Random Sampling of Labeled Tournaments

Lisa McShine
School of Mathematics, Georgia Institute of Technology
Atlanta GA 30332-1060, USA
mcshine@math.gatech.edu *

Submitted: January 5, 2000; Accepted February 23, 2000

Abstract

This note extends a recent result of Kannan, Tetali and Vempala to completely solve, via a simple proof, the problem of random generation of a labeled tournament with a given score vector. The proof uses the method of path coupling applied to an appropriate Markov chain on the set of labeled tournaments with the same score vector. MRS: 65C05, 05C20.

1 Introduction

A *tournament* is an oriented complete graph, which models a real-life round robin tournament with no ties. The *score vector* of a tournament is the sequence of out-degrees of the vertices of the tournament. In general, given a score vector with n scores, there can be an exponential (in n) number of labeled (and unlabeled) tournaments which have the given score vector (see e.g. [7], [6], and references therein). Although the problem of counting the number of labeled (or unlabeled) tournaments with the same score vector *exactly* seems very difficult, no #P-hardness result is known to that effect.

In this work we show that a simple algorithm based on the Markov chain Monte Carlo method is actually a very efficient algorithm for generating a labeled tournament uniformly at random from among the set of all tournaments with a given score vector.

The problem we consider is one of two problems studied in a recent paper of Kannan, Tetali and Vempala [6]: namely, randomly generating labeled bipartite graphs with a given degree sequence, and randomly generating labeled tournaments with a given score vector. For each of the two problems they designed an appropriate Markov

*This research is in part supported by P. Tetali's NSF Grant, DMS-9800351.

chain whose stationary distribution is uniform on the state space of the desired structures (namely, bipartite graphs and tournaments), and under certain assumptions of "near regularity" of the degree sequences (and score vectors, respectively) they showed that the chains converged to stationarity in polynomial time.

In this note we solve the tournament problem in full generality – we show that without any restrictions on the score vector, given a score vector, a simple Markov chain on the set of all labeled tournaments with the given score vector converges to the (uniform) stationary distribution in time $O(n^3 \log n)$. Thus our main contribution is the extension of the result of [6] to all score vectors with a simple proof.

Following [6], we use the Markov chain Monte Carlo method to solve this problem. However, in contrast to [6], which made use of canonical path techniques, we use the technique of path coupling which was introduced by Bubley and Dyer [2, 3]. The presentation of our proof follows the notation and terminology used by Jerrum [5]. For brevity, we omit the motivation behind the problem and refer the reader to [6] and to references therein. Section 2 introduces all the terms used in this paper along with the notation. Section 3 describes the Markov chain formally and also the proof of rapid convergence to stationarity.

2 Preliminaries on Markov Chains

This section will review the definitions for *rapid mixing* of a discrete time Markov chain, following the exposition in Sinclair [8], and Jerrum [5]. Let (Ω, P, π) denote an irreducible and aperiodic Markov chain, $\mathfrak{M}$, (thus, the chain is ergodic), with finite state space Ω, transition probability matrix P, and stationary distribution π. Here P is non-negative and *row-stochastic.*

We assume that the chain is reversible, i.e., that it satisfies the detailed balance conditions, $\pi(x)P(x,y) = \pi(y)P(y,x)$, for all $x, y \in \Omega$. For $x, y \in \Omega$, and $t \in \mathbb{Z}^+$, $P^t(x,y)$ will denote the t-step probability of going from x to y.

The time a Markov chain takes to reach the stationary distribution can be measured at time t by the *total variation distance* between P^t and π, which is given by

$$||P^t, \pi||_{tv} = \max_{x \in \Omega} \frac{1}{2} \sum_{y \in \Omega} |P^t(x,y) - \pi(y)|.$$

For $\epsilon > 0$, the *mixing time* $\tau(\epsilon)$ is defined

$$\tau(\epsilon) = \min\{t : ||P^{t'}, \pi||_{tv} \le \epsilon, \ \forall t' \ge t\}$$

One method of bounding the mixing time of $\mathfrak{M}$ is given by the following lemma, the "Coupling Lemma", originally due to Doeblin [4], and our formulation is from [5].

