# CSE 421 Spring 2025: Set 1

Instructor: Chinmay Nirkhe
Due date: April 9th, 2025 11:59pm

**Instructions:** Solutions should be legibly handwritten or typeset (ideally in LaTeX). Mathematically rigorous solutions are expected for all problems unless explicitly stated.

You are encouraged to collaborate on problems in small teams but everyone must individually submit solutions. Solutions for the problems may be found online or in textbooks – but do not use them.

For grading purposes, list, with each problem, the names of your collaborators. Please start each problem on a new page.

**Problem 1** (Big-O notation)**.**

1. **[5 points]** Prove or disprove the following statement: There exist functions $f, g : \mathbb{N} \to \mathbb{N}$ such that $f$ is not $O(g)$ *and* $g$ is not $O(f)$. If it is true, give an example. If it is false, prove that no such functions exist.

2. **[5 points]** Let $f_1, f_2, \ldots, f_k : \mathbb{N} \to \mathbb{N}$ be a set of functions. Show there exists a function $g : \mathbb{N} \to \mathbb{N}$ with the following properties:

   (a) $f_i$ is $O(g)$ for each $i = 1, \ldots, k$

   (b) For any function $g' : \mathbb{N} \to \mathbb{N}$ such that $f_i$ is $O(g')$ for each $i = 1, \ldots, k$, it follows that $g$ is $O(g')$.

   In other words, there exists an asymptotic least upper bound for any finite collection of functions.

**Problem 2** (Probabilistic argument)**. [8 points]** Consider an undirected graph $G = (V, E)$ with average degree $d$. Prove that there exists an independent set of size $\geq n/(2d)$. An independent set is a collection of vertices $S \subseteq V$ such that no two are adjacent in $E$.

Hint: Sample a random subset $S \subseteq V$ with each vertex included in $S$ with i.i.d. probability $p = 1/d$. Let $G'$ be the graph $G$ restricted to $S$. Then prune $S$ by removing one vertex per edge in $G'$ to find an independent set.

**Problem 3** (The shipping problem)**.** A shipping company has $n$ boats and services $n$ remote ports. The ports are so small that one boat can visit a port per day. On day 0, the boats leave the main terminal under a schedule $T$ for how the boats will travel for the next $m$ days for $m \gg n$. The schedule $T$ satisfies that

(a) each boat visits at most one port per day,

(b) each port is visited by at most one boat per day,

(c) each boat visits each port exactly once,

(d) and a boat is "at sea" if it is not visiting a port that day.

Equivalently, the schedule specifies a matrix $\{T_{ip}\}$ which is the date that boat $i$ visits port $p$.

Some time after the boats leave the main terminal but before any boat arrives at its first port, the government declares a pandemic and requires that the company halt their activities by truncating each boat's schedule to halt at some port. Meaning if boat $i$ is designated to halt at port $p$, the boat will continue its schedule until it docks at port $p$ upon which it will stay at $p$ indefinitely. Furthermore, no other boat is allowed to dock at port $p$ after this time due to the size constraints of the port.

The company is interested in choosing halting ports in a way that maximizes the number of boats that arrive at each port subject to the condition that no two boats are at the same port on the same day.

1. **[15 points]** Construct an algorithm from input matrix $\{T_{ip}\}$ for deciding which port each boat should truncate its schedule. Furthermore, argue that the algorithm is **optimal** meaning that it maximizes the number of boats that arrive at each port. What is the runtime of your algorithm?

   If you choose to use an algorithm we have studied in case as a subroutine, you must argue how the resulting solution is both **feasible** (meaning the output satisfies the condition that no two boats are at the same port on the same day) and **optimal**.

2. **[2 points]** Now argue if $m = n$, there is an even simpler algorithm for this problem. Prove that it is both feasible and optimal.

**Problem 4** (Stable matching implementation)**.** For any stable matching problem defined by matrices $P$ and $R$ and a matching $M$, define the average-proposer-rank as

$$f(n) = \mathop{\mathbf{E}}_{\text{proposer } p} P(p, M(p)). \tag{1}$$

and the average-receiver-rank as

$$g(n) = \mathop{\mathbf{E}}_{\text{receiver } r} R(r, M(r)). \tag{2}$$

**[10 points]** In any programming language of your choosing, implement a program that estimates $f(n)$ and $g(n)$ over uniformly random preference matrices $P$ and $R$ for the matching $M = \text{GaleShapley}(P, R)$.

Produce **separate** scatter plots up to $n = 1000$ of $f(n)$,$g(n)$, $f(n)/n$ and $g(n)/n$. For this problem, you only need to produce the scatter plots and provide the code you use to generate the plots. What asymptotic functions do $f$ and $g$ resemble? (You do not need to give justification.)