NATIONAL UNIVERSITY OF SINGAPORE
SCHOOL OF COMPUTING

Endterm Assessment for CS5330 – Randomized Algorithms

14 April 2020 Time Allowed: 180 minutes

INSTRUCTIONS:

- This paper consists of **THREE** questions.
- This is an **OPEN BOOK/NOTES** examination.
- Do **NOT** use the internet for help while taking the examination. Any evidence otherwise will result in a deduction of **at least 50%** for that problem.
- Do **NOT** collaborate while taking the examination. Any evidence otherwise will result in a deduction of **at least 50%** for that problem.
- Either type or hand-write your solutions. Please submit your solutions in PDF format to the LumiNUS submission folder before **10:00 PM**. The PDF file name should be your student number.

QUESTIONS:

1. (25 points)

   (a) (10 points) Give two functions $h_1, h_2 : \{1, 2, 3, 4\} \to \{0, 1\}$ such that $\{h_1, h_2\}$ is 2-universal.

   (b) (10 points) A random walk on a directed graph, at each step, chooses uniformly at random an outgoing edge from its current node and travels along it to the next node. Describe a directed graph G on n nodes for which the expected number of steps to reach a node v starting from another node u is $2^{\Omega(n)}$. You need not give a formal analysis; an informal justification is fine.

   (c) (5 points) Explain how we used the definition of total variation distance in the analysis for approximate counting of k-colorings done in class.

2. (55 points) Denote $[n] = \{1, \dots, n\}$, and let $\Omega = \{S \subseteq [n] : |S| = k\}$, subsets of size k from a universe of size n. Consider the following Markov chain $(S_0, S_1, \dots)$ where each $S_t \in \Omega$.

- With probability 1/2, let $S_{t+1} = S_t$.
- Otherwise, pick i uniformly from S_t, pick j uniformly from[1] $[n] \setminus S_t$, and let $S_{t+1} = S_t \cup \{j\} \setminus \{i\}$.

(a) (5 points) Show that this chain is irreducible and aperiodic. Therefore, it has a unique stationary distribution.

(b) (5 points) Show that the stationary distribution is uniform over Ω.

(c) (5 points) Consider the following joint process (X_t, Y_t).

- With probability 1/2, let $X_{t+1} = X_t, Y_{t+1} = Y_t$.
- Otherwise, let $A = X_t \setminus Y_t$ and $B = Y_t \setminus X_t$. Note that $|A| = |B|$. Fix an arbitrary bijection $\pi : A \to B$. Pick $i \in X_t$ uniformly at random and $j \in [n] \setminus X_t$ uniformly at random. Set $X_{t+1} = X_t \cup \{j\} \setminus \{i\}$. Define i' and j' as follows:
  - If $i \in X_t \cap Y_t$, then $i' = i$, else $i' = \pi(i)$.
  - If $j \notin Y_t$, $j' = j$, else $j' = \pi^{-1}(j)$.

  Set $Y_{t+1} = Y_t \cup \{j'\} \setminus \{i'\}$.

Show that (X_t, Y_t) is a valid coupling for the Markov chain (S_t).

(d) (10 points) Let $D_t = \{i : i \in X_t \setminus Y_t \text{ or } i \in Y_t \setminus X_t\}$, and let $q_t = |D_t|$. In the next two parts, you will be asked to show that for $k \geq 2$:

$$\mathbb{E}[q_{t+1} \mid D_t] \leq \left(1 - \frac{1}{k}\right) q_t.$$

Why does this imply that $t_{\text{mix}}(\varepsilon) = O(k \log(k/\varepsilon))$ for the Markov chain (S_t)?

(e) (20 points) With probability $\frac{1}{2}$, $X_{t+1} = X_t, Y_{t+1} = Y_t$ and so, $q_{t+1} = q_t$. Now suppose the coupling transitions using the "otherwise" step above. Analyze $q_t - q_{t+1}$ in each of the following cases:

(i) $i \in X_t \cap Y_t, j \in Y_t \setminus X_t$.
(ii) $i \in X_t \setminus Y_t, j \in [n] \setminus (X_t \cup Y_t)$.
(iii) $i \in X_t \setminus Y_t, j \in Y_t \setminus X_t, j \neq \pi(i)$.
(iv) All other cases

(f) (10 points) Conclude that $\mathbb{E}[q_{t+1} - q_t \mid D_t] = -\frac{n-2}{k(n-k)} q_t \leq -\frac{1}{k} q_t$, where for the last inequality, we assumed $k \geq 2$.

[1] $S \setminus T$ denotes the set difference between S and T, i.e., elements contained in S but not in T.

3. (20 points) You are in front of a large playpen filled with R red balls and G green balls. You uniformly sample n balls without replacement. In other words, you draw the first ball uniformly from the playpen and remove it, the second ball uniformly from the remaining balls in the playpen and remove it, and so on for n times.

   (a) (10 points) For $i \geq 0$, let R_i be the number of red balls remaining after i draws. So, $R_0 = R$. Let $r_i = R_i/(R+G-i)$ be the fraction of red balls in the playpen after i steps. Show that the sequence $r_0, \ldots, r_n$ forms a martingale with respect to $R_0, \ldots, R_n$.

   (b) (10 points) Show that:
   $$\Pr\left[\left|r_n - \frac{R}{R+G}\right| > \lambda\right] \leq 2e^{-\lambda^2/2n}.$$
   Conclude that the number of balls removed concentrates around $nR/(R+G)$.