# The minimum feedback arc set problem is NP-hard for tournaments

Pierre Charbit*, Stéphan Thomassé† and Anders Yeo‡

**Abstract**

Answering a question of Bang-Jensen and Thomassen [4], we prove that the minimum feedback arc set problem is NP-hard for tournaments.

A *feedback arc set* (fas) in a digraph $D = (V, A)$ is a set $F$ of arcs such that $D \setminus F$ is acyclic. The size of a minimum feedback arc set of $D$ is denoted by $mfas(D)$. A classical result of Lawler and Karp [5] asserts that finding a minimum feedback arc set in a digraph is NP-hard. Bang-Jensen and Thomassen [4] conjectured that finding a minimum fas in a tournament is also NP-hard. A very close answer was given by Ailon, Charikar and Newman in [1] where they prove that the problem is NP-hard under randomized reductions. Our approach is similar but the reduction we use is simpler and therefore easily derandomized via parity-check matrices (see Alon and Spencer [3], p.255). Finally we prove that the minimum fas for tournaments is polynomially equivalent to the minimum fas for digraphs, and thus NP-hard.

The following lemma is just Chebyschev inequality applied to the parity matrix of subset-intersection.

**Lemma 1** *Let $z$ be an integer. We denote by $A$ the $2^z \times 2^z$ matrix whose rows and columns are indexed by the subsets $F_i$ of $\{1, \dots, z\}$ (in any order) and whose entries are $a_{ij} = (-1)^{|(F_i \cap F_j)|}$. For any subset $J$ of $r$ columns, we have:*

$$\sum_{i=1}^{2^z} |\sum_{j \in J} a_{ij}| \leq 2^z \sqrt{r}$$

**Proof.** Observe that $\sum_{i=1}^{2^z} a_{ip}a_{iq} = 0$ when $p \neq q$. Indeed, if $F_p \neq F_q$, there are exactly $2^{z-1}$ subsets $F_i$ for which $|F_i \cap (F_p \Delta F_q)|$ is even (or equivalently $a_{ip} = a_{iq}$). Now by Cauchy-Schwarz:

$$\sum_{i=1}^{2^z} \frac{|\sum_{j \in J} a_{ij}|}{2^z} \leq \sqrt{\frac{\sum_{i=1}^{2^z} (\sum_{j \in J} a_{ij})^2}{2^z}} = \sqrt{\frac{\sum_{i=1}^{2^z} (\sum_{p \in J} a_{ip}^2 + 2\sum_{p \neq q \in J} a_{ip}a_{iq})}{2^z}} = \sqrt{r}$$

■

**Lemma 2** *Let $z$ be any positive integer divisible by three. Let $k = 2^z$ and let $A$ be the $k \times k$ matrix introduced in Lemma 1. Let $B = (b_{ij})$ be the matrix obtained from $A$ by an arbitrary permutation of the columns. Define $q_i$ as follows.*

*Laboratoire Camille Jordan, Université Claude Bernard, Lyon 1, France (email:charbit@univ-lyon1.fr)

†Laboratoire Camille Jordan, Université Claude Bernard, Lyon 1, France (email:thomasse@univ-lyon1.fr)

‡Department of Computer Science, Royal Holloway, University of London, Egham; Surrey TW20 0EX, UK (email:anders@cs.rhul.ac.uk)

$$q_i = \max\{|\sum_{j=1}^{p} b_{ij}| \,:\, p = 1,2,\ldots,k\}$$

*We have* $\sum_{i=1}^{k} q_i \leq 2k^{5/3}$.

**Proof.** Define the integers $l = k^{2/3}$ and $s = k^{1/3}$. For all $i = 1,2,\ldots,k$ and $j = 1,2,\ldots,s$, we let $c_i^j = |\sum_{j'=(j-1)l+1}^{jl} b_{ij'}|$. By Lemma 1 we have $\sum_{i=1}^{k} c_i^j \leq k\sqrt{l}$ for all $j = 1,2,\ldots,s$. Therefore $\sum_{i=1}^{k}\sum_{j=1}^{s} c_i^j \leq ks\sqrt{l} = k^{5/3}$.

