CSE 6140 / CX 4140 Assignment 5
due Nov 27, 2018 at 11:59pm on Canvas

1. Please upload a *single* PDF named `assignment.pdf` for all of your answers/report containing:
   (a) a typed preamble that contains (-2.5 pts if not included):
      i. the list of people you worked with people for each question (if applicable),
      ii. the sources you used,
      iii. and if you wish, your impressions about the assignment (what was fun, what was difficult, why...);
   (b) your solutions for all problems, typed (not handwritten).
2. **Not abiding by the submission instructions will cause you to lose up to all of the points for a problem.**
3. If you do not understand the question, please ask on Piazza or come to office hours. Misunderstanding the question is not a valid excuse for losing points.
4. Remember the academic honor code: you are not allowed to copy from other students, you may not use material from prior classes, and googling questions on the Internet is not allowed.

1 Fiber Connectivity

You have been hired as a consultant for Google. They wish to install fiber optic cables throughout the state of California, where a cable would go from one city to another. Imagine we are provided with a map of the state where we are given the locations of all the cities and all the roads connecting the cities. Google has asked you to determine the most efficient placement of these cables along the roads such that all cities are connected by these cables. However, you must also keep in mind one very important piece of information: due to the nature of the technology, a city cannot have more than k cables connecting through it.

Formally, suppose we have: n cities, $C = \{c_1, c_2, \cdots, c_n\}$; m roads, $R = \{r_1, r_2, \cdots, r_m\}$, each connecting a pair of cities; and a constant $k \in \{2, ..., |C| - 1\}$. If we are given a subset T of roads, $T \subseteq R$, then define $d_T(u)$ as the number of roads that city u is connected to using only roads in T. We want to find a subset T such that $|T| \leq |C| - 1$ such that $1 \leq d_T(u) \leq k$ for all nodes $u \in C$, and all cities in C are connected together.

Prove that this problem is NP-Complete.

2 Minimum Set Cover

The Minimum SET COVER problem (or SET COVER for short) is defined as follows: Given a set of elements $\mathcal{U} = \{u_1, u_2, \ldots, u_m\}$ and a set of subsets of $\mathcal{U}$, $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}, S_i \subseteq \mathcal{U}, \forall i = 1, \ldots, n$, find a subset $\mathcal{S}'$ of $\mathcal{S}$ such that the union of the selected subsets covers all elements, $\bigcup_{S_i \in \mathcal{S}'} S_i = \mathcal{U}$ and the number of selected subsets, $|\mathcal{S}'|$, is minimized.

1. Devise a branch-and-bound algorithm for the Minimum SET COVER problem. This entails deciding:
   (a) What is a subproblem?
   (b) How do you choose a subproblem to expand?
   (c) How do you expand a subproblem?
   (d) What is an appropriate lower bound?

   Do you think that your choices above will work well on typical instances of the problem? Why?

2. Outline a simple greedy heuristic for the SET COVER problem, and explain why it finds a valid solution and its running time.

3. Imagine you wanted to use a local search method to solve Minimum SET COVER such as Simulated Annealing or Iterated Local Search. Imagine a candidate solution is a subset of the sets that might or might not cover all elements in the Universe set U.
   (a) What could be a possible scoring function for such candidate solutions?
   (b) What would be a Neighborhood (or Moves) you would consider using for your local search to move from one candidate solution to other 'nearby' solutions? How many potential neighbors can a candidate solution have under your Neighborhood (using Big-Oh)?
   (c) Why would you consider adding Tabu Memory and what would be remembered in your Tabu Memory?

3 Optimizing Amazon's Operations

Amazon is considering building a set of warehouses in a new market. There are n possible locations at which a warehouse can be built, and the cost of building a warehouse at location $i \in \{1, \ldots, n\}$ is $f_i \in \mathbb{R}_{>0}$. To serve this new market efficiently, Amazon would like the warehouses to be close to the cities. In particular, there are m cities, and the distance between city $j \in \{1, \ldots, m\}$ and warehouse $i \in \{1, \ldots, n\}$ is $d(i,j) \in \mathbb{R}_{\geq 0}$. Of course, Amazon could build a warehouse at every one of the n possible locations, but that may be too expensive. Instead, the company would like to construct warehouses at a *subset* W of the n locations, $W \subseteq \{1, \ldots, n\}$, such that the sum of the total construction cost and the minimum distances to the cities is minimized. More formally, Amazon would like to find the set W that minimizes:

$$\sum_{i \in W} f_i + \sum_{j \in \{1, \ldots, m\}} \min_{i \in W} d(i,j)$$

1. Devise a branch-and-bound algorithm for this problem. This entails deciding:
   (a) What is a subproblem?
   (b) How do you choose a subproblem to expand?
   (c) How do you expand a subproblem?
   (d) What is an appropriate lower bound?

2. Outline a simple greedy heuristic for the problem, and explain why it finds a valid solution and its running time.

3. Imagine you wanted to use a local search method such as Simulated Annealing or Iterated Local Search to solve Amazon's problem. Imagine a candidate solution is a subset of the locations.

   (a) What could be a possible scoring function for such candidate solutions?

   (b) What would be a Neighborhood (or Moves) you would consider using for your local search to move from one candidate solution to other 'nearby' solutions? How many potential neighbors can a candidate solution have under your Neighborhood (using Big-Oh)?

   (c) Why would you consider adding Tabu Memory and what would be remembered in your Tabu Memory?