Short certificates for tournaments

Noga Alon *　　Miklós Ruszinkó †

Submitted: November 6, 1996; Accepted: March 13, 1997.

Abstract

An *isomorphism certificate* of a labeled tournament T is a labeled subdigraph of T which together with an unlabeled copy of T allows the errorless reconstruction of T. It is shown that any tournament on n vertices contains an isomorphism certificate with at most $n \log_2 n$ edges. This answers a question of Fishburn, Kim and Tetali. A *score certificate* of T is a labeled subdigraph of T which together with the score sequence of T allows its errorless reconstruction. It is shown that there is an absolute constant $\epsilon > 0$ so that any tournament on n vertices contains a score certificate with at most $(1/2 - \epsilon)n^2$ edges.

1 Introduction

A *tournament* is an oriented complete graph. An *isomorphism certificate* of a labeled tournament T is a labeled subdigraph D of T which together with an unlabeled copy of T allows the errorless reconstruction of T. More precisely, if $V = \{v_1, \ldots, v_n\}$ denotes the vertex set of T, then a subdigraph D of T is such a certificate if for any tournament T' on V which is isomorphic to T and contains D, T' is, in fact, identical to T. The *size* of the certificate D is the number of its edges, and D is a *minimum* certificate if no *isomorphism certificate* has a smaller size.

Note that the unique directed Hamilton path in a transitive tournament on n vertices is an *isomorphism certificate* of size $n-1$ for the tournament. It is also not difficult to check that any edge

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel. Email: noga@math.tau.ac.il. Research Supported in part by a USA Israeli BSF grant.

†Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest P.O.Box 63, Hungary-1518. Email: ruszinko@lutra.sztaki.hu. Research supported in part by OTKA Grants T 016414 and W 015796 and the "Magyar Tudományért" Foundation.

of the cyclic triangle is an *isomorphism certificate* for it, and that there are three edges of the regular tournament on 5 vertices which form an *isomorphism certificate* for it. Besides these examples, it seems that any other tournament on n vertices does not have certificates with less than $n-1$ edges. This was conjectured by Rubinstein [5], motivated by certain questions in Economics.

Conjecture 1.1 ([5]) *There exists an integer n_0 such that the minimum isomorphism certificate of any tournament on $n > n_0$ vertices is of size at least $n-1$.*

As observed by the first author (cf. [5] for a proof), the assertion of the conjecture is at least nearly correct, in the sense that for any $\epsilon > 0$ there exists some $n_0 = n_0(\epsilon)$ so that the minimum isomorphism certificate of any tournament on $n > n_0(\epsilon)$ vertices is of size at least $(1-\epsilon)n$. Fishburn, Kim and Tetali [2] showed that the only tournaments with $n \leq 7$ vertices that contain isomorphism certificates of size smaller than $n-1$ are the regular tournaments on 3 and on 5 vertices, and it is thus reasonable to suspect that one may take $n_0 = 5$ in the above conjecture.

Kim, Spencer and Tetali [4] proved that *most* tournaments on n vertices contain *isomorphism certificates* of size at most $O(n \log n)$, and Fishburn, Kim and Tetali [2] wondered whether there are any tournaments on n vertices in which the size of the minimum *isomorphism certificate* is much larger. Here we show that there are no such tournaments.

Theorem 1.2 *Any tournament on n vertices contains an* isomorphism certificate *of size at most $\log_2 n! \leq n \log_2 n$.*

The *score* of a tournament on n vertices is the vector $(d_1, d_2, \ldots, d_n)$ of outdegrees of its vertices, ordered so that $d_1 \geq d_2 \geq \ldots \geq d_n$. A *score certificate* of a labeled tournament T on a set V of n vertices is a subdigraph D of T such that any tournament on V that contains D and has the same score sequence as T is identical to T. A *score certificate* is *minimum* if no other *score certificate* has less edges. This notion was introduced by Kim, Tetali and Fishburn [3], who proved that the minimum size of a *score certificate* of any tournament on $n > 5$ vertices is at least $n-1$. They also showed, together with the first author (see [2]), that there are tournaments on n vertices whose minimum *score certificates* contain at least $(7/24 + o(1))n^2$ edges, that is, significantly more than half the edges of the tournaments. The proof combines the fact that the quadratic tournaments on p vertices do not contain *score certificates* with less than $(1/2 - o(1))p^2$ edges, as follows easily from Theorem 1.1 in Chapter 9 of [1], with some additional arguments.

Here we show that the maximum possible size of a minimum *score certificate* of a tournament on n vertices is a fraction which is bounded away from that of the total number of edges. This is stated in the following result.

Theorem 1.3 *There exists an $\epsilon > 0$ so that any tournament on n vertices contains a* score certificate *of size at most $(1/2 - \epsilon)n^2$ edges.*

In the rest of this note we prove the above two theorems. All logarithms from now on are in base 2.

2 Isomorphism certificates

In this section we prove Theorem 1.2. The proof is short, and implies a more general statement, as described in the end of the section.

Proof of Theorem 1.2.

