**CS 325- Homework Assignment 4**

**Problem 1**: *(5 points)* **Class Scheduling**:

Suppose you have a set of classes to schedule among a large number of lecture halls, where any class can class place in any lecture hall. Each class cj has a start time sj and finish time fj. We wish to schedule all classes using as few lecture halls as possible. Verbally describe an efficient greedy algorithm to determine which class should use which lecture hall at any given time. What is the running time of your

algorithm?

We start with an array of classes with start and end times. Also, we start with an empty array that will have a min-heap structure representing lecture halls, prioritized by the end time of the lecture. We will use m to represent the total number of lecture halls needed.

First, we sort the array of classes by start time, using merge sort. Then, we loop through the array of classes. Since the min-heap is empty on the first iteration, we increment m and add the first class’s end time fj to the priority queue. On all remaining iterations, we check if the start time of the current class sj conflicts with (is earlier than) the end time of the earliest ending class, the root node of the min-heap. If there is a conflict, we increment m and add fj to the min-heap. If there is no conflict, we remove the root of the min-heap, and we add fj to the min-heap without incrementing m. At the end of execution, m will be the fewest number of lecture halls needed to accommodate all of the classes.

Time complexity: Merge sort takes nlgn time. Also, it takes n time to cycle through the array of classes. And, in the worst case, it takes lgn time to add an element to the min-heap. We then have nlgn + nlgn = 2nlgn. Based on the rules of asymptotic analysis, the runtime for this algorithm is ϴ(nlgn).

**Problem 2**: *(5 points)* **Road Trip**:

Suppose you are going on a road trip with friends. Unfortunately, your headlights are broken, so you can only drive in the daytime. Therefore, on any given day you can drive no more than d miles. You have a map with n different hotels and the distances from your start point to each hotel x1< x2< ... < xn. Your final destination is the last hotel. Describe an efficient greedy algorithm that determines which hotels you should stay in if you want to minimize the number of days it takes you to get to your destination. What is the running time of your algorithm?

We call a road\_trip function passing in the set of distances x and the max distance d. Within this function, we would call a \*modified\_binary\_search (described below) to locate the largest possible distance in x that is not greater than d. When a distance xi is returned by our modified binary search, that is the distance to the next destination that we travel to. We then take the sum of d and xi and that becomes our new max (d = d + xi). The road\_trip function is then executed again using d and the subset of distances to the left of xi. The recursive calls continue until our modified binary search returns xn. When this happens, we are close enough to our final destination to get there without driving more than d miles from the previous location, and we now know which locations we need to travel to in order to minimize the number of days it takes to get to our final destination.

\*Modified\_binary\_search: Binary search is called with parameters d and the array of distances x. If the middle index (size/2) holds a value that is greater than d, we recursively call the modified binary search on all values to the left of the current index. If the middle index is less than d, we check the index to the right. If the index to the right is greater than d, we have found our destination. If the index is less than d, we continue the search on the left side of the current index. We continue the search until: we arrive at an index that has a value that is less than d with the one to the right being more than d, or we arrive at an index that has a value that is less than d and no value to the right of it, or we reach the end of the binary search without finding a value.

Time-complexity: In a worst-case scenario, we would need to travel to every location x before reaching our final destination. Also, binary search is lgn time. Then, in the worst case, the time to execute this algorithm would be ϴ(nlgn).

**Problem 3:** *(5 points)* **Scheduling jobs with penalties**:

For each 1 ≤ 𝑖 ≤ 𝑛 job 𝑗𝑖 is given by two numbers 𝑑𝑖 and 𝑝𝑖, where 𝑑𝑖 is the deadline and 𝑝𝑖 is the penalty.

The length of each job is equal to 1 minute and once the job starts it cannot be stopped until completed. We want to schedule all jobs, but only one job can run at any given time. If job i does not complete on or before its deadline, we will pay its penalty 𝑝𝑖. Design a greedy algorithm to find a schedule such that all jobs are completed and the sum of all penalties is minimized. What is the running time of your algorithm?

