CS 341: Algorithms @ Waterloo **Last Updated:** Mar 28, 2023
Homework 5 **Due Date: Apr 8, 2023, 11:00pm**

You are allowed to discuss with others but are not allowed to use any references other than the course notes and the three reference books. Please list your collaborators for each question. You must write your own solutions. See the course outline for the homework policy.

The full mark is 50. This homework is counted 8% of the course grade.

---

1. **Written Problem: Flu Detection Machines** (10 marks)

   There are a lot of flights flying back and forth between country A and country B. Since it is a high season of flu, both governments would like to cooperate to protect their countries from wide range of infections. The governments plan to install powerful flu detection machines in some airports, such that all passengers are scanned by the machine at least once either at the departure airport or the arrival airport. Since that machine is quite expensive, the governments would like to minimize the number of machines they must install in the airports. Given all the flight information between the two countries, our task is to decide which airports to install the machines.

   Input: The airports in country A are called $\{a_1, a_2, \ldots, a_M\}$ and the airports in country B are called $\{b_1, b_2, \ldots, b_N\}$. There are F flights between the two countries, where each flight is denoted by (a_i, b_j) for some $1 \leq i \leq M$ and $1 \leq j \leq N$.

   Output: The minimum number of airports to install the machines so that for each flight there is a flu machine either at the departure airport of the arrival airport (or both).

   Model this as a graph problem and design an efficient algorithm to solve the problem. You will get full marks if the time complexity is polynomial and the proofs are correct.

2. **Written Problem: Facility Scheduling** (10 marks)

   There are m facilities and n people. Each person requests to use a subset of facilities. Each facility can only be used by one person in one day, and each person may only use at most one facility on any given day. Suppose each person requests to use at most d facilities, and each facility is requested by at most d people. Prove that there is a schedule (which person to use which facility on which day) that uses at most d days so that every person can use all the facilities that they requested. Model this as a graph problem and design an efficient algorithm to output such a schedule. You will get full marks if the time complexity is polynomial and the proofs are correct.

   Table 1: Input: Facility Requests

	Adam	Mary	John	Peter
Soccer	✓		✓	
Basketball	✓	✓		
Tennis			✓	✓

   Table 2: Output: Final Schedule

	Day 1	Day 2
Soccer	Adam	John
Basketball	Mary	Adam
Tennis	John	Peter

3. **Written Problem: Shortest Simple Path** (10 marks)

   Prove that the following problem is NP-complete.

Input: A directed graph $G = (V, E)$ where each edge $e \in E$ has a length l_e, an integer L, and two vertices $s, t \in V$. Note that both l_e and L could be negative numbers.

Output: Does there exist a *simple* path from s to t with total length at most L? A path is simple if it visits every vertex at most once.

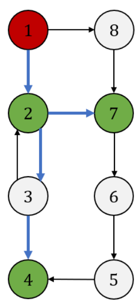
4. **Written Problem: Directed Subgraph** (10 marks)

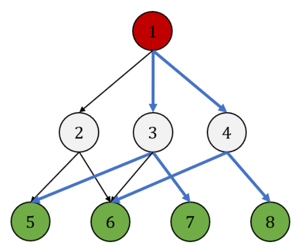
   Prove that the following problem is NP-complete.

   Input: A directed graph $G = (V, E)$, a source vertex $s \in V$, sink vertices $t_1, \ldots, t_l \in V$, and a positive integer k.

   Output: Does there exist a subgraph H of G with at most k edges such that there is a directed path from the source vertex s to each sink vertex t_i for $1 \le i \le l$ in H? Recall that H is a subgraph of G if the edge set of H is a subset of the edge set of G.

   Here are two examples:





   Explanation: The red vertex is the source and the green vertices are the sinks. If the graph on the left is the input graph and $k = 4$, then the answer is YES, by choosing the four highlighted edges. Check that there is a directed path from the source to each sink using only the chosen edges.

   If the graph on the right is the input graph and $k = 5$, however, then the answer is NO. The subgraph H needs to contain at least six edges, and one optimal solution is as shown.

5. **Written Problem: Job Scheduling** (10 marks)

   We are given n jobs, each with a release time $r_i \in \mathbb{Z}_+$, a deadline $d_i \in \mathbb{Z}_+$ and a processing time $p_i \in \mathbb{Z}_+$. Our task is to output a feasible schedule to finish the jobs so that all the deadlines are met, or determine that such a schedule does not exist.

   A schedule is a set of disjoint time intervals $[s_1, f_1], [s_2, f_2], \ldots, [s_n, f_n]$ (with $s_1 < f_1 \le s_2 < f_2 \le \ldots \le s_n < f_n$), and an assignment of each job i to one time interval $t_i \in \{1, \ldots, n\}$ such that we process job i during interval $[s_{t_i}, f_{t_i}]$. A schedule is feasible if each job is assigned to a distinct time interval and the following conditions hold for every job $1 \le i \le n$:

   (a) $r_i \le s_{t_i}$, i.e. we can only start job i after it is released in time r_i;

   (b) $f_{t_i} \le d_i$, i.e. we have to finish job i before its deadline in time d_i;

   (c) $f_{t_i} - s_{t_i} \ge p_i$, i.e. the length of the interval assigned to job i is at least its processing time p_i.

   Either design a polynomial time algorithm in terms of n to solve this problem (and provide a proof of correctness and analyze its time complexity as usual), or prove that this problem is NP-complete.