**CSC373 Assignment 2**

Note: all pseudo-code will be written using Python-style semantics

**Terminology and Notation:**

|  |  |
| --- | --- |
| **Term / Notation** | **Meaning** |
|  | Directed edge that goes from node to node |
|  | Undirected edge that connects node and node |
|  | The shortest path from node to node |
|  | A path from node to t |
|  | If statement has multiple correct possibilities, this is the set of all correct possibilities that satisfy (e.g. is the set of all possible paths from node to node ) |
| Constrictor | A constricting edge |
| Augmentation | Augmenting path |
| CC | Connected component |
|  | Set of CCs in graph |
|  | A spanning tree of graph |
|  | Set of irrational numbers () |
| FF | The Ford-Fulkerson method |
|  | Denotes that an method is being run times for a total runtime of |
| Parent Node | Node is a parent of if |

1. True / False Short Answer

|  |  |  |
| --- | --- | --- |
| **Claim** | **Verdict** | **Justification / Counter Example** |
| An undirected graph with vertices and at most edges has at least connected components. | True | Consider a graph with nodes and zero edges – the number of CCs is (the number of nodes).  If we add an edge to connect two CCs into one CC, we reduce the number of CCs by 1 (from to , where ).  Thus, if we add edges to graph where  , we reduce from  to . |
| The shortest-paths tree computed by Dijkstra is necessarily an MST. | False | We note that the total weight of the Dijkstra tree is 6, while the total weight of the MST is 4. Clearly the Dijkstra tree is not minimal. |
| Suppose that we have computed an MST. If the weight of each edge in the graph is increased by 1, the computed spanning tree remains minimum with respect to the new weights. | True | Let the graph in question be , the MST be and the set of spanning trees be .  We know that the number of edges in a tree with nodes is , so if we increase the weight of every edge in a tree by , the total weight increases by .  Thus, the relative total weights of the trees in remain the same under the new weights, therefore is still minimal in , with respect to the weights. |

|  |  |  |
| --- | --- | --- |
| **Claim** | **Verdict** | **Justification / Counter Example** |
| In a weighted directed graph with positive weights, Dijkstra might call the procedure (aka procedure, see CLRS 24.3) on the same edge more than once. | False | We inspect the pseudo-code given in CLRS 24.3 – clearly, Dijkstra only considers the leaving edges of every vertex only once for relaxation (lines 7-8).  Since edges can only leave one vertex, Dijkstra’s algorithm must only consider one edge for relaxation. |
| Maximum flow in a network with integral capacities is necessarily integral. | False |  |
| We are given a weighted graph and . If all edge weights in the graph are multiplied by a positive constant, remains shortest from s to t with respect to the new weights. | True | Let the graph in question be , and  , where .  If we multiply all the weights of by some , the total weights of the paths in are also multiplied by .  Since was minimal in with respect to total weight (by definition), must also be minimal (the shortest path) in . |
| Suppose we run DFS on a directed graph and we find a vertex with discovery time 1 and finishing time . Then the entire graph must be strongly connected. | True | We note that the (relative) finishing time of the head of the DFS tree is equal to double the number of nodes connected to the head of the tree (DFS spends 2 “turns” on each node, one going “forward” and one “backward”)  Since , we can say that the number of nodes connected to the head of the tree is .  If nodes are connected to the head, we can say there is a way of getting from any node to another. |

|  |  |  |
| --- | --- | --- |
| **Claim** | **Verdict** | **Justification / Counter Example** |
| Ford-Fulkerson method runs in polynomial time assuming that capacities are positive integers. | False | We note that FF runs in time, where is the maximum flow of the target network (CLRS 26.2). This is clearly not polynomial with respect to or .  Example: the following takes 1000 iterations to complete with traditional FF |
| Ford-Fulkerson method terminates on all input flow networks with capacities that are positive real numbers. | False | Consider graph and capacity function  , where (there is an edge with