Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute of Theoretical Computer Science
Bernd Gärtner, Mohsen Ghaffari, Rasmus Kyng, David Steurer

Algorithms, Probability, and Computing **Midterm Exam** **HS20**

Candidate

First name: ....

Last name: ....

Student ID (Legi) Nr.: ....

I attest with my signature that I was able to take the exam under regular conditions and that I have read and understood the general remarks below.

Signature: ....

General remarks and instructions

1. Check your exam documents for completeness (6 two-sided pages with 5 exercises).
2. You have **2 hours** to solve the exercises.
3. You can solve the exercises in any order. You should not be worried if you cannot solve all exercises! Not all points are required to get the best grade.
4. Immediately inform an assistant in case you are not able to take the exam under regular conditions. Later complaints will not be accepted.
5. Pencils are not allowed. Pencil-written solutions will not be reviewed.
6. No auxiliary material allowed. All electronic devices must be turned off and are not allowed to be on your desk. We will write the current time on the blackboard every 15 minutes.
7. Attempts to cheat lead to immediate exclusion from the exam and can have judicial consequences.
8. Provide only one solution to each exercise. Please clearly mark/scratch the solutions that should not be graded.
9. **All solutions must be understandable and well-founded. Write down the important thoughts in clear sentences and keywords. Unless stated otherwise, no points will be awarded for unfounded or incomprehensible solutions.**
10. You may use anything that has been introduced and proved in the lecture or in the exercise sessions. You do not need to re-prove it, but you need to state it clearly and concretely. However, if you need something *different* than what we have shown, you must write a new proof or at least list all necessary changes.
11. Write your student-ID (**Legi-number**) on **all** sheets (and your name only on this cover sheet).

	achieved points (maximum)	reviewer's signature
1	(14)	
2	(12)	
3	(12)	
4	(12)	
5	(10)	
Σ	(60)	

Exercise 1: Five Short Questions (14 points)

No justification is required.
(There are no negative points for wrong answers.)

(a) Let G denote an n-node weighted graph with distinct edge weights. Assume that we sample each edge in G independently with probability $p = 1/n^{3/4}$ and we call the resulting graph G'. Give an asymptotic upper bound (which is tight in the worst-case) on the expected number of edges in G that are not T'-heavy, where T' denotes the MSF of G'. (3 points)

Answer: ....

(b) Let G denote an n-node simple, unweighted and undirected graph with minimum cut size $k \geq 1$. Assume that we sample each edge in G independently with probability $p = 1/k$ and we denote the sampled set of edges by E'. Now, we consider an arbitrary order of the edges in E', and we contract the corresponding edges sequentially in that order (omitting edges whose corresponding edge is not present anymore). Among the four answers, choose the asymptotically tightest lower bound for the probability that the minimum cut size of G is equal to the minimum cut size of the contracted graph. (3 points)
a) $\Omega(1/n^2)$ b) $\Omega(1/n)$ c) $\Omega(1/\log(n))$ d) $1/1000$.
During the midterm, the following announcement was made:
1.) You can assume $k \geq 2$
2.) One should not treat k as a constant
3.) $1 - x \geq e^{-2x}$ for $0 \leq x \leq 0.5$

Answer: ....

(c) Given a set of n points in the plane, we have devised a data structure that can answer the following query: Given a non-vertical query line, report all of the k points below the query line. What is the storage space and query time for the data structure? (3 points)

Storage space: ....

Query time: ....

(d) Consider a set S of n non-crossing segments in general position and let $\mathcal{T}(S)$ denote its trapezoidal decomposition. Let S' denote the set obtained from S by removing a segment chosen uniformly at random from S. What is the expected number (asymptotically) of trapezoids that are contained in $\mathcal{T}(S)$, but not in $\mathcal{T}(S')$.
(3 points)

Answer: Θ ....

(e) An n-dimensional ellipsoid is a set of the form $E = ...$ (2 points)

Answer: ....

Exercise 2: Approximate Minimum Cuts in Hypergraphs (12 points)

Devise a (randomized) algorithm that takes as input a multihypergraph G of rank at most 4 (i.e. every edge joins at most 4 vertices) and outputs a random cut of G. Your algorithm should output each 2-approximate minimum cut C of the graph G with probability $\Omega(1/n^8)$. A 2-approximate minimum cut of G is a cut that contains at most twice as many edges as the minimum cut size of G.

You only need to analyze the correctness, but not the runtime of your algorithm.

Exercise 3: Random Binary Search Tree (12 points)

For a random binary search tree with n distinct keys, we are interested in the distance between the smallest key and the second smallest key, i.e., the number of edges along the path between the two corresponding nodes. For $n \geq 2$, let X_n denote the random variable for the distance between the smallest and the second smallest key and let $x_n := \mathbb{E}[X_n]$. The goal of this exercise is to show that x_n can be upper bounded by a constant independent of n.

Part (a). (6 points) Devise a recursive formula for x_n for every $n \geq 2$.
Hint: The expected height of the smallest key in a random binary search tree with n keys is $H_n - 1$, where $H_n := \sum_{i=1}^{n} \frac{1}{i}$.

Part (b). (6 points) Derive a closed-form solution for x_n for every $n \geq 2$. Simplify as much as possible!

Exercise 4: Point Location (12 points)

Let $\mathcal{H}$ denote a collection of $n \geq 4$ (hyper)planes in $\mathbb{R}^3$, i.e., each $h \in \mathcal{H}$ is a subset of $\mathbb{R}^3$ that can be expressed as $h = \{x \in \mathbb{R}^3 \colon a^\mathsf{T} x = b\}$ for some $a \in \mathbb{R}^3$ and $b \in \mathbb{R}$, where a is not the all-zero vector.

We assume that every three distinct hyperplanes contained in $\mathcal{H}$ intersect at a single point and there don't exist four distinct hyperplanes in $\mathcal{H}$ with non-empty intersection. A point $p \in \mathbb{R}^3$ is called an intersection point of $\mathcal{H}$ if it is contained in the intersection of three distinct hyperplanes in $\mathcal{H}$. We additionally assume that the pairwise distances of intersection points of $\mathcal{H}$ are distinct.

We now consider a random ordering of the hyperplanes in $\mathcal{H}$, i.e., each of the $n!$ possible orderings is equally likely.

For $i \geq 4$, we denote with $\mathcal{H}_i$ the collection that contains the first i hyperplanes in the random ordering. For each $i \in \{4, 5, \ldots, n\}$, let d_i denote the distance between the two closest intersection points of $\mathcal{H}_i$. We define $D := \{d_i \colon i \in \{4, 5, \ldots, n\}\}$.

Prove that there exists some constant c such that the expected size of the set D is at least $4H_n - c$ and at most $6H_n$.

Exercise 5: Linear Programming (10 points)

Assume that the following LP has optimal value Z.

$$\begin{array}{lrcl} \text{minimize} & x_1 & & \\ \text{subject to} & 2x_1 + x_2 & \geq & 4 \\ & -x_2 & \geq & -4 \\ & x_1 - x_2 & \geq & 1 \\ & x_1, x_2 & \geq & 0 \end{array} \tag{1}$$

What is the optimal value of the following LP? Provide a proof!

$$\begin{array}{lrcl} \text{maximize} & -4y_1 + y_2 + 4y_3 & & \\ \text{subject to} & -y_1 - y_2 + y_3 & \leq & 0 \\ & y_2 + 2y_3 & \leq & 1 \\ & y_1, y_2, y_3 & \geq & 0 \end{array} \tag{2}$$