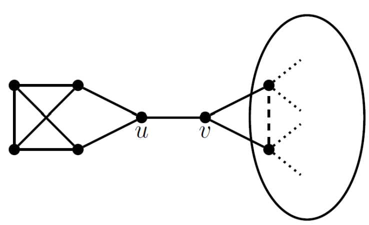
CS5330: Sample Endterm

April 14, 2020

Here are brief solution sketches to the problems.

1. Give an example of a family $\mathcal{H}$ of hash functions $h : [n] \to \{0,1\}$ such that $\mathcal{H}$ is strongly $(n-1)$-universal but not strongly n-universal.

   A function $h : [n] \to \{0,1\}$ can be viewed as a bitstring of length n: the i'th coordinate equals $h(i)$. Consider the family $\mathcal{H}$ of strings $\{z \circ a : z \in \{0,1\}^{n-1}, a = z_1 \oplus z_2 \oplus \cdots \oplus z_{n-1}\}$ where $\circ$ denotes string concatenation and $\oplus$ denotes the XOR operation. You can check that if a string is drawn uniformly from $\mathcal{H}$, the first $n-1$ coordinates are independent while clearly the last coordinate depends on the rest. This is true for any subset of $n-1$ coordinates. Hence, $\mathcal{H}$ describes family of hash functions that is $(n-1)$-universal but not n-universal.

2. 

   In the graph above, the ellipse on the right side is a 3-regular graph with $n-6$ vertices where the dashed edge is absent. Show that $h_{u,v} = O(1)$ but $h_{v,u} = \Omega(n)$. (**Hint**: What can you say about $h_{u,u}$?)

   The hitting time from u to v in the given graph G is the same as the hitting time from u to v in the constant-sized graph that doesn't contain the $n-6$ vertices to the right. The hitting time between two vertices in any constant sized graph is clearly $O(1)$.

   The stationary distribution on a regular graph is the uniform distribution. Hence, $h_{u,u} = n$. If the neighbors of u are v, w, and x, then we get $n = h_{u,u} = \frac{1}{3}(h_{v,u}+1) + \frac{1}{3}(h_{w,u}+1) + \frac{1}{3}(h_{x,u}+1)$. Clearly, $h_{w,u}$ and $h_{x,u}$ are $O(1)$ by the same logic as above. It follows that $h_{v,u} = \Theta(n)$.

3. Consider a Markov chain on n points $\{0, 1, \ldots, n-1\}$ lying in order on a circle. At each step, the chain stays at the current point with probability $1/2$ or moves to the next point in the clockwise direction with probability $1/2$. Find the stationary distribution and show that, for any $\varepsilon > 0$, the mixing time $t_{\mathrm{mix}}(\varepsilon)$ is $O(n^2 \log(1/\varepsilon))$.

   Use the coupling in which you randomly choose $i \in \{1, 2\}$, you keep the i'th walk fixed while you move the other walk to the next point in the clockwise direction. The distance between the two walks increases or decreases by one with equal probability. The result follows from our hitting time analysis for the path.

4. Design a Markov chain on the independent sets of a graph where in the stationary distribution, $\pi_x \propto \lambda^{|I_x|}$ where I_x is the independent set corresponding to the state x and λ is a parameter.

   Solved in section 11.4.1 of MU. You can use the Metropolis-Hastings method discussed in class.

5. Consider an n-cube with $N = 2^n$ nodes. Let S be a nonempty set of vertices on the cube, and let x be a vertex chosen uniformly at random among all vertices of the cube. Let $D(x, S)$ be the minimum number of coordinates in which x and y differ over all points $y \in S$.

   (a) Show that

   $$\Pr[|D(x,S) - \mathbb{E}[D(x,S)]| \leq \lambda] \leq 2\exp\left(-\frac{2\lambda^2}{n}\right)$$

   S is fixed, so $D(x, S)$ is a function of the n independent coordinates of x. The function is 1-Lipschitz. So, the probability bound is $2e^{-2\lambda^2/n}$.

   (b) By plugging in $\lambda = \mathbb{E}[D(x,S)]$ into part (a), show that if $|S| \geq \alpha \cdot 2^n$, then show that:

   $$\mathbb{E}[D(x,S)] \leq c_\alpha \sqrt{n}$$

   where c_α is a constant depending on α.

   Plugging in $\lambda = \mathbb{E}[D(x,S)]$, we get that $\Pr[x \in S] \leq 2\exp\left(-\frac{2(\mathbb{E}[D(x,S)])^2}{n}\right)$. Solving, $\mathbb{E}[D(x,S)] \leq \sqrt{\frac{1}{2}\ln\frac{2}{\alpha}}$.