HOMEWORK 5 - GREEDY ALGORITHMS

Problem 1

Decide if the following statement is true or false. If true, give a short explanation. If false, give a counterexample.

> Let $G = (V, E)$ be an arbitrary connected, undirected graph with a distinct cost $c(e)$ on every edge. Let $|E| = m$, $|V| = n$, and $m = n+1$. Let $e*$ be the second heaviest edge in G. Then the minimum spanning tree T of G contains edge $e*$.

Problem 2

Decide if the following statement is true or false. If true, prove it. If false, give a counterexample.

> A graph with distinct edge weights has a single minimum spanning tree.

Problem 3

Background. Netflix is the world's largest consumer of bandwidth, taking up 15% of all Internet bandwidth globally. Given that its users reside in 190 different countries, it must take into consideration the large discrepancies in bandwidth worldwide and how this affects download speeds and limits on data usage. Viewers in a location with very low bandwidth may have to wait hours for a movie to download, whereas the same movie might be ready in seconds for viewers in an area with high bandwidth. We will not discuss Netflix's architecture in detail, but we will be working with a hugely simplified representation of its content delivery system.

Premise. Assume Netflix has a central server from which video files are transmitted. There is a network of routers that deliver the files from the server to users, and video files are downloaded in smaller chunks over time. Imagine there are 2 Netflix users in different locations trying to watch the same movie. The paths taken by the movie file from the Netflix server to the users can be represented by the graph in Figure 1, where:

- the graph is a directed acyclic graph with source N and sinks A and B (user A and user B)
- each node represents a router and keeps track of any file chunks it currently holds

- each edge is decorated with 2 different weights: "speed" and "slots"
  - "speed" represents how fast a file can be delivered along an edge, measured in megabits per second (Mbps) (e.g. If an edge has speed = 5, then 5 Mb can be transmitted across that edge at a time.)
  - "slots" represents the number of file chunks that an edge can simultaneously accommodate (e.g. If slots = 3 and speed = 5 for some edge, then 3 chunks of size 5 Mb can be delivered along that edge at one time.)
- transmission of data across one edge takes 1 second.

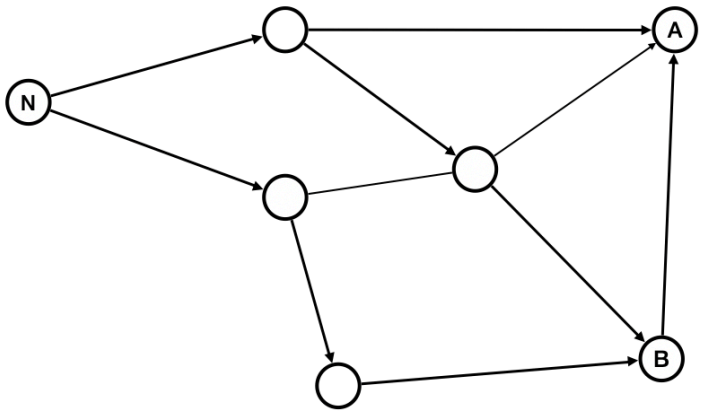

FIGURE 1. The example graph.

Suppose we want to deliver a 4 GB (= 4,000 MB = 32,000 Mb) movie to users A and B at the same time. 2 copies of this video file start at node N and are transmitted in chunks across the graph until each user has received the full file. If an edge has slots > 1, then the edge may transmit *slots* number of chunks from either or both files all at that edge's speed, (i.e. speed = 5 with slots = 3 means 3 chunks of 5 Mb may be transmitted along that edge at once). When one chunk is delivered from one node to the next, each subsequent chunk will follow its path. As a result, for each file, the first chunk will find the best path to the user, and the rest of the chunks will retrace that path until all chunks totaling 4 GB have reached the user.

Tasks. It may happen that the 2 best paths that you have chosen require both files to use the same edge at the same time. We refer to this a "collision." You must decide the best way to handle this situation. Given the above background and premise:

write a "collision policy" detailing how you would handle collisions.

Keep in mind that there could be any number of slots so you must decide how you will use them. For example, you could just randomly choose one user and allocate all slots to that user every time there is a collision. You are not permitted to use this solution. Briefly justify your solution in about a paragraph. There is no universal answer to this question – your answer will be graded on thoughtfulness. Careless, sloppy answers that do not engage with the problem will not receive credit.

Reflection question. Please answer the following reflection question in addition to submitting a paragraph on your "collision policy."

> *What if Netflix designed their service to always prioritize the streams of the users with the largest bandwidth? Would this be fair? Would users with high bandwidth perceive this as fair? Would users with low bandwidth perceive this as fair? How do you personally decide what is fair in this problem?*