Philip Wellnitz, Barış Can Esmer **Winter 2022/23**

– Randomized Algorithms and Probabilistic Analysis of Algorithms, Exercise Sheet 4/5 –

https://www.mpi-inf.mpg.de/departments/algorithms-complexity/teaching/winter22/random

Total Points: 100 + 100 Due: Wednesday, **January 11**, 2023

You are allowed to collaborate on the exercise sheets. Justify your answers. Cite all external sources that you use (books, websites, research papers, etc.). You need to collect at least 50% of all points on exercise sheets to be admitted to the exam. Please hand in your solutions before the lecture on the day of the deadline.

Exercise 4.1 — **20** points

For $0 < \varepsilon, \delta < 1$, demonstrate a randomized algorithm that gives an (ε, δ)-approximation of π.

(Hint: Sample a random point from the square $(-1,-1)(1,-1)(1,1)(-1,1)$ *and consider the event that said sample lies inside of the unit circle.)*

Exercise 4.2 — **10 + 20** points

Your favorite game center offers a new game where in each round, you can equally likely win 1€ or loose 1€. You start with an initial budget of 200€ and you may play as long as your balance is positive (that is, you can play as long as you are not bankrupt).

a) What is the probability that you win 200€ before going bankrupt?

b) For positive ℓ_1, ℓ_2, what is the probability that you own 200€ $+ \ell_1$€ before you own 200€ $- \ell_2$€?

Exercise 4.3 — **10** points

Show that the cover time of the complete graph on n vertices is $\Theta(n \ln n)$.

Exercise 4.4 — **10** points

Consider a finite, irreducible, and aperiodic Markov chain with transition matrix $\mathbb{P}$. Show that if there are nonnegative numbers $\bar{\pi} = (\pi_1, \ldots, \pi_n)$ that sum to 1 and satisfy for any i, j

$$\pi_i P_{i,j} = \pi_j P_{j,i}, \tag{1}$$

then $\bar{\pi}$ is the stationary distribution corresponding to $\mathbb{P}$.

Note: A Markov chain that satisfies (1) is also called *time-reversible*.

Exercise 4.5 — **10 + 15 + 5** points

Consider a finite Markov chain with n states, transition matrix $\mathbb{P}$, and stationary distribution $\bar{\pi}$. Imagine running the chain for r steps, yielding the sequence $X_0, X_1, \ldots, X_r$ and consider the reverse sequence $X_r, \ldots, X_0$.

a) Argue that the reverse sequence is Markovian.

b) Argue that for the reverse sequence, the transition probabilities $Q_{i,j}$ are given by

$$Q_{i,j} = \frac{\pi_j P_{j,i}}{\pi_i}.$$

c) Argue that for a time-reversible Markov chain, we have $P_{i,j} = Q_{i,j}$ for all states i, j.

In this exercise sheet, we develop several techniques to prove the *existence* of objects. Collectively, these methods are also called the *probabilistic method*.

Exercise 5.1 — **5 + 5 + 5 + 5** points

In this exercise, we show that for integers n and k that satisfy $\binom{n}{k}2^{-\binom{k}{2}+1} < 1$, we can color the edges of the complete graph K_n with two colors such that no complete subgraph K_k is monochromatic.

a) Define a sample space Ω of all valid colorings of K_n with two colors. What is the size of said sample space?

b) Compute the probability that in a coloring sampled uniformly from Ω, a specific k-clique K_k is monochromatic.
(Hint: Think of uniformly sampling a coloring from Ω as assigning each edge one of the two colors uniformly at random.)

c) Assuming $\binom{n}{k}2^{-\binom{k}{2}+1} < 1$, conclude that the probability that all k-cliques in K_n are monochromatic is less than 1. Further conclude that hence, there is a coloring of K_n that ensures that no k-clique subgraph is monochromatic.

d) Discuss how to (efficiently) construct such a coloring of K_n.

Exercise 5.2 — **10 + 10** points

In this exercise, we see an averaging argument that is sometimes easier to apply.

a) Prove that for a probability space $\mathcal{S}$ and a random variable X defined on $\mathcal{S}$ with $\mathbb{E}[\,X\,] = \mu$, we have $\Pr[\,X \geq \mu\,] > 0$ and $\Pr[\,X \leq \mu\,] > 0$.

b) Prove that in any undirected graph $G = (V, E)$ with m edges, there is a partition of V into disjoint sets A and B such that at least $|E|/2$ edges connect a vertex in A to a vertex in B.
(Hint: Assign each vertex to A or B uniformly at random and compute the expected number of edges crossing a cut constructed in this way.)