Lemma 1 *Suppose that $\mathfrak{M}$ is a countable ergodic Markov chain with transition probability matrix P, and let $((X_t, Y_t) : t \in \mathbb{N})$ be a Markovian coupling, i.e., a Markov process satisfying*

$$\mathbb{P}(X_{t+1} = x' \mid X_t = x \wedge Y_t = y) = P(x, x')$$

and

$$\mathbb{P}(Y_{t+1} = y' \mid X_t = x \wedge Y_t = y) = P(y, y')$$

Suppose further that $t : (0, 1] \to \mathbb{N}$ is a function such that $\mathbb{P}(X_{t(\epsilon)} \neq Y_{t(\epsilon)}) \leq \epsilon$ for all $\epsilon \in (0, 1]$, uniformly over the choice of initial state (X_0, Y_0). Then the mixing time $\tau(\epsilon)$ of $\mathfrak{M}$ is bounded above by $t(\epsilon)$.

The last definition needed is that of rapid mixing, as defined in Sinclair [8].

Definition 2 *A family of Markov chains, the size of whose state spaces is parameterized by n, is rapidly mixing iff there exists a polynomially bounded function $q : \mathbb{N} \times \mathbb{R}^+ \to \mathbb{N}$ such that the mixing time $\tau_n(\epsilon)$ of the chain on the nth member of the family satisfies*

$$\tau_n(\epsilon) \leq q(n, \ln \epsilon^{-1})$$

for all n and $\epsilon \in (0, 1]$.

3 The Markov Chain on Tournaments

This Markov chain will be studied by using the method of *path coupling*, introduced by Bubley and Dyer [2, 3]. Path coupling is an extension of the coupling method, in which the coupling is first defined on pairs of "adjacent" states and then is extended to arbitrary pairs of states using shortest paths. The Coupling Lemma mentioned above is still used. The state space must have an adjacency structure and a distance function which will be defined below.

3.1 The State Space

Let T_n denote a labeled tournament on n vertices. The *score vector*, $S(T_n)$, of a tournament T_n is a sequence $\mathbf{s} = (s_1, s_2, \ldots, s_n)$, such that $0 \leq s_1 \leq s_2 \leq \cdots \leq s_n$, and each s_i is the out-degree of one of the vertices of T_n. Performing a Δ-*reversal* is reversing the orientation of the arcs in a 3-cycle in the tournament. Let the state space $\Omega = \mathcal{T}(\mathbf{s})$ be the set of all labeled tournaments on n vertices with score vector $\mathbf{s} = (s_1, s_2, \ldots, s_n)$. We will use the facts that $\mathcal{T}(\mathbf{s})$ is closed under Δ-reversals, that is, Δ-reversals preserve the score vector, and also that any tournament on n vertices can be transformed to any other tournament on n vertices with the same score vector by performing successive Δ-reversals. Note that T_n has precisely $N_{\mathbf{s}} = \binom{n}{3} - \sum_i \binom{s_i}{2}$ labeled 3-cycles (see [1]). The set of 3-cycles will be denoted by $\mathcal{N}_\Delta$.

For completeness we provide here a short proof that the state space is irreducible. (This is proved in a similar way in [1]).

Lemma 3 *If two tournaments on n vertices, T and T', have the same score vector, then T can be transformed to T' by a finite sequence of Δ-reversals.*

Proof. The key observation is that two different tournaments sharing the same score vector differ in the orientation of an edge-disjoint collection of cycles. The proof will proceed by induction on the length of the cycle. If the cycle is length 3, then reversing its orientation is performing a Δ-reversal.

Now suppose that we can reverse a cycle of length $\leq k$ by using a finite sequence of Δ-reversals. Consider a cycle of length $k+1$, passing through the vertices $v_1, \ldots, v_k, v_{k+1}, v_1$, in that order. There are two cases: either $v_i \to v_1$ for some $i \neq k+1$, or $v_1 \to v_i$ for some $i \neq 2$.

