, for example, when the tournament represents the outcome of a paired-comparison experiment [1]. The object in this note is to obtain bounds for $f(n)$, the greatest integer k such that every tournament T_n contains a set of k consistent arcs.

2. A lower bound for $f(n)$. In this section we show that for all positive integers n,

$$f(n) \geqq [\tfrac{n}{2}] \cdot [\tfrac{n+1}{2}] , \tag{1}$$

where, as usual, $[x]$ denotes the largest integer not exceeding x.

This is trivially true when $n = 1$; suppose it has been established for all n such that $1 \leqq n \leqq m - 1$, and consider any tournament T_m. Since such a tournament has a total of $\frac{1}{2}m(m-1)$ arcs, there must exist some node, say u_m, from which at least

$[\frac{1}{2}m]$ arcs issue. By definition, the tournament defined by the remaining $m-1$ vertices contains a set S of at least $f(m-1)$ consistent arcs. It is clear that the arcs issuing from u_m and the arcs in S are consistent; therefore, appealing to the induction hypothesis, it follows that T_m contains a set of at least

$$[\frac{m}{2}] + [\frac{m-1}{2}] \cdot [\frac{m}{2}] = [\frac{m}{2}] \cdot [\frac{m+1}{2}]$$

consistent arcs. This suffices to complete the proof of (1) by induction.

3. An upper bound for $f(n)$. In this section we show that for any fixed positive ϵ and all sufficiently large values of n,

$$f(n) \leqq \frac{1+\epsilon}{2} \binom{n}{2} . \tag{2}$$

Let $\epsilon > 0$ be chosen. In a tournament T_n there are $n!$ ways of relabelling the nodes and $N = \binom{n}{2}$ pairs of distinct nodes. Hence, there are at most $n! \binom{N}{k}$ such tournaments whose largest set of consistent arcs contains k arcs. So, an upper bound for the number of tournaments T_n which contain a set of more than $(1+\epsilon)N/2$ consistent arcs is given by

$$\begin{aligned} n! \sum_{k>(1+\epsilon)N/2} \binom{N}{k} &< n!\, N \binom{N}{[(1+\epsilon)N/2]} \binom{N}{[N/2]} \binom{N}{[N/2]}^{-1} \\ &< n!\, N 2^N \binom{N}{[(1+\epsilon)N/2]} \binom{N}{[N/2]}^{-1} \\ &= n!\, N 2^N \frac{(N-[N/2])(N-[N/2]-1) \cdots (N-[(1+\epsilon)N/2]+1)}{([N/2]+1)([N/2]+2) \cdots [(1+\epsilon)N/2]} \\ &< n!\, N 2^N e^{-\epsilon^2 N/4} \end{aligned} \tag{3}$$

The last inequality of (3) follows from a simple computation using the fact that $1 - x < e^{-x}$ for $0 < x < 1$. But for all sufficiently large n the last quantity in (3) is easily seen to be less than 2^N, the total number of tournaments with n nodes. Hence, there must be at least one tournament T_n which does not contain any set of more than $(1 + \epsilon)N/2$ consistent arcs. This proves (2), by definition. With a more careful analysis of inequality (3) this argument actually implies that

$$f(n) < 1/2 \binom{n}{2} + (1/2 + o(1)) (n^3 \log n)^{1/2} . \tag{4}$$

It would be desirable to obtain a better estimate for $f(n)$.

The argument employed in the preceding paragraph illustrates the usefulness of probabilistic methods in extremal problems in graph theory, for while we can easily infer the existence of a tournament with a certain required property we are unable to give an explicit construction actually exhibiting such a tournament in general.

4. A more general problem. Let $G(n, m)$ denote an incomplete tournament, or oriented graph, with n nodes and m arcs. Let $f(n, m)$ denote the greatest integer k such that every incomplete tournament $G(n, m)$ contains a set of at least k consistent arcs. If it is assumed that $n \log n/m \to 0$ as n and m tend to infinity then it can be shown, by arguments similar to those used above, that

$$\lim_{n\to\infty} f(n, m)/m = 1/2 . \tag{5}$$

REFERENCE

1. M. G. Kendall and B. Babington Smith, On the method of paired comparisons, Biometrika, 31 (1939) 324-345.

University College London