ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute of Theoretical Computer Science
Mohsen Ghaffari, Angelika Steger, David Steurer, Emo Welzl, Peter Widmayer

Algorithms, Probability, and Computing **Special Assignment 2** **HS18**

- The solution is due on **Tuesday, December 4, 2018** by **2:15 pm**. Please bring a print-out of your solution with you to the lecture. If you cannot attend (**and please only then**), you may alternatively email your solution as a PDF, likewise **until 2:15 pm** to the head assistant, Ahad N. Zehmakan. We will send out a confirmation that we have received your file. Make sure you receive this confirmation within the day of the due date, otherwise complain timely.
- Write an exposition of your solution using a computer, where we strongly recommend to use LaTeX. **We do not grade hand-written solutions.**
- For geometric drawings that can easily be integrated into LaTeX documents, we recommend the drawing editor IPE, retrievable at `http://ipe7.sourceforge.net/` in source code and as an executable for Windows.
- Write short, simple, and precise sentences.
- This is a theory course, which means: if an exercise does not explicitly say "you do not need to prove your answer" or "justify intuitively", then a formal proof is **always** required. You can of course refer in your solutions to the lecture notes and to the exercises, if a result you need has already been proved there.
- We would like to stress that the ETH Disciplinary Code applies to this special assignment as it constitutes part of your final grade. The only exception we make to the Code is that we encourage you to verbally discuss the tasks with your colleagues. It is strictly prohibited to share any (hand)written or electronic (partial) solutions with any of your colleagues. We are obligated to inform the Rector of any violations of the Code.
- As with all exercises, the material of the special assignments is relevant for the final exam.

Exercise 1 (20 points)

As input, we are given an undirected graph $G = (V, E)$, and edge cost $c_e \geq 0$ for each $e \in E$, a selected root vertex $r \in V$, and a penalty $\pi_v \geq 0$ for each $v \in V$. We want to find a tree T, as a subgraph of G, which contains the root vertex r and has the minimum cost. The cost of a tree T is the cost of the edges in T plus the penalties of all vertices not in T. In other words, the goal is to minimize $\sum_{e \in E(T)} c_e + \sum_{v \in V \setminus V(T)} \pi_v$, where $V(T)$ and $E(T)$ are the set of vertices and edges in the tree T.

To model this problem as an integer linear program, we define a 0-1 variable y_v for each $v \in V$ and x_e for each $e \in E$. Variable y_v is set to 1 if v is in the solution tree and 0 if it is not, while x_e is set to 1 if e is in the solution tree and 0 otherwise.

(a) **(5 points)** Justify that the optimal value of the following integer linear programming is equivalent to the minimum cost of a tree which contains the root r in the aforementioned problem.

$$\begin{array}{lrl} \text{minimize} & \sum_{e \in E} c_e x_e + \sum_{v \in V}(1 - y_v)\pi_v & \\ \text{subject to} & \sum_{e \in \delta(S)} x_e \geq y_v & \forall S \subseteq V \setminus \{r\}, S \neq \emptyset, \forall v \in S, \\ & y_r = 1, & \\ & y_v \in \{0,1\} & \forall v \in V \\ & x_e \in \{0,1\} & \forall e \in E. \end{array}$$

(b) **(5 points)** Relax the integer programming formulation from (a) and prove that you can apply the ellipsoid method to solve the resulting linear program in polynomial time.

Hint: You might define a separation oracle building on the network flow problem.

(c) **(5 points)** Given an optimal solution $(\mathbf{x}^*, \mathbf{y}^*)$ to the linear programming relaxation and some value $\alpha \in [0, 1)$, define the set $U := \{v \in V : y_v^* \geq \alpha\}$. Let tree T be a subgraph of G which includes all vertices in U and has the minimum possible cost. Prove that

$$\frac{1}{1-\alpha} \sum_{v \in V} \pi_v (1 - y_v^*) \geq \sum_{v \in V \setminus V(T)} \pi_v.$$

(d) **(5 points)** Assume that one can find the tree T from part (c) in polynomial time and the following inequality holds for T

$$\sum_{e \in E(T)} c_e \leq \frac{2}{\alpha} \sum_{e \in E} c_e x_e^*.$$

By selecting a suitable α and combining this statement and part (c) prove that this provides us with a polynomial time 3-approximation algorithm for the above problem, i.e., the output will not be more than a factor 3 times the value of a minimum solution.

Exercise 2 (15 points)

(a) **(10 points)** Assume that we have a collection of m different kinds of raw materials. Specifically, we have a_j units of raw material of kind j. There are n different kinds of products that we can generate from these raw materials. Each unit of product i takes b_{ij} units of raw material j to create and can be sold at the price p_i. It is possible to sell fractional amounts of any product. We want to produce the most profitable set of products with our available materials. Formulate this optimization problem as a linear program and write its dual.

(b) **(5 points)** Let $\mathbf{x}^* = (x_1^*, \cdots, x_n^*)$ be a feasible solution of the linear program

$$\text{maximize } \mathbf{c}^T \mathbf{x} \text{ subject to } A\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \quad (\text{P})$$

and let $\mathbf{y}^* = (y_1^*, \cdots, y_m^*)$ be a feasible solution of the dual linear program

$$\text{minimize } \mathbf{b}^T \mathbf{y} \text{ subject to } A^T\mathbf{y} \geq \mathbf{c},\ \mathbf{y} \geq \mathbf{0} \quad (\text{D}).$$

Prove that the following two statements are equivalent:

(1) $\mathbf{x}^*$ is optimal for (P) and $\mathbf{y}^*$ is optimal for (D).

(2) For all $i = 1, \cdots, m$, $\mathbf{x}^*$ satisfies the i-th constraint of (P) with equality or $y_i^* = 0$; similarly, for all $j = 1 \cdots, n$, $\mathbf{y}^*$ satisfies the j-th constraint of (D) with equality or $x_j^* = 0$.

Exercise 3 (25 points)

(a) **(15 points)** Assume that Hansel and Gretel each have an n-bit binary number. Gretel would like to know whether they both have the same number or not. Thus, she asks Hansel to send her his number by a letter. However, Hansel prefers not to write the whole number since it is quite long. Provide them with a randomized strategy so that Hansel only needs to send $\mathcal{O}(\log n)$ bits, but Gretel is still able to figure out the answer to her question with probability at least $1 - \mathcal{O}(\frac{1}{n^2})$. The strategy does not need to be computationally efficient since both Hansel and Gretel have unlimited computational power.

You might use the two following propositions, without proving them.

Proposition 1: The number of distinct prime divisors of any number less than 2^n is at most n.

Proposition 2: For any integer $k \geq 2$, the number of distinct primes less than k is $\Omega(\frac{k}{\log k})$.

Hint: You might ask Hansel to pick a prime number p uniformly at random from a suitable range and then divide his number by p.

(b) **(10 points)** Let $A \in \mathsf{GF}(2)^{m \times n}$ and $b \in \mathsf{GF}(2)^m$. We know that the set $\{x \in \mathsf{GF}(2)^n \mid Ax = b\}$ has either size 0 or 2^i for some $i \in \mathbb{N}_0$. Building on this statement, prove that for any graph G, the number of Pfaffian orientations $\overrightarrow{G}$ is either 0 or 2^i for some $i \in \mathbb{N}_0$.