In the first case, $(v_1, v_2, \ldots, v_i)$ is a cycle of length $i \leq k$, which may be reversed by a sequence of Δ-reversals. After this we have reversed the orientation of the cycle $(v_1, v_2, \ldots,, v_i)$. Now since the orientation of arc (v_i, v_1) has been reversed, we have a cycle $(v_1, v_i, v_{i+1}, \ldots, v_{k+1})$. Again this is a cycle of length $\leq k$, so by the inductive hypothesis we can reverse its orientation using Δ-reversals. Now all the arcs in the cycle $(v_1, \ldots, v_k, v_{k+1})$ have been reversed and the arc (v_1, v_i) has been reversed twice, so it remains in it original orientation.

In the second case, the first cycle to reverse is $(v_1, v_i, v_{i+1}, \ldots, v_{k+1})$, and after that, since we will then have $v_i \to v_1$, reverse the cycle $(v_1, v_2, \ldots,, v_i)$ using the inductive hypothesis. □

Remark 1. The proof of Lemma 3 in fact shows that we can reverse a cycle of length m in $m-2$ Δ-reversals. This is easy to see, once it is noted that a 4-cycle can be reversed in 2 Δ-reversals. A cycle of length m is reversed (following the inductive step above) by reversing a cycle of length $i+1$ and then a cycle of length $m-i+1$. This takes $(i+1-2)+(m-i+1-2) = m-2$ Δ-reversals.

Let $T \oplus T'$ denote the edge-disjoint collection of cycles in which two tournaments from $\mathcal{T}(\mathbf{s})$ differ. Let cycle $i \in T \oplus T'$ have length m_i. Then the diameter of the interchange graph on $\mathcal{T}(\mathbf{s})$ is $\leq \sum_{i \in T \oplus T'} (m_i - 2)$. Since $\sum_i m_i \leq \binom{n}{2}$, the diameter is $\leq \binom{n}{2} - 2(\#\text{ cycles in } T \oplus T') \leq \binom{n}{2} - 2$.

Remark 2. Brualdi and Qiao [1] give an example of a class of tournaments on n vertices for which the diameter of the interchange graph is $n-2$, and showed that the diameter of the interchange graph on regular (resp. near-regular) tournaments on n vertices with score vector $\mathbf{s}$ is at least $(n-1)^2/4$, (resp. $n(n-2)/4$). They conjecture that this is the actual diameter of those interchange graphs. The difficulty in determining the diameter of the interchange graph (and consequently a lower bound on the mixing time of the Markov chain $\mathfrak{M}$) reduces to determining the number of cycles and their lengths in a tournament from the set $\mathcal{T}(\mathbf{s})$.

We will define an adjacency structure on the state space so that we may apply path coupling. The most natural definition is to declare two states x and x' (labeled tournaments on n vertices with the same score vector) adjacent iff x' may be obtained from x with precisely one Δ-reversal. We will identify the state space with the vertices of a graph whose edges join adjacent states, as defined above. This graph will be referred to as the *interchange graph.* The distance between two states in the state space, namely, the distance $d(x, x')$, is the least number of Δ-reversals needed to transform x into x'. With this definition, the interchange graph is an unweighted, undirected graph, with vertex set Ω, and d is the ordinary graph distance on it.

3.2 The Markov Chain

The natural Markov chain $\mathfrak{M}$ on $\Omega = \mathcal{T}(\mathbf{s})$, which is a random walk on the interchange graph, is defined as follows:

- Let X_t be the current state of the chain, $X_t \in \Omega$, so that X_t is a tournament with score vector $\mathbf{s}$.
- Pick a 3-cycle uniformly at random (u.a.r) from the $N_{\mathbf{s}}$ 3-cycles in X_t.
- Pick $r \in \{0, 1\}$ u.a.r. If $r = 0$, $X_{t+1} = X_t$, if $r = 1$, perform the Δ-reversal, which we will denote $X_{t+1} = X_t \circ \Delta$.