We now evaluate $q_i$. Assume that $p$ is defined so that $q_i = |\sum_{j'=1}^{p} b_{ij'}|$. Let $j$ such that $(j-1)l \leq p < jl$. Note that $q_i \leq c_i^1 + c_i^2 + \ldots + c_i^{j-1} + l$, the term $l$ being an upper bound on $|\sum_{i=(j-1)l+1}^{p} b_{ij}|$. Thus $\sum_{i=1}^{k} q_i \leq (\sum_{i=1}^{k}\sum_{j=1}^{s} c_i^j) + kl \leq 2k^{5/3}$. ∎

**Theorem 1** *Let $z$ be any positive integer divisible by three and let $k = 2^z$. There exists a bipartite tournament $G_k$, whose partite sets both have $k$ vertices ($|V(G_k)| = 2k$) and mfas$(G_k) \geq \frac{k^2}{2} - 2k^{5/3}$. Furthermore, we can construct $G_k$ in polynomial time.*

**Proof.** Let $A = (a_{ij})$ be the $k \times k$ matrix given in Lemma 2. Observe that $A$ has $k(k+1)/2$ positive entries since every column has $k/2$ positive entries, except the emptyset column which has $k$. Let the partite sets of $G_k$ be $\{r_1, r_2, \ldots, r_k\}$ and $\{s_1, s_2, \ldots, s_k\}$ respectively. Now add an arc from $r_i$ to $s_j$ if $a_{ij} = -1$ in $A$, and add an arc from $s_j$ to $r_i$ if $a_{ij} = 1$ in $A$. This clearly defines a bipartite tournament, which can be constructed in polynomial time.

Let $\pi$ be a minimum feedback arc set order of $G_k$, i.e. an enumeration of the vertices for which the number of backward arcs is *mfas*$(G_k)$. Without loss of generality we may assume that the order of the $s_j$'s in $\pi$ is $s_1, s_2, \ldots, s_k$. Let $i \in \{1,2,\ldots,k\}$ be arbitrary and define $p$ such that $s_1, s_2, \ldots, s_p$ come before $r_i$ in $\pi$ and $s_{p+1}, s_{p+2}, \ldots, s_k$ come after $r_i$ in $\pi$. Let $m_i$ denote the number of "1" in row $i$ and note that the number of backward arcs adjacent to $r_i$ is the following:

$$|\{a_{ij} : a_{ij} = -1,\, j \leq p\}| + |\{a_{ij} : a_{ij} = 1,\, j > p\}| = |\{a_{ij} : a_{ij} = -1,\, j \leq p\}| + (m_i - |\{a_{ij} : a_{ij} = 1,\, j \leq p\}|)$$

Let $q_i = \min\{\sum_{j=1}^{p} a_{ij} : p = 1,2,\ldots,k\}$ and note that the minimum feedback arc set of $G_k$ is at least $\sum_{i=1}^{k}(m_i + q_i)$, which by Lemma 2 implies that *mfas*$(G_k) \geq \frac{k(k+1)}{2} - 2k^{5/3} > \frac{k^2}{2} - 2k^{5/3}$. ∎

**Theorem 2** *The minimum feedback arc set for tournaments is NP-hard.*

**Proof.** We will reduce from the minimum feedback arc set in general digraphs, so let $D$ be any digraph of order $n$. We may assume that $D$ has no cycles of length two, as deleting such a cycle decreases the minimum feedback arc set by exactly one. We may also assume that $D$ has no loops. Let $V(D) = \{v_1, v_2, \ldots, v_n\}$ and let $k = 2^{6\lceil 1+log_2(n)\rceil}$. Note that $k \in O(n^6)$ and $k \geq 64n^6$.

Let the vertices in the partite sets of $G_k$, which was defined in Theorem 1, be $\{r_1, r_2, \ldots, r_k\}$ and $\{s_1, s_2, \ldots, s_k\}$ respectively.

We now construct the tournament $T$ with vertex set $\{w_i^j : i = 1,2,\ldots,n$ and $j = 1,2,\ldots,k\}$ and the arc set described below. Let $a, b \in \{1,2,\ldots,n\}$ and $i, j \in \{1,2,\ldots,k\}$ be arbitrary. An arc between the veritices $w_a^i$ and $w_b^j$ is in $T$ according to the following rules.

