ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute of Theoretical Computer Science
Bernd Gärtner, Rasmus Kyng, Angelika Steger, David Steurer

Algorithms, Probability, and Computing — Midterm Exam — HS24

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 (8 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	(10)	
2	(10)	
3	(10)	
4	(10)	
5	(10)	
Σ	(50)	

Exercise 1: Short Questions (10 points)

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

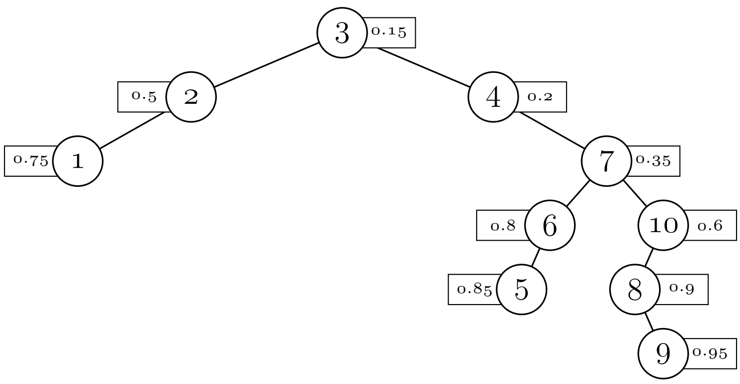
(a) (2 points)

- Draw the binary search tree obtained by inserting the keys 1, 2, 3, 4, 5, 6, and 7 into an initially empty BST in that order.
  Answer:

- Draw the binary search tree with the smallest possible height containing the keys 1, 2, 3, 4, 5, 6, 7.
  Answer:

(b) (2 points)

Consider the treap shown below. We remove node 4. How many rotations are performed during the removal of node 4?

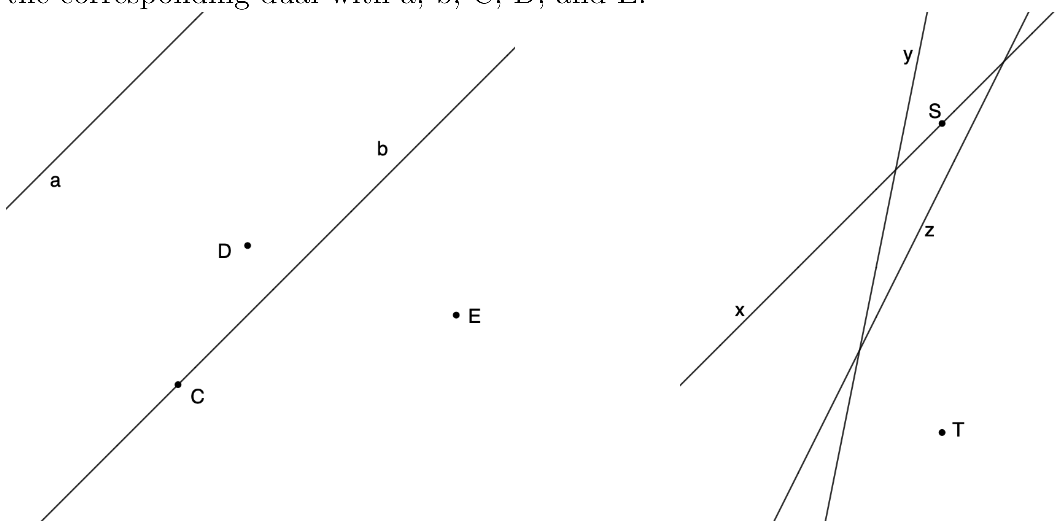


Answer: ....

(c) (2 points) Let $n > 0$ and $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$ be n points in the plane. Suppose that each point has a different distance from the origin. We add the n points in a uniformly random order. What is the expected number of times that the closest point to the origin changes during the process? The answer can be expressed using the $O(\cdot)$ notation.

Answer: ....

(d) (2 points) The picture on the right shows the dual of the picture on the left. Fill in the corresponding dual with a, b, C, D, and E.



Answer: S:......; T:......; x:......; y:......; z:......

