CS 330, Fall 2020, Homework 10
Due Wednesday, November 18, 2020, 11:59 pm Eastern Time

Student: Justin DiEmmanuele
Collaborators: Shilpen Patel, George Padavick, Matthew Gilgo

1. **Cell tower problem**

   The company operating the cell towers is at it again. This time they need to install towers along a long country road connecting two cities. Help them find the best location for the towers! Each installed tower covers an area of 5 miles' radius (i.e. 5 miles on either side of the tower). There needs to be cell coverage along the entire road. There is infrastructure to install a tower at every milestone along the road. The total distance between the cities is n miles. (You are allowed to place a tower in either city.) The cost of installing towers varies by location. The costs of the different locations are given to you in an array `cost[ ]` of the length $n + 1$. (The costs are positive numbers.) You need to find a set of building sites such that the entire road has cell coverage and the total cost is minimized.

   (a) Observe that it is possible that the minimum cost solution consists of stretches of road that are covered by multiple towers. Prove however, that in any *optimal* solution, each stretch is covered by at most 2 towers.

   - Suppose we have a stretch of road that is covered by three towers. One possibility, is that two of the towers are in the same position. In this case it is clear that one of the towers is redundant because the removal of one does not change the stretch of road covered.
   - The second, less clear case, is where all three of the towers are in different locations. Again, at least one point along the road is covered by three towers. If this were the case, one tower would be the left most tower and one would be the right most. The middle tower will always be redundant as it will not stretch as far left as the left most tower and also will not stretch as far right as the right most tower. Since there is one space where all three towers cover, we know that the left and right most towers can span the entire covered interval on their own. This means the middle tower will always be redundant. This extends to any number of towers larger than 2.

(b) Consider a greedy algorithm where we add locations to our solution set one at a time. We consider the stretches of road (the road between two neighboring tower locations) in increasing order of distance from the first city. Each time we encounter a so far uncovered road stretch r_i (that is between miles $i-1$ to i), we will add a tower location j that (1.) covers r_i (2.) has the best cost-coverage ratio. That is, we choose a location j such that the relative cost $\frac{\texttt{cost[j]}}{\texttt{additional miles}}$ of every additional mile covered by this tower is minimized.

Show an example where this algorithm fails.

- Drawn in the following figure is a counterexample to the stated rule. The tower we are focused on here is tower b. Assuming towers a and c are in the lowest-cost positions they could be in, the lowest cost position for b to have complete coverage over the interval is clearly where it is placed with a cost of 3. This would give a value of $\frac{3}{6}$ since the tower has a cost of 3 and covers 6 more miles than what was previously covered. Notice that if we place tower b at location 15 instead, the cost is 4 and 10 more miles are covered. this gives a relative cost of $\frac{4}{10}$ which is lower than that of the present location, yet the present location has a lower cost of 3. The proposed rule would not work in this case.

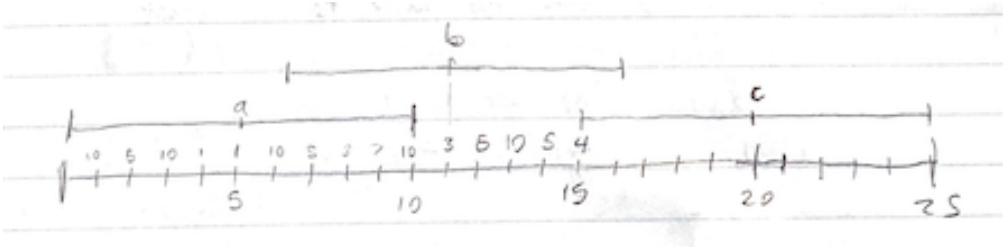

Figure 1: Counterexample

(c) Here is the pseudocode for a memoized algorithm to find the *minimum cost* of locations. (Note that it does not give you the actual set of locations.)

Algorithm 1: DPMinCostTowers(`cost`, n)

```
1 M ← empty array of length n;
2 for i = 0 ... n do
3 |   M[i] ← {some appropriate formula to be defined by you};
4 return M[n];
```

What does the entry `M[i]` contain (that is, state the subproblem for which it provides a solution)?

- $M[i]$ contains the cheapest solution in which there is a tower at position i. This is represented by the following formula:

$$M[i] = cost[i] + min\{M[i-10], \ldots, M[i-1]\}.$$

Write the recursive formula to compute `M[i]` in line 3 of the algorithm. (Write the formula only, no need for explanation.)

Algorithm 2: DPMinCostTowers(`cost`, n)

```
1 M ← empty array of length n;
2 for i = 0 ... n do
3 |   if i ≤ 5 then
4 |   |   M[i] = cost[i];
5 |   else
6 |   |   M[i] ← cost[i] + min([M[max(n + i, 0)] for i in range(−10, 0)]);
7 return M[n];
```

The algorithm essentially implements the formula above but taking into consideration the base cases and possible out of bounds errors when accessing arrays.

(d) Write an algorithm that takes the table M filled in by Algorithm 2 and returns a list of the locations where towers should be built.

2. **Game tournament**

You are participating in an online gaming tournament. The tournament consists of n competitions. You are eligible to participate in any of them. However, you can only be in one competition at a time, there can be no overlap between the tournaments you enter. Sadly, you'll have to make a choice in which competitions you participate. As an input to this problem you will be given an array `start[ ]` of length n containing the start time of each competition, and another array `duration[ ]` with the duration of each game. You may assume that the games are indexed 0 through $n-1$ in increasing order of their start times.

Since you play for the pure enjoyment of the game, you don't care about your placement in any of the competitions. Your goal is to pick games so that you **maximize** the time that you spend playing. (If you start a game, you have to finish it.)

(a) What algorithm from class could be used to solve this problem? How would you transform the inputs to this problem into the ones needed for the algorithm from class? Make sure you understand how that algorithm works before proceeding.

- Weighted interval scheduling in which the weights are the duration of each game.

(b) Due to the large number of entries, the tournament organizers have decided to limit the number of competitions each player can enter to k (so the input now consists of the arrays `start` and `duration`, and the number k). Your goal is to find a set of at most k tournaments that do not overlap, and whose total duration is as long as possible. Design a polynomial-time dynamic programming algorithm for the problem. Please organize your answer as follows:

i. Define the set of subproblems that your algorithm will solve. (For example, for the Knapsack problem, the subproblems were: "for each $i \in \{1, ..., n\}$ and $w \in \{1, ..., W, \}$: compute the maximum value one can obtain by using only items 1 to i and a knapsack with capacity w.")
[*Hint:* Consider having one subproblem for each $i \in \{1, ..., n\}$ and $j \in \{1, ..., k\}$. Think carefully about the order in which to consider the tournaments.]

ii. How many such subproblems are there?

iii. Suppose we create a table M to store the value of the subproblems we solve. Write a recursive formula to compute `M[i,j]` in terms of other entries of M (that is, the solutions to smaller subproblems). (Be mindful to include the base cases.)

iv. Explain briefly why your formula is correct. (This explanation is really the proof of correctness for your algorithm.)

v. Write pseudocode for your algorithm (it should fill the memoization table and find the optimal subset of tournaments).

vi. What is the asymptotic running time of your algorithm? (Only write the Θ formula, no proof needed.)