18.317 Combinatorics, Probability, and Computations on Groups 28 September 2001

Lecture 9

Lecturer: Igor Pak *Scribe: Christopher Malon*

Generalized Random Subproducts

Today, we start to upgrade the Erdős–Rényi machine to show that "most" generating sets of size $O(\log |G|)$ have a mixing time of $O(\log |G|)$. We'll be more precise about this in due course.

First, we want to show that Erdős–Rényi is robust when we insert junk in the middle of our strings. Fix $\overline{g} = (g_1, g_2, \ldots, g_k) \in G^k$. On Monday, we defined a probability distribution over G

$$Q_{\overline{g}}(h) = \Pr_{\overline{\epsilon}}\left[g_1^{\epsilon_1} \cdots g_k^{\epsilon_k} = h\right]$$

where $\overline{\epsilon} = (\epsilon_1, \ldots, \epsilon_k)$ is picked randomly, uniformly from $\{0,1\}^k$. Now fix group elements $x_1, \ldots, x_l$ and integers $\gamma_1, \ldots, \gamma_l$. We insert the x_i at intervals γ_i: Let

$$R_{\overline{g}}(h) = \Pr_{\overline{\epsilon}}\left[g_1^{\epsilon_1} \cdots g_{\gamma_1}^{\epsilon_{\gamma_1}} x_1 g_{\gamma_1+1}^{\epsilon_{\gamma_1+1}} \cdots g_{\gamma_1+\gamma_2}^{\epsilon_{\gamma_1+\gamma_2}} x_2 \cdots x_l \cdots g_k^{\epsilon_k} = h\right]$$

Let's look at this in the case where $l = 1$. In this case, we are considering products of the form

$$g_1^{\epsilon_1} \cdots g_i^{\epsilon_i} x g_{i+1}^{\epsilon_{i+1}} \cdots g_k^{\epsilon_k} = g_1^{\epsilon_1} \cdots g_i^{\epsilon_i} \left(g_{i+1}^{x}\right)^{\epsilon_{i+1}} \cdots \left(g_k^{x}\right)^{\epsilon_k} x$$

where $i = \gamma_1$, $x = x_1$, and g^x denotes xgx^{-1}. Declare $\overline{z}(x, \gamma) = (1, \ldots, 1, x, \ldots, x)$, where 1's appear in the first i positions, and write

$$(g_1, \ldots, g_k)^{(z_1, \ldots, z_k)} = (g_1^{z_1}, \ldots, g_k^{z_k})$$

Evidently,

$$R_{\overline{g}}(h) = Q_{\overline{g}^{\overline{z}(x,\gamma)}}(hx^{-1})$$

If $l > 1$, we just have to repeat these maneuvers to define a string $\overline{z}(\overline{x}, \overline{\gamma})$ and a function $f(\overline{x})$ so that

$$g_1^{\epsilon_1} \cdots g_{\gamma_1}^{\epsilon_{\gamma_1}} x_1 g_{\gamma_1+1}^{\epsilon_{\gamma_1+1}} \cdots g_{\gamma_1+\gamma_2}^{\epsilon_{\gamma_1+\gamma_2}} x_2 \cdots x_l \cdots g_k^{\epsilon_k} = \left(\overline{g}^{\overline{z}(\overline{x},\overline{\gamma})}\right)^{\overline{\epsilon}} \cdot (f(\overline{x}))^{-1}$$

Then we have

$$R_{\overline{g}}(h) = Q_{\overline{g}^{\overline{z}(\overline{x},\overline{\gamma})}}(h \cdot f(\overline{x})) \tag{1}$$

We want to show $R_{\overline{g}}$ is usually close to uniform.

Definition 1 *A probability distribution Q on a finite group G is ϵ–*uniform *if*

$$Q(h) > \frac{1-\epsilon}{|G|}$$

for all $h \in G$.

Two lectures ago, we proved:

Theorem 2 *(Erdős–Rényi) For all $\epsilon, \delta > 0$, $Q_{\overline{g}}$ is ϵ–uniform for more than $1 - \delta$ proportion of $\overline{g} \in G^k$, given $k > 2\log_2 |G| + 2\log_2 \frac{1}{\epsilon} + \log_2 \frac{1}{\delta}$.*

Multiplication by $f(\overline{x})$ is a bijection on the elements of G, so ϵ–uniformity of $Q_{\overline{g}}(h)$ over h implies ϵ–uniformity of $Q_{\overline{g}}(h \cdot f(\overline{x}))$ over h. Because conjugation by $\overline{z}(\overline{x}, \overline{\gamma})$ is a bijection on G^k, the $R_{\overline{g}}$ will be ϵ–uniform for the same proportion of $\overline{g}$ as $Q_{\overline{g}}$, applying equation (1). Thus, we have Erdős–Rényi for $R_{\overline{g}}$:

