18.317 Combinatorics, Probability, and Computations on Groups	5 December 2001
Lecture 33	
Lecturer: Igor Pak	*Scribe: Christopher Malon*

Blind Algorithms and Product Replacement

Recall the Product Replacement Algorithm:

- Start at a generating k-tuple $\langle g_1, \ldots, g_k \rangle = G$.
- Run a random walk on $\Gamma_k(G)$ for T steps.
- Output a random component g_i of the vertex you arrive at.

So that we know how long to take the random walk in this algorithm, it would be helpful to know whether the mixing time of $\Gamma_k(G)$ is polynomial in $\log|G|$.

We can make some trivial observations in response to this question:

- $\Gamma_k(G)$ need not even be connected, so the mixing time could be infinite.
- If $k > d(G) + m(G)$, then $\Gamma_k(G)$ is connected, and its mixing time is finite.
- The diameter of $\Gamma_k(G)$ is $O(\log^2|G|)$ for $k = 2\log|G|$. The mixing time must be at least as big as the diameter, but we don't know how much bigger.

We will prove:

Theorem 1 *Given $c, c' > 0$, there is a constant $c'' > 0$ so that if $c \log|G| \log\log|G| \le k \le c' \log|G| \log\log|G|$, then the mixing time $\tau_4 \le c'' \log^{14}|G| \log\log^5|G|)$.*

Blind Algorithms

Suppose $R_1, \ldots, R_k$ are reversible Markov chains on $\{1, \ldots, n\}$, and let π be a stationary distribution, *i.e.*, $R_i\pi = \pi$ for all i. (If π is a uniform distribution, then reversibility means that the R_i are symmetric matrices.) Define $M = \frac{1}{k}(R_1 + \ldots + R_k)$, which is again a Markov chain satisfying $M\pi = \pi$.

Let $\overline{a} = (a_1, \ldots)$ be a finite sequence with each $a_i \in \{1, \ldots, k\}$. Let $l(\overline{a})$ denote the length of the sequence $\overline{a}$. Let $\mathcal{A}$ be the set of all such sequences $\overline{a}$, and A be a probability distribution on $\mathcal{A}$. Let $T = E_A(l(\overline{a}))$ be the expectation value of the length. For each $\overline{a}$, define $R_{\overline{a}} = R_{a_1} \cdots R_{a_{l(\overline{a})}}$.

Definition 2 *A defines a* blind algorithm *if, for all $i \in \{1, \ldots, n\}$, we have $\|E_A(R_{\overline{a}}(i)) - \pi\| < \frac{1}{4}$.*

A special case of a blind algorithm arises when we have a labeled graph on n vertices, the transition probabilities in each R_i are positive only between vertices that are joined by an edge, and the R_i are symmetric (so that the uniform distribution is stationary with respect to all R_i). If we fix a starting vertex i, each sequence $\overline{a}$ defines a probability distribution on the vertices of the graph, namely, the probability distribution over the endpoints of paths of length $l(\overline{a})$ from i, in which we use R_{a_j} to decide where to go on the jth step. If we, furthermore, impose a probability distribution A on the sequences $\overline{a}$, then we get a probability distribution Q_i on all the vertices of the graph. To say that A defines a "blind" algorithm means that for all i, the separation distance $\|Q_i - U\| < \frac{1}{4}$.

Recall the fourth definition of mixing time for a random walk whose probability distribution is Q^t at the tth step (Lecture 12, October 5):

$$\tau_4 = \min\{t : \|Q^t - U\| < \frac{1}{4}\}$$

Note that neither A nor $\overline{a}$ defines a random walk in the usual sense, because the transition probabilities at each step depend on more than our location in the graph. However, $M = \frac{1}{k}(R_1 + \dots R_k)$ does define a random walk, and we have the following theorem.

Theorem 3 *Let $M = \frac{1}{k}(R_1 + \ldots + R_k)$. If A defines a blind algorithm and T is the expected length of a path chosen from $\mathcal{A}$ via A, then the mixing time $\tau_4(M) = O(T^2 k \log \frac{1}{\pi_0})$, where π_0 is the minimum of the entries appearing in the stationary distribution π.*

We won't prove this theorem, but we'll apply it in a special case.