Assume we have an empty set z that will contain the jobs in the order they should be executed. We will use merge sort to sort jobs by penalty pi from largest to smallest. This sorted array will be x. We will loop through x, starting with the first index i, and use the deadline di­ of the job at index i to calculate which indexes we check in z. We will then check for free indexes in z, starting at j = di – 1 and moving toward j = 0. We want to find the highest index in z that is less than our deadline di. If an index in z is free, we assign the job to execute during that minute by adding it to that index of z. If the index is not free, we continue checking for a free index that is less than our deadline, moving towards the 0th index. If there are no free indexes, we add di to a penalty set y, to be executed only after the jobs that will not incur a penalty. We would continue looping through x until all jobs have either been added to z or y.

Time-complexity: The outer loop will iterate n times and the inner loop will iterate n or less times. Since we are dealing with a nested loop, the run-time will be O(n2).

We do not need to consider the run-time of merge sort in this case, since it is a lower order term.

**Problem 4:** *(5 points)* **CLRS 16-1-2** **Activity Selection Last-to-Start**

Suppose that instead of always selecting the first activity to finish, we instead select the last activity to start that is compatible will all previously selected activities. Describe how this approach is a greedy algorithm, and prove that it yields an optimal solution.

References: Followed lecture 3 from this week to help understand proofs for this problem.

We must show that the algorithm uses the greedy choice property and has optimal substructure.

Greedy Choice Property: We must show that there is an optimal solution that begins with a greedy choice (the activity with the latest start time). Suppose we have a set S of activities, and A is a subset of S that is an optimal solution to the Activity Selection problem for S. The activities in A are ordered by start time in descending order (later start times come first). Also, suppose that k is the first activity in set A. If ks  (the start time of k) is the latest start time, then obviously this is the greedy choice.

Now, suppose ks is less than or equal to the start time of another activity xs. Since we know that xs ≥ ks, we could remove k and replace it with x for an alternative optimal solution.

B = A – {k} U {x}

xs ≥ ks so B is also an optimal solution

Optimal substructure: Suppose that set A is an optimal solution for set S. Also, A’ = A – {1} and S’ is all elements in S with lower start times than the first item in A. Then, A’ must be an optimal solution to S.

If there was a solution to S’ that was more optimal than A’, then we could add {1} to that solution and have a solution to S that was more optimal to A. However, this contradicts what we already know, that A is an optimal solution to S. Therefore, this problem has optimal substructure.

1

**CS 325- Homework Assignment 4**

**Problem 5:** *(10 points)* **Activity Selection Last-to-Start** **Implementation**

*Submit a copy of all your files including the txt files and a README file that explains how to compile and run your code in a ZIP file to TEACH. We will only test execution with an input file named act.txt.*

You may use any language you choose to implement the activity selection last-to-start algorithm described in problem 4. Include a verbal description of your algorithm, pseudocode and analysis of the theoretical running time. *You do not need to collected experimental running times.*

The program should read input from a file named “act.txt”. The file contains lists of activity sets with number of activities in the set in the first line followed by lines containing the activity number, start time & finish time.

**Example act.txt:**

11

1. 1 4
2. 3 5
3. 0 6
4. 5 7
5. 3 9
6. 5 9
7. 6 10
8. 8 11
9. 8 12
10. 2 14
11. 12 16

3

3 1 8

2 1 2

1 3 9

In the above example the first activity set contains 11 activities with activity 1 starting at time 1 and finishing at time 4, activity 2 starting at time 3 and finishing at time 5, etc.. The second activity set contains 3 activities with activity 3 starting at time 1 and finishing at time 8 etc. **Note**: the activities in the file are not in any sorted order.

Your results including the number of activities selected and their order should be outputted to the terminal. For the above example the results are:

Set 1

Number of activities selected = 4

Activities: 2 4 8 11

Set 2

Number of activities selected = 2

Activities: 2 1

2