18.317 Combinatorics, Probability, and Computations on Groups	3 December 2001
Lecture 32	
Lecturer: Igor Pak	*Scribe: Dennis Clark*

Proving Bias

In this lecture, we set out to demonstrate that there are examples where the product replacement algorithm produces bias in its output. Continuing with the example of last time, we let $G = A_n^{n!/8}$, $G =< g_1, \dots g_k >$, Q the probability distribution of g_1 in a random generating k-tuple, and U the uniform distribution.

Then, we have:

Theorem 1 *Let $k = o(n)$. Then $||Q - U|| \to 1$ as $n \to \infty$.*

Proof: Let $B \subset G$ such that $Q(B) \to 0$ and $U(B) = \frac{|B|}{|G|} \to 1$ as $n \to \infty$. The existence of such a set was demonstrated last class. Then

$$||Q - U|| = \max_{B \subset G} |Q(B) - U(B)| \to 1 \tag{1}$$

as $n \to \infty$, as needed. ■

But we need to do better than this. If we were federal inspectors attempting to prove boxes of group elements to be non-uniform, only knowing that elements in those boxes were unlikely to belong to B would not be good enough. We need the following theorem.

Theorem 2 *Let $k = o(n)$. Then there exists a word $w(x_1 \dots x_r)$ of length $o(\log \log |G|)^c$ such that $w[Q] = 1$ with high probability (i.e. probability $> 1 - \frac{1}{n^{\alpha n}}$) and $w[U] \neq 1$ with high probability.*

The proof comes in three parts. One part is like last time, one part is a statistical story, and one part is combinatorics. We'll do the not-too-combinatorial combinatorial part first.

Proof: We're going to try to implement something like predicate logic in $H = A_n$. Our basic predicate will be the compare-to-identity predicate.

Let A be the event $h_1 = 1$ and B be the event $h_2 = 1$, where h_1 and h_2 are elements drawn according to some distribution P you don't know. Then, it would be almost right to say that $A \vee B$ is given by $[h_1, h_2] = 1$ with high probability, and $A \wedge B = h_1 h_2$ with high probability.

Now, we take an element h to another element v, where

$$v = (h)^{u_1} (h_1)^{u_2} \dots (h_r)^{u_r} \tag{2}$$

wheter the u_i are sampled uniformly and independently, and the h_j are independently sampled from P. The product should be nearly uniform if $h \neq 1$.

Let v_1, v_2 both be either the identity or random with high probability. Then we get $v_1 v_2 = 1$ with high probability if and only if both elements are 1 with high probability. We then compute:

$$Pr([v_1, v_2] = 1) = \frac{\#conj.classes(H)}{|H|} < \frac{2\#partitionsofn}{\frac{n!}{2}}. \tag{3}$$

This is approximately equal to $\frac{1}{n^{\alpha n}} \to 0$. ■

Now we think of $h \to v_r$ as a random walk. The probability distribution of h^{v_i} is invariant on conjugacy classes.

Now we make a claim about the mixing time.

Claim 3 *Let $h = \sigma \in c_\lambda$ with high probability. Then the mixing time of a random walk on A_n with generating set c_λ is less than $diam^2(A_n, c_\lambda) \log |A_n| \approx n \log n$.*

Proof: First,

$$diam(A_n, c_\lambda) = \frac{n}{2} \tag{4}$$

where the generating set is pairs of transpositions, and

$$diam(S_n, c_\lambda) \leq n - 1 \tag{5}$$

where the generating set is all transpositions. Also,

$$diam(S_n, c_{(n)}) = 3 \tag{6}$$

since any transposition is the product of two long cycles. Also, σ is a product of two cycles, of length λ_1, λ_2, whose sum is k. S_k is given by two cylces, and S_{2k} is given by 4.

Then the diameter is some polynomial in n, so we're done, since the mixing time is polynomial in n. ■

Now, we go on to the statistical question: how do we demonstrate that there is bias in some sample? We take our example from the idea of proving discrimination to a judge, and we set it on α- Centauri.

On α-Centauri, there are three kinds of humanoids: red-haired ones, blue-haired ones, and green-haired ones. The CSGHH (Centaurian Society for Green Haired Humanoids) is suing company XYZ for hair-discrimination, claiming that green-haired humanoids get paid less. The lawyers go to court with statistics showing that the average pay for a green-haired humanoid is 12.3 units, while the average for other hair colors is 12.5 units. The number of employees at company XYZ is huge, so this is a statistically significant difference, but it's difficult to convince a judge of that. So the lawyers employ a never-before-seen argument, and win the case. The argument runs as follows:

There are clearly two hypotheses: either there is or is not discrimination. Suppose not. Divide the company randomly into small groups numbered with positive integers, and for each group number i, let

$$\epsilon_i = \begin{cases} 1 & \text{if the group's highest earner has green hair} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Let p be the probability that the highest earner in a single group has green hair if there is no discrimination, and take $l = \frac{1}{p}$. Then, divide the groups into blocks of l and take the and of small sets of blocks. Then or together these results in groups, and those in groups, or those in groups, and continue alternating until you get the entire company:

$$[(\epsilon_1 \vee \cdots \vee \epsilon_l) \wedge (\epsilon_{l+1} \vee \dots) \dots] \wedge \dots \tag{8}$$

The result will, with high probability, be 1 if there is no discrimination, and 0 otherwise.

This converts the problem into a similar combinatorial situation considered by Ajtai. We have a sequence of ones and zeros, and we want to figure out if it has "too many" zeros. We take intervals of appropriate length and proceed as above, giving a function that will tell if there are too many. This is called bias amplification.

Back in our group theoretical world, we create the word w based on the boolean formula we just created, using our translation of logic into the group, and we're done. The word ends up being close to what a group theorist would think of as a law on the group.