(e) (2 points) Suppose that a linear program P has a feasible solution x. Let D be the dual program of P. Then

- D always have a feasible solution; [ ] False [ ] True
- The optimal solution of D always exists and has the same cost as the optimal solution of P. [ ] False [ ] True

Exercise 2: Randomized MST (10 points)

Let us consider a variation of the randomized MST algorithm from the lecture that performs $i \geq 3$ iterations of Borůvka and selects edges with some constant probability $0 < c < 1$ instead of $\frac{1}{2}$. Note that i and c do not depend on m or n.

> RANDOMIZED MINIMUM SPANNING TREE ALGORITHM(G):
> Perform i iterations of Borůvka's algorithm
> In the new graph:
>   Select edges with probability c and compute recursively
>   an MSF for the graph consisting of the selected edges.
>   Call this forest T.
>   Use FINDHEAVY to find all unselected edges that are not T-heavy.
>   Add all edges that are *not* T-heavy to T and delete all other edges.
> recurse (until the graph contains only one vertex)

(a) Prove that the upper bound on the expected number of edges found by FINDHEAVY is $\frac{1-c}{c}\frac{n}{2^i}$.

(b) Prove that $\forall c \in (\frac{1}{2^i-2}, 1)$ the algorithm has *expected* runtime of $O(m+n)$.

Hint: You can use the following result without proof.

Proposition 1. *Let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of independent Bernoulli random variables with $\Pr[X_i = 1] = p$, $\forall i \in \mathbb{N}$. For an integer $k \geq 0$, let T be the random variable of the (first) point in time where we have seen the k-th 0. Formally $T := \min\{n \in \mathbb{N} |\ |\{i \leq n | X_i = 0\}| \geq k\}$.*
Let Q be the number of ones in the sequence up to T, formally $Q := \sum_{i=1}^{T} X_i$. Then

$$\mathrm{E}[Q] = k\frac{p}{1-p}.$$

Exercise 3: Random Binary Search Tree (10 points)

(a) Prove that, in a random binary search tree with $n \geq 2$ nodes, the probability that the node with rank r is a leaf is $\frac{1}{2}$ if $r = 1$ or $r = n$ and $\frac{1}{3}$ otherwise.

(b) Compute the expected number of leaves in a random binary search tree. (You can assume that the statement in part (a) is true even if you did not prove it.)

Exercise 4: Point Location (10 points)

You are given a set of n line segments L on the plane and a fixed point x. The segments in L do not intersect each other. For a given query point q, you need to determine if q can "see" x. In other words, we want to check if the line segment connecting x and q intersects any of the line segments in L.

- If the line segment between x and q does not intersect any segments in L, the algorithm should return "Yes" to indicate that q can see x.
- If there is an intersection with one or more segments in L, the algorithm should return "No" along with one of the intersecting segments from L as a "certificate" of the obstruction.

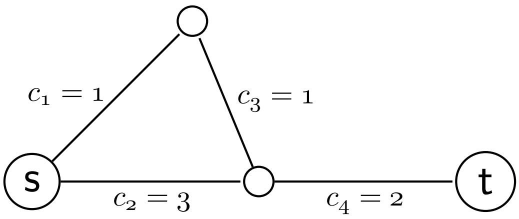
To attain a full score, your algorithm should preprocess the data in $O(n^2)$ time, and each query should be answered in $O(\log n)$ time. Assume that among the endpoints of the segments, the fixed point x, and any query points, no three points are collinear.

Exercise 5: Linear Programming (10 points)

We are given a network represented by a connected graph $G = (V, E)$. In the network, there is a special node $s \in V$ that is called the source that produces an amount f of outflow and a node $t \in V \setminus \{s\}$ that can absorb an unlimited amount of flow. Finally, for each edge $e_j \in E$, we know the cost c_j of the transportation of one flow unit on edge e_j (in either direction). We are interested in finding the flow that routes all the flow produced by s with minimum cost.

(a) Formulate an LP problem that finds the cost of the best flow in a general network.

Consider now the following network with the costs reported in the picture and suppose that we want to route an amount of flow $f = 5$ from node s to node t.



(b) Formulate an LP problem that finds the cost of the best flow for this particular network.

(c) Find all the basic feasible solutions of the LP of part (b). Find the optimal solution of the LP of part (b).