Suppose G is a finite group, $S = S^{-1} = \{s_1, \dots, s_k\}$ is a symmetric generating set, and $\Gamma = \Gamma(G, S)$ is the corresponding Cayley graph. Take the R_i to be the permutation matrix given by right multiplication $g \to gs_i$ (a deterministic Markov chain). Given any sequence $\overline{a} = (a_1, \dots, a_l)$, $R_{\overline{a}}$ sends $g \to gs_{a_1} \cdots s_{a_l}$. For every element $g \in G$, fix a path from the identity e to g of minimal length. Define a probability distribution A on $\mathcal{A}$ to be $\frac{1}{|G|}$ at $\overline{a}$ if $s_{a_1}, s_{a_1}s_{a_2}, \dots, s_{a_1}s_{a_2}\cdots s_{a_l}$ is the selected path from e to the group element $s_{a_1}\cdots s_{a_l}$, and zero otherwise.

In $G = \mathbb{Z}_n$ with $S = \{\pm 1\}$, there are only one or two ways to fix these paths (the shortest decomposition of each element, except possibly $\frac{n}{2}$, is unique). The matrix R_1 corresponds to moving left through the cycle, and R_2 to moving right. The expected length T of a path is $O(n)$, and $\pi_0 = \frac{1}{n}$ because the uniform distribution is stationary under R_1 and R_2. By Theorem 3, the mixing time for this Cayley graph is $O(n^2 \log n)$. This result is close to what we know ($O(n^2)$).

For any finite group G with A defined as above, we have $T = E_A(l(\overline{a})) \leq d$ where $d = \operatorname{diam}(\Gamma(G, S))$. Thus, the mixing time for a random walk on Γ where we apply generators $s \in S$ uniformly at random is $O(d^2 \log |G|)$.

A Blind Algorithm on the Product Replacement Graph

Finally, we sketch the proof of Theorem 1. Recall that the edges in the graph $\Gamma_k(G)$ are given by

$$R_{ij}^{\pm} : (g_1, \dots, g_k) \to (g_1, \dots, g_i g_j^{\pm}, \dots, g_k)$$

where $g_i g_j^{\pm}$ appears in the ith position. There are $O(|G|^k)$ vertices in the graph, and $O(k^2)$ edges emanate from every vertex. Theorem 3 will give us a bound on the mixing time of $\Gamma_k(G)$ if we construct a blind algorithm A with respect to the $R_{ij}^{\pm}$.

Let $(g_1, \dots, g_r)$ be a generating r–tuple for G, where $r = O(\log |G|)$, and consider $(g_1, \dots, g_r, 1, \dots, 1) \in \Gamma_k(G)$. Instead of following a random walk on the product replacement graph $\Gamma_k(G)$, we're going to embed

Babai's algorithm for generating uniform random group elements into an algorithm on $\Gamma_k(G)$. Start by setting $s = r$. We will define the probability distribution A on $\mathcal{A}$ as follows. For each of the first L steps, choose $i \in \{1, \ldots, s\}$ and $\pm$ uniformly at random, and apply $R^{\pm}_{s+1,i}$. After these L steps, increment s and repeat. Analysis of the Babai algorithm (November 14) shows that if we take $L = O(\log^3 |G|)$ and do this $l = O(\log |G|)$ times, we should have g_{r+l} close to uniform in G.

By multiplying every position by a nearly uniform group element in this manner, we can obtain a nearly uniform element of $\Gamma_k(G)$ in $T = O(k \cdot \log^4 |G|)$ steps. There are a lot of technical details to work through here, and they weren't covered in class.

As $k \leq c' \log |G| \log\log |G|$, Theorem 3 yields

$$\begin{aligned} \tau_4 &= O(T^2 k^2 \log \frac{1}{\frac{1}{|G|^k}}) \\ &= O(k^5 (\log |G|)^9) \\ &= O(\log^{14} |G| \log\log^5 |G|) \end{aligned}$$

as desired.