$\mathcal{T}(\mathbf{s})$ is closed under Δ-reversals, and the interchange graph on $\mathcal{T}(\mathbf{s})$ is connected under Δ-reversals, so $\mathfrak{M}$ is irreducible. Hence it has a unique stationary distribution π. Since $P(x, y) = P(y, x)$ for all $x, y \in \Omega$, $\mathfrak{M}$ is symmetric and hence the stationary distribution is uniform over the state space. The self-loop probabilities are defined to be non-zero, so $\mathfrak{M}$ is also aperiodic. Now we are ready to state our main result formally:

Theorem 4 *The mixing time of the Markov chain on $\Omega = \mathcal{T}(\mathbf{s})$ described above is bounded from above by $c_1 N_{\mathbf{s}}(\ln n + \ln \epsilon^{-1})$, for $\epsilon \in [0, 1)$, where $c_1 > 0$ is an absolute constant and $N_{\mathbf{s}}$ is number of 3-cycles in any tournament with $\mathbf{s}$ as the score vector.*

3.3 The Coupling

The Markov chain on the set of all labeled tournaments on n vertices with the same score vector is given above.

It is only necessary to define the path coupling for adjacent states, because the coupling can then be extended via shortest paths in the interchange graph to arbitrary pair of states.

Suppose the current pair of states is (X_t, Y_t), and that X_t is adjacent to Y_t, i.e., $Y_t = X_t \circ \Delta_i$ for some 3-cycle $\Delta_i \in \mathcal{N}_\Delta(X_t)$. For convenience, denote $\Delta_i = (a, b, c)$, that is, $a \to b \to c$. Let $S_\Delta(X_t, Y_t)$ denote the set of 3-cycles which is common to both X_t and Y_t. Let $D_\Delta(X_t) = \mathcal{N}_\Delta(X_t) \setminus S_\Delta(X_t, Y_t)$ and likewise, $D_\Delta(Y_t) = \mathcal{N}_\Delta(Y_t) \setminus S_\Delta(X_t, Y_t)$. There is a one-to-one correspondence between $D_\Delta(X_t)$ and $D_\Delta(Y_t)$, because X_t and Y_t are both tournaments with the same fixed score vector.

We define the coupling via a map $\sigma : \mathcal{N}_\Delta(X_t) \to \mathcal{N}_\Delta(Y_t)$, where $\sigma(\Delta) = \Delta$ if $\Delta \in S_\Delta(X_t, Y_t)$, and the other cases are illustrated in Figure 1. The latter six cases arise from considering all possible relationships between v and the vertices in the 3-cycle (a, b, c).

Recall that the only arcs that differ between X_t and Y_t are those in the 3-cycle $a \to b \to c$ (resp. $a \to c \to b$).

The transition from (X_t, Y_t) to (X_{t+1}, Y_{t+1}) is given by the following experiment:

- Pick an ordered pair $(\Delta, \sigma(\Delta))$ uniformly at random from $N_\Delta(X_t) \times N_\Delta(Y_t)$.
- Pick $r_X \in \{0, 1\}$ uniformly at random
- If $(\Delta, \sigma(\Delta)) = ((a, b, c), (a, c, b))$, set $r_Y = 1 - r_X$, otherwise set $r_Y = r_X$
- $$\begin{array}{ll} r_X = r_Y = 0 & (X_{t+1}, Y_{t+1}) = (X_t, Y_t) \\ r_X = r_Y = 1 & (X_{t+1}, Y_{t+1}) = (X_t \circ \Delta, Y_t \circ \sigma(\Delta)) \\ r_X = 1,\ r_Y = 0 & (X_{t+1}, Y_{t+1}) = (X_t \circ (a, b, c), Y_t) \\ r_X = 0,\ r_Y = 1 & (X_{t+1}, Y_{t+1}) = (X_t, Y_t \circ (a, c, b)) \end{array}$$

Figure 2 shows (X_t, Y_t) and (X_{t+1}, Y_{t+1}) for one of the cases mentioned above. This is typical of all the cases, except where $\sigma(\Delta) = \Delta$ (in which case X_{t+1} and Y_{t+1} also differ precisely in the 3-cycle (a, b, c)), and where $(\Delta, \sigma(\Delta)) = ((a, b, c), (a, c, b))$, (in which case X_{t+1} and Y_{t+1} are identical).