**(a):** $w_a^i w_b^j \in A(T)$ if $a = b$ and $i < j$.

**(b):** $w_a^i w_b^j \in A(T)$ if $v_a v_b \in A(D)$.

**(c):** If $v_a$ and $v_b$ have no arc between them in $D$ and $a < b$ then $w_a^i w_b^j \in A(T)$ if $r_i s_j \in A(G_k)$ and $w_b^j w_a^i \in A(T)$ if $s_j r_i \in A(G_k)$.

Roughly speaking, we blow-up every vertex of $D$ by a transitive tournament of size $k$, and we fill-in the bipartite gaps resulting from non-edges of $D$ by copies of $G_k$.

We will now bound $mfas(T)$ from both above and below. Without loss of generality assume that $|\{v_a v_b : v_a v_b \in A(D),\, a > b\}| = mfas(D)$. Note that Theorem 1 implies that the arcs generated by (c) above will always contribute at least $(\binom{n}{2} - |A(D)|)(\frac{k^2}{2} - 2k^{5/3})$ to $mfas(T)$ and at most $(\binom{n}{2} - |A(D)|)(\frac{k^2}{2} + 2k^{5/3})$. Now consider the following order of the vertices in $T$.

$$w_1^1, w_1^2, \ldots, w_1^k, w_2^1, w_2^2, \ldots, w_2^k, w_3^1, w_3^2, \ldots \ldots, w_n^k$$

This order implies the following bound on $mfas(T)$.

$$mfas(T) \leq k^2 mfas(D) + \left(\binom{n}{2} - |A(D)|\right)\left(\frac{k^2}{2} + 2k^{5/3}\right)$$

In order to bound $mfas(T)$ from below let $\pi$ be an ordering of the vertices in $T$, such that exactly $mfas(T)$ arcs are backward in the ordering. Let $i_1, i_2, \ldots, i_n$ all be integers from $\{1, 2, \ldots, k\}$ and note that there are at least $mfas(D)$ arcs between vertices in $\{w_1^{i_1}, w_2^{i_2}, w_3^{i_3}, \ldots, w_n^{i_n}\}$ which are backward arcs in $\pi$, as this set of vertices induce a digraph isomorphic to $D$. By summing over all possible values of $i_1, i_2, \ldots i_n$ we get $k^n mfas(D)$ backward arcs, where each arc can be counted at most $k^{n-2}$ times. This implies the following bound.

$$mfas(T) \geq \frac{k^n mfas(D)}{k^{n-2}} + \left(\binom{n}{2} - |A(D)|\right)\left(\frac{k^2}{2} - 2k^{5/3}\right)$$

Note that as $k^{1/3} \geq 64^{1/3} n^2 = 4n^2$ we get that $(\binom{n}{2} - |A(D)|) \times 2k^{5/3} < k^2 \frac{2n^2}{k^{1/3}} \leq \frac{k^2}{2}$. The above two bounds now imply the following.

$$mfas(D) - \frac{1}{2} < \frac{mfas(T)}{k^2} - \frac{1}{2}\left(\binom{n}{2} - |A(D)|\right) < mfas(D) + \frac{1}{2}$$

So if we could compute $mfas(T)$ in polynomial time, we would also have computed $mfas(D)$. As our reduction is polynomial, this implies the result. ∎

After submission of this paper the authors became aware of an independent and similar proof of our main result. This paper by Noga Alon (see [2]) was submitted six month earlier than our paper. It uses the same reduction as we do, but uses quadratic residues to produce the bipartite tournaments with high minimum feedback arc set. The argument is slightly more involved than ours, but the bound on the minimum feedback arc set in the bipartite tournament is much stronger.

# References


[1] N. Ailon, M. Charikar and A. Newman, Aggregating Inconsistent Information: Ranking and Clustering, preprint.

[2] N. Alon, Ranking tournaments, preprint.

[3] N. Alon and J. Spencer, The probabilistic method. Second edition. *Wiley-Interscience Series in Discrete Mathematics and Optimization*, New York (2000).

[4] J. Bang-Jensen and C. Thomassen, A polynomial algorithm for the 2-path problem for semicomplete digraphs, *SIAM J. Discrete Math.*, **5** (1992), 366–376.

[5] R. Karp, Reducibility among combinatorial problems, *Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y.*, (1972), 85–103.