Let T be a fixed *unlabeled* tournament on n vertices. For an arbitrary set H of labeled edges on the set $V = \{v_1, \ldots, v_n\}$ of n vertices, we say that a labeled tournament T' on V is *consistent* with H if T' is isomorphic to T and contains all edges in H. Consider the following procedure for producing an *isomorphism certificate* . Initially, define $H_0 = \emptyset$ and let $\mathcal{T}_0$ be the set of all tournaments on V which are consistent with H_0 (that is; the set of all tournaments which are isomorphic to T.) Note that $\mathcal{T}_0$ contains $n!/|Aut(T)|$ tournaments, where $Aut(T)$ is the automorphism group of T. Assuming $i \geq 1$ and assuming H_{i-1} is a set of $i-1$ edges that has already been defined, and $\mathcal{T}_{i-1}$ is the set of all tournaments on V which are consistent with H_{i-1}, proceed as follows. If $|\mathcal{T}_{i-1}| = 1$ stop; H_{i-1} is an *isomorphism certificate* for the unique copy of T which lies in $\mathcal{T}_{i-1}$. Otherwise, pick an arbitrary pair $j < k$ such that there are tournaments T_1 and T_2 in $\mathcal{T}_{i-1}$, with (v_j, v_k) being a directed edge of T_1 and (v_k, v_j) being a directed edge of T_2. Define, now, $H_i = H_{i-1} \cup \{(v_j, v_k)\}$ if the number of tournaments consistent with $H_{i-1} \cup \{(v_j, v_k)\}$ is at most $|\mathcal{T}_{i-1}|/2$. Otherwise, define $H_i = H_{i-1} \cup \{(v_k, v_j)\}$. Note that $\mathcal{T}_{i-1}$ is the disjoint union of tournaments consistent with $H_{i-1} \cup \{(v_j, v_k)\}$ and those consistent with $H_{i-1} \cup \{(v_k, v_j)\}$. Therefore, if $\mathcal{T}_i$ is the set of all tournaments consistent with H_i it follows that $|\mathcal{T}_i| \leq |\mathcal{T}_{i-1}|/2$ for all $i \geq 1$. Moreover, by our choice, no $\mathcal{T}_i$ is empty. Since $|\mathcal{T}_0| = n!/|Aut(T)|$ it follows that there exists some $i \leq \log(n!/|Aut(T)|)$ ($\leq \log n!$) for which $|\mathcal{T}_i| = 1$. The corresponding set of labeled edges H_i is of cardinality at most $\log n!$ and forms an *isomorphism certificate* for the unique copy of T in $\mathcal{T}_i$. Since T was an arbitrary tournament on n vertices, this completes the proof. □

Remark. The argument above is general and has little to do with tournaments. In fact, a similar argument applies for providing small certificates for arbitrary combinatorial structures. Instead of stating the most general result of this type, we mention here only one additional example, and leave the formulation of the obvious generalizations to the reader. A *colored graph* is a graph together with an assignment of a color to each of its edges. Two such graphs are isomorphic if there is a color-preserving isomorphism between them. An *isomorphism certificate* for a labeled colored complete graph K on a set of vertices V is a labeled colored subgraph H of it, such that any colored complete graph on V which is isomorphic to K and contains H is identical to K. The argument above clearly shows that any labeled colored complete graph on n vertices contains an *isomorphism certificate* of size at most $\log n! = O(n \log n)$. Moreover, this estimate is tight, up to a constant factor. To see this, consider the following example. Let U denote the set of all 2^k binary vectors of length k, and let $V = \{x_1, \ldots, x_k\} \cup \{y_u\}_{u \in U}$ be a set of $n = k + 2^k$ vertices. Let K be the colored, complete graph on V in which all the edges connecting two vertices x_i or two vertices y_u are colored red, and the color of each edge of the form $x_i y_u$ is black if $u_i = 1$ and white if $u_i = 0$. We claim that each *isomorphism certificate* for K contains at least $k2^{k-1} = \Omega(n \log n)$ edges. To see this, fix an i, $1 \leq i \leq k$ and let u^0 and u^1 be two vectors in U which are identical in all coordinates besides the $i-th$ coordinate, where $u^0_i = 0$ and $u^1_i = 1$. Note that even if the colors of all edges besides those of the two edges $x_i y_{u^0}$ and $x_i y_{u^1}$ are given, the colors of these two edges are not determined. This means that any *isomorphism certificate* must contain at least one of these two edges. Since there are $k2^{k-1}$ pairwise disjoint pairs of edges of this form this proves the above claim. It is worth noting that the problem of finding a similar example using only two colors (as well as that of showing that the assertion of Theorem 1.2 is tight) seems to be a lot harder.

3 Score certificates

In this section we prove Theorem 1.3. We make no attempt to optimize our estimate for ϵ and prove the theorem for $\epsilon = 1/160$ and $n \geq 80$. (The last inequality can clearly be omitted by reducing ϵ). To simplify the presentation, we omit all floor and ceiling signs whenever these are not crucial.