Exercise 5.3 — **10 + 10** points

It turns out that it is sometimes easier to first construct an intermediate structure and then modify it to have the desired properties. In this exercise, we see an example of this "sample and modify" paradigm.

Given a connected graph $G = (V, E)$ with $|E| \geq |V|/2$, write $d := 2|E|/|V|$ to denote the average degree of G. Consider the following randomized algorithm for obtaining an independent set in G.

- Delete each vertex v of G (and all edges incident to v independently with probability $1 - 1/d$.
- For each remaining edge, remove the edge and one of the vertices adjacent to the edge (arbitrarily picking the vertices to delete).

a) Compute the expected number of vertices and edges that remain in G after the first step of the algorithm.

b) Compute the expected size of the independent set that remains in G after the second step of the algorithm. Conclude that any graph with $|E| \geq |V|/2$ edges has an independent set of size at least $|V|/4|E|$ vertices.

Exercise 5.4 — 10 + 10 + 10 points

In this exercise, we prove (a special case of) the Lovász Local Lemma, which is a useful tool for applying the probabilistic method. Intuitively, we are concerned with a set of bad events $E_1, \dots, E_n$ (in some probability space) and wish to show that with positive probability, none of the events occur. The catch is, that the events E_i need not be independent (but the dependency is limited).

Formally, we say an event E_{n+1} is *mutually independent* of the events $E_1, \dots, E_n$ if for any subset $I \subseteq \{1, \dots, n\}$, we have

$$\Pr\left[E_{n+1} \mid \bigwedge_{j \in I} E_j \right] = \Pr[\, E_{n+1} \,] .$$

Next, we define the *dependency graph* of the events $E_1, \dots, E_n$ as follows: there is a vertex for i for each event E_i, and two vertices are connected by an edge if they are not mutually independent. The degree of the dependency graph is the maximum degree of any of the vertices of this graph.

We proceed to prove the following result.

Theorem 1 (Lovász Local Lemma)**.** *Suppose $E_1, \dots, E_n$ are events that satisfy all of the following*

- *for all i, we have $\Pr[\, E_i \,] \le p$,*
- *the degree of the dependency graph of $E_1, \dots, E_n$ is bounded by d, and*
- *we have $4dp \le 1$.*

Then, $\Pr[\, \bigwedge_{i=1}^{n} \bar{E}_i \,] > 0$.

Consider a subset $S \subseteq \{1, \dots, n\}$; we proceed to prove Theorem 1 by induction on an upper bound s on the size of $|S|$, that is, for $s = 0, \dots, n-1$, we show that all $k \notin S$ satisfy

$$\Pr\left[E_k \mid \bigwedge_{j \in S} \bar{E}_j \right] \le 2p. \tag{2}$$

a) Prove that for nonempty S, we have $\Pr\left[\bigwedge_{j \in S} \bar{E}_j \right] > 0$. Also argue about an induction base.

b) Consider the set $S_1 := N(k) \cap S$ of of all neighbors of k in the dependency graph that are also in S and write $S_2 := S \setminus S_1$. If $S_2 = S$, show that $\Pr\left[E_k \mid \bigwedge_{j \in S} \bar{E}_j \right] \le p$.

c) For $|S_2| < s$, define $F_S := \bigwedge_{j \in S} \bar{E}_j$ and analogously define F_{S_1} and F_{S_2}. Prove that $\Pr[\, E_k \wedge F_{S_1} \mid F_{S_2} \,] \le p$ and $\Pr[\, F_{S_1} \mid F_{S_2} \,] \ge 1/2$. Conclude the inductive step and conclude Theorem 1.

Exercise 5.5 — 10 points

In the k-SAT problem, we are given a Boolean formula in conjunctive normal form where each clause has exactly k literals. Using the Lovász Local Lemma, prove that if in a k-SAT formula, no variable appears in more than $T = 2^k/4k$ clauses, then the formula has a satisfying assignment.

(Hint: Consider the probability space of all random assignments to variables and consider the events that a fixed clause is not satisfied. Finally, observe that the degree of the dependency graph between the events is bounded and apply the Lovász Local Lemma.)