Lemma 5 *For adjacent states X_t and Y_t,*

$$\mathbb{E}(d(X_{t+1}, Y_{t+1}) | X_t, Y_t) \le (1 - \rho) d(X_t, Y_t), \tag{1}$$

where $\frac{1}{\rho} = N_{\mathbf{s}}$.

Proof. As explained above and illustrated in Fig. 1, in all except one of the $N_{\mathbf{s}}$ cases, $d(X_{t+1}, Y_{t+1}) = d(X_t, Y_t) = 1$, and in the last case $d(X_{t+1}, Y_{t+1}) = d(X_t, Y_t) - 1 = 0$, so that we obtain the inequality stated in the lemma. This is basically because in a strong tournament on four vertices, there are two distinct 3-cycles and if one of them is where X_t and Y_t differ, then the coupling is designed in such a way as to map the reversal of the other 3-cycle of X_{t+1} with that of Y_{t+1}; and this either preserves the distance between X_t and Y_t or reduces it. □

The following lemma is due to Bubley and Dyer, and our formulation of it follows Jerrum [5]:

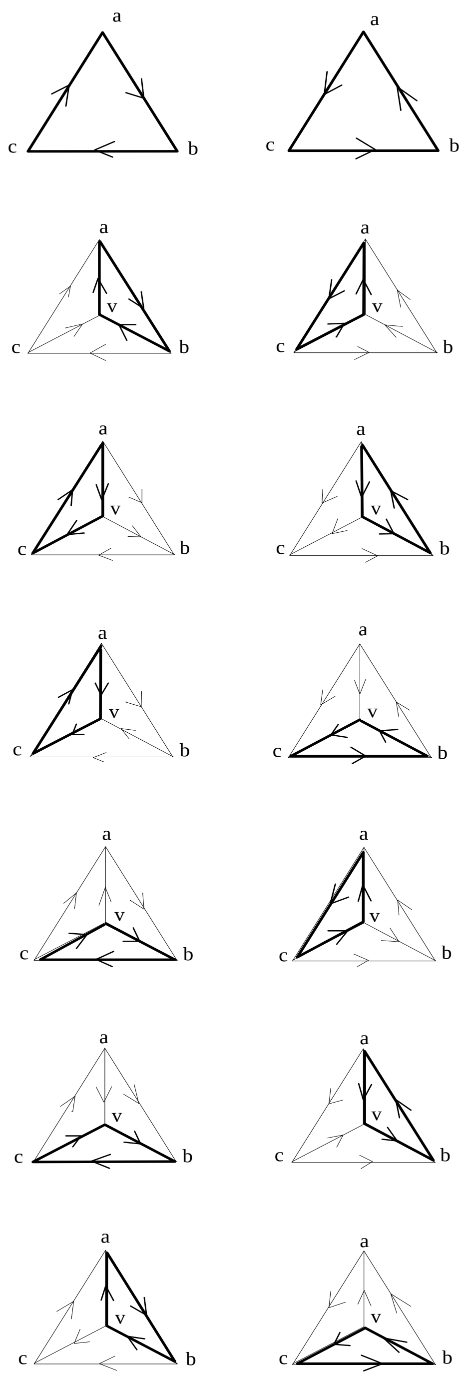

Figure 1: The map $\sigma : N_\Delta(X_t) \to N_\Delta(Y_t)$

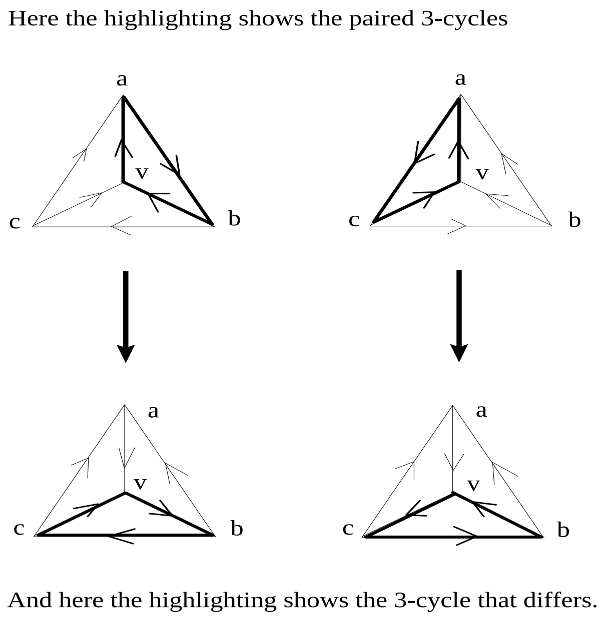

Figure 2: The top picture shows (X_t, Y_t), and the bottom shows (X_{t+1}, Y_{t+1})