Theorem 3 *For all $\epsilon, \delta > 0$, $R_{\overline{g}}$ is ϵ–uniform for more than $1 - \delta$ proportion of $\overline{g} \in G^k$, given $k > 2\log_2 |G| + 2\log_2 \frac{1}{\epsilon} + \log_2 \frac{1}{\delta}$.*

Now, we look at a probability distribution of group elements over a larger sample space, describing what happens when the x_i may or may not be inserted. Define

$$Q_{\overline{g},\overline{x}}(h) = \Pr_{\overline{\epsilon},\overline{\alpha}}\left[g_1^{\epsilon_1} \cdots g_{\gamma_1}^{\epsilon_{\gamma_1}} x_1^{\alpha_1} g_{\gamma_1+1}^{\epsilon_{\gamma_1+1}} \cdots g_{\gamma_1+\gamma_2}^{\epsilon_{\gamma_1+\gamma_2}} x_2^{\alpha_2} \cdots x_l^{\alpha_l} \cdots g_k^{\epsilon_k} = h\right]$$

where $\overline{\alpha}$ is picked uniformly from $\{0,1\}^l$. For fixed $\overline{\alpha}$, let

$$R_{\overline{g},\overline{x},\overline{\alpha}}(h) = \Pr_{\overline{\epsilon}}\left[g_1^{\epsilon_1} \cdots g_{\gamma_1}^{\epsilon_{\gamma_1}} x_1^{\alpha_1} g_{\gamma_1+1}^{\epsilon_{\gamma_1+1}} \cdots g_{\gamma_1+\gamma_2}^{\epsilon_{\gamma_1+\gamma_2}} x_2^{\alpha_2} \cdots x_l^{\alpha_l} \cdots g_k^{\epsilon_k} = h\right]$$

Then we have

$$Q_{\overline{g},\overline{x}} = \frac{1}{2^l} \sum_{\overline{\alpha}} R_{\overline{g},\overline{x},\overline{\alpha}} \tag{2}$$

Suppose $k > 2\log_2 |G| + 2\log_2 \frac{1}{\epsilon} + \log_2 \frac{1}{\delta}$ is fixed. Draw a grid whose rows represent the choices of $\overline{g}$ from G^k, and whose columns represent the choices of $\overline{\alpha}$. Keep l, $\overline{\gamma}$, and $\overline{x}$ fixed. Mark the $(\overline{g}, \overline{\alpha})$ position in this grid if $R_{\overline{g},\overline{x},\overline{\alpha}}$ is ϵ–uniform. Theorem 3 applies to $R_{\overline{g},\overline{x},\overline{\alpha}}$, saying that in every column (*i.e.* for any fixed $\overline{\alpha}$), the proportion of unmarked squares is less than δ. Consequently, less than $\sqrt{\delta}$ of the rows have more than $\sqrt{\delta}$ of their positions unmarked. By equation (2), for more than $1 - \sqrt{\delta}$ of the $\overline{g}$,

$$Q_{\overline{g},\overline{x}}(h) > \frac{1-\sqrt{\delta}}{|G|}(1-\epsilon)$$

This proves:

Theorem 4 *For every $\epsilon, \delta > 0$, $Q_{\overline{g},\overline{x}}$ is $(\epsilon + \delta)$–uniform for more than $1 - \sqrt{\delta}$ proportion of $\overline{g}$, given $k > 2\log_2 |G| + 2\log_2 \frac{1}{\epsilon} + \log_2 \frac{1}{\delta}$.*

Lazy Random Walks

Now we return to our question about mixing time.

Definition 5 *Fix $\overline{g} = (g_1, \ldots, g_k)$, not necessarily generating G. A* lazy random walk *through $\overline{g}$ is a sequence of group elements X_t such that $X_0 = 1$ and $X_{t+1} = X_t \cdot g_{i_{t+1}}^{\epsilon_{t+1}}$, where each i_{t+1} is picked randomly, uniformly from $\{1, \ldots, k\}$, and ϵ_{t+1} is picked randomly, uniformly from $\{0,1\}$. Thus*

$$X_t = g_{i_1}^{\epsilon_1} \cdots g_{i_t}^{\epsilon_t}$$

Let $P_{\overline{g}}^t$ be the probability distribution of the lazy random walk through $\overline{g}$ after t steps. Today's work has shown us: If, after t steps, we have selected k' distinct i's (where k' is at least as big as in theorem 4), then $P_{\overline{g}}^t$ will be close to uniform for most $\overline{g}$; the redundant generators chosen in the lazy random walk will take the place of $\overline{x}$.

From the Coupon Collector's Problem, we can compute how big t must be in order to provide enough generators. If we need to collect all the g_i and $\overline{g} = (g_1, \ldots, g_k)$, the expected waiting time is

$$1 + \frac{k}{k-1} + \frac{k}{k-2} + \ldots + k = k \log k + O(k)$$

We'll fill in more details in the next lecture.