Proof of Theorem 1.3. Let T be a labeled tournament on the n vertices $v_1, v_2, \ldots, v_n$, where the outdegree of v_i is d_i and $d_1 \geq d_2 \geq \ldots \geq d_n$. Call an edge (v_i, v_j) a *back* edge if $i > j$. A *score reversible set* is a subset E' of the set of edges of T so that the tournament obtained by reversing the direction of all edges in E' has the same score sequence as T. Obviously, any *score certificate* has to intersect all score reversible sets of a given tournament and vice versa: any set of edges that intersects all score reversible subsets is a *score certificate* . To complete the proof it thus suffices to show that T contains a set of ϵn^2 edges which does not contain (as a set) any score reversible subsets, since this implies that the set of all edges besides those form a *score certificate* of the required size.

Claim 1: A score reversible set of edges cannot contain only back edges.

Proof. Suppose this is false, and reversing a subset E' of back edges one can get a tournament with the same score sequence. Let v_i $(1 \leq i \leq n)$ be the vertex of smallest index for which a back edge of the form $(v_j, v_i) \in E'$ has been reversed. Then in the new score sequence the sum of the outdegrees of the first i vertices is strictly bigger than in the original one, supplying the desired contradiction. $\Box$

Therefore, if the number of back edges is at least ϵn^2, the desired result follows. Thus we may and will assume that there are less than ϵn^2 back edges.

Claim 2: There exist at least $n/2$ vertices each of which is incident with at most $4\epsilon n$ back edges.

Proof. Otherwise, there are more than $\frac{1}{2}\frac{n}{2}4\epsilon n = \epsilon n^2$ back edges, contradicting the preceding assumption. $\Box$

Let V' denote such a set of $n/2$ vertices. Clearly, for every $v_i \in V'$

$$n - i - 4\epsilon n \leq d_i \leq n - i + 4\epsilon n.$$

Note that the number of non-back edges in the graph spanned on the vertices V' is at least $\binom{n/2}{2} - \epsilon n^2$. For a non-back edge (v_i, v_j), call the quantity $j - i$ the *length* of the edge. Note that the number of non-back edges of length at least $17\epsilon n$ in the induced subgraph on V' is at least

$$\binom{n/2}{2} - \epsilon n^2 - \frac{n}{2} \cdot 17\epsilon n \geq n^2/16,$$

where we used the fact that ϵ is sufficiently small (say, $\epsilon = 1/160$) and n is sufficiently large (say, $n \geq 80$.)

It is not difficult to partition the set of all non-back edges in V' into $n/2 - 1$ classes, where in each class the maximum indegree and maximum outdegree is at most 1. (In fact, the edges of any digraph D in which all indegrees and all outdegrees are at most h can be partitioned into at most h such classes. To see this, construct a bipartite graph H whose color classes are two copies $A = \{a_1, \ldots, a_m\}$ and $B = \{b_1, \ldots, b_m\}$ of the vertex set of D, and for each directed edge ij of D, add the edge $a_i b_j$ to H. By the Hall-König Theorem the edges of of H can be partitioned into at most h matchings, which give the desired partition of the edges of D.) Therefore, by the pigeon-hole principle there are some $8\epsilon n$ classes which contain together at least

$$\frac{n^2}{16} \cdot \frac{8\epsilon n}{n/2 - 1} \geq \epsilon n^2$$

non-back edges among those of length at least $17\epsilon n$ on V'. Let E' denote the set of these edges. We complete the proof by showing that all edges besides those in E' form a *score certificate* . Suppose this is false. Then there exists a score reversible set $E^\star \subseteq E'$. Let v_k be the vertex with smallest index incident with an edge of $E^\star$. Then v_k is the initial vertex of each such edge, and after reversing the edges in $E^\star$ the outdegree of v_k will decrease. This means that in the new tournament some other vertex v_p must have its outdegree increased to the value of the outdegree of v_k. However, by construction, reversing edges in $E^\star$ may increase the outdegree of a vertex by at most $8\epsilon n$ and if v_p is any vertex whose outdegree increases at all then $p - k \geq 17\epsilon n$. This implies that

$$d_k - d_p > 17\epsilon n - 2 \cdot 4\epsilon n > 8\epsilon n,$$

and shows that the outdegree of v_p in the new tournament cannot increase to reach that of v_k in the original one. This completes the proof. $\square$

Acknowledgment. We would like to thank Imre Leader and Svante Janson for fruitful discussions.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley, 1992.

[2] P. Fishburn, J. H. Kim and P. Tetali, Tournament Certificates, *Technical memorandum, AT& T Bell Laboratories,* February 1994, DIMACS Technical Report No. 94-05.

[3] J. H. Kim, P. Tetali and P. Fishburn, Score Certificates for Tournaments, *J. Graph Theory* 24 (1997), 117-138.

[4] J. H. Kim, J. Spencer and P. Tetali, Certificates for Random Tournaments, personal communication (1996).

[5] A. Rubinstein, Why are certain properties of binary relations relatively more common in natural language ?, *Econometrica*, in press.