Lemma 6 *Suppose a coupling (X_t, Y_t) has been defined for $\mathfrak{M}$ on adjacent pairs of states, and suppose that the coupling satisfies the contraction condition (1) on adjacent pairs. Then the coupling can be extended to all pairs of states in such a way that (1) holds unconditionally.*

Proof of Theorem 4. As shown in Lemma 5, the Markov chain satisfies the contraction property required by Lemma 6. Iteration of Lemma 5 gives $\mathbf{E}(d(X_t, Y_t)|X_0, Y_0) \leq (1-\rho)^t d(X_0, Y_0)$. Next, the interchange graph is connected via Δ-reversals, and as shown in Remark 1, the diameter of the interchange graph is $\leq \binom{n}{2} - 2$. Thus $d(X_0, Y_0) < \binom{n}{2}$, and

$$\mathbb{P}(X_t \neq Y_t) \leq \mathbb{E}(d(X_t, Y_t)) \leq (1-\rho)^t n^2/2.$$

The latter quantity is less than ϵ if $t \geq N_{\mathbf{s}}(2\ln n + \ln(2\epsilon)^{-1})$. $\square$

Remark 3. We may use Theorem 6.11 from the survey article on tournaments by Reid and Beineke [7] to bound $N_{\mathbf{s}}$. Let $C(n,3)$ denote the maximum number of 3-cycles in a tournament on n vertices. Then Theorem 6.11 [7] yields $C(n,3) = \frac{1}{24}n(n^2-1)$, if n is odd, and $C(n,3) \leq \frac{1}{24}n(n^2-4)$, if n is even. Hence $N_{\mathbf{s}} \leq C(n,3) \leq n^3/24$, and so

$$\tau(\epsilon) \leq (n^3/24)(2\ln n + \ln(2\epsilon)^{-1}),$$

as claimed in the introduction.

Acknowledgments. The author thanks Prasad Tetali and an anonymous referee for their helpful comments.

References

[1] R.A. Brualdi and L. Qiao. The interchange graph of tournaments with the same score vector, *Progress in Graph Theory*, Academic Press, Canada, 1984, pp. 128–151.

[2] R. Bubley and M. Dyer. Path coupling, Dobrushin uniqueness, and approximate counting, Report 97.04, School of Computer Studies, University of Leeds, 1997.

[3] R. Bubley and M. Dyer. Path coupling: a technique for proving rapid mixing in Markov chains, *Proc. 38th IEEE Symposium on Foundations of Computer Science*, IEEE Computer Society Press, 1997, pp. 223–231.

[4] W. Doeblin. Exposé de la théorie des chaines simples constantes de Markov à un nombre fini d'états, *Revue Mathématique de l'Union Interbalkanique*, **2** (1933), pp. 77–105.

[5] M. Jerrum. Mathematical foundations of the Markov chain Monte Carlo method, *Probabilistic Methods for Algorithmic Discrete Mathematics*, Algorithms and combinatorics **16**, eds. M. Habib, C. McDiarmid, J. Ramirez-Alfonsin and B. Reed, Springer, 1998, pp. 116–165.

[6] R. Kannan, P. Tetali, S. Vempala. Simple Markov-chain algorithms for generating bipartite graphs and tournaments, *Random Struct. Alg.*, **14**, 1999, pp. 293–308.

[7] K.B. Reid and L.W. Beineke. Tournaments, Ch. 7, *Selected Topics in Graph Theory*, eds. L.W. Beineke and R.J. Wilson, Academic Press, New York, 1978, pp. 169–204.

[8] A.J. Sinclair. *Algorithms for random generation and & counting: a Markov chain approach.* Birkhäuser, Boston, 1993.