MALMÖ UNIVERSITY
Computer Science

Advanced Programming
Autumn 2021

# Advanced Programming Exam

**October 10, 2021. 8.15–13.15**

**Instructions:** Treat at most one question per sheet, write only on one side of each sheet, number and mark each page with your initials. **Do not use red ink!** Write legibly and comment your work extensively, this can give you extra marks even if the end result is wrong. The exam contains 5 question and each question is worth 10 points.

**Allowed Material:** The course books (Dasgupta, Papadimitriou, Vazirani. *Algorithms*; Skiena, Revilla. *Programming Challenges*; Bentley. *Programming Pearls*), course notes and other printed material, pen or pencil and paper.

**Hand-in:** Answers should arrive in my email at `bengt.nilsson.TS@mau.se` as *one* document at the latest 13:15. Allowed document formats are: word, `.txt` (if you trust me) or `.jpg`, `.png`, PDF/A (if you don't). If you have questions during the exam, you can contact me via email. **Good Luck!**

---

1. Given two sets $S_1$ and $S_2$, each containing $n$ numbers, and a number $x$, devise an algorithm for finding whether a number from $S_1$ and a number from $S_2$ sum to $x$. Your algorithm should take $O(n \log n)$ time. (10p)

2. Superman can see through objects made of different materials using x-ray vision. However, the amount of energy he requires varies according to the opacity of the material, e.g., steel requires more energy than wood. In addition, the amount of energy he must use is proportional to his distance from the object, the farther away, the more energy he must use.

   Along a straight coastline, $n$ bunkers made of different materials have been built. Each bunker is placed at the one-dimensional coordinate position $p_1, p_2, \ldots, p_n$, with $p_i < p_{i+1}$, along the coastline so you can assume that the bunkers are given in the order they appear along the coastline. Assume further that you have $w_i$, $1 \leq i \leq n$, the amount of energy (a non-zero, positive value) superman needs to see into the $i^{\text{th}}$ bunker if he stands next to it.

   Find an algorithm that computes the point $s$ along the coastline for which superman can exert the minimum amount of energy to see into all the bunkers, i.e., $s$ is the position for which
   $$\max_{1 \leq i \leq n} \{|s - p_i| + w_i\}$$
   is minimized. Your algorithm should take linear time. (10p)

3. The $n$th Catalan number is defined for any non-negative integer $n$ as:
   $$C(n) = \begin{cases} 1 & \text{if } n = 0, \\ \sum_{k=0}^{n-1} C(n-1-k) \cdot C(k) & \text{if } n > 0. \end{cases}$$

   The $n$th Catalan number represents the number of possible sequences of $n$ pairs of correctly matching parentheses (for $n = 2$, the possibilities are `(())` and `()()` so in this case $C(2) = 2$) or equivalently the number of possible binary search trees with $n$ nodes.

Please turn page!

(a) Construct an efficient algorithm that computes $C(n)$ with the use of dynamic programming. (5p)

(b) Analyze the time complexity of your algorithm in terms of $n$. (2p)

(c) Use your algorithm to compute $C(4)$. (3p)

4. An amateur lepidopterist (someone who studies butterflies) friend of yours has collected a set of $n$ butterflies that your friend believes come from two different species, $A$ and $B$. It is very hard to directly label a specimen as belonging to one of $A$ or $B$ but it is possible to carefully compare two specimens to see if they are the *same* or *different*.

   After having performed $m$ comparisons of pairs of butterfly specimens (having outcome either *same* or *different*), your friend wishes to know if the results are consistent with the hypothesis that they come from two species. Design an algorithm that determines whether the $m$ comparisons are consistent with the hypothesis or not. Your algorithm should take $O(n+m)$ time. (10p)

   (Hint: Consider the graph with specimens as nodes and edges between a pair if they are different. What property does the graph obey if specimens come from two species?)

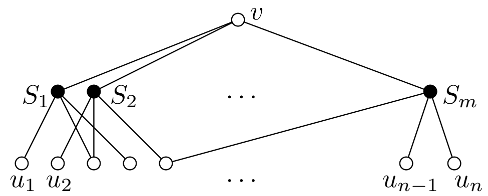
5. A variant of the GST problem can be stated formally as follows.

   INPUT: an undirected weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ with vertices $\mathcal{V}$, edges $\mathcal{E}$, and edge weights $w$. We are also given a subset $\mathcal{T} \subseteq \mathcal{V}$ of the vertices called the *terminals* and a positive integer $t$.
   OUTPUT: a subtree of weight at most $t$ that spans all the terminal vertices in $\mathcal{T}$, if such a tree exists.

   The following problem, denoted Set Cover, is known to be NP-hard.

   INPUT: a set $\mathcal{U}$ of items called the *universe*, $\mathcal{U} = \{u_1, \dots, u_n\}$, a family $\mathcal{F} = \{S_1, \dots, S_m\}$ consisting of subsets of the universe, $S_j \subseteq \mathcal{U}$, and a positive integer $k$.
   OUTPUT: a subset $\mathcal{R} \subseteq \mathcal{F}$ such that $|\mathcal{R}| \leq k$ and $\bigcup_{S \in \mathcal{R}} S = \mathcal{U}$, if such a subset exists.

   We can construct a polynomial time reduction Set Cover $\longrightarrow_P$ GST to prove that our simplified variant of GST is NP-hard by constructing a graph from a set cover instance as follows:




   $\mathcal{G}$ is constructed so that it has a vertex for each item $u_i$ in the set cover universe, one for each subset $S_j$ in the set cover family of subsets, plus an extra vertex $v$. The terminals, the subset $\mathcal{T}$, are chosen to be those vertices that correspond to items $u_i$ in the set cover universe and the vertex $v$, indicated with white centers in the figure above.

   The edges of $\mathcal{G}$ are constructed so that there is one edge from $v$ to each vertex corresponding to a subset $S_j$ in the set cover family and there is an edge from a vertex corresponding to a subset $S_j$ to a vertex corresponding to an item $u_i$ iff $u_i \in S_j$ in the set cover instance.

   What should be the values of the non-negative edge weights (they are not all the same) and what should be the positive value $t$ in order to finalize the proof? Provide arguments for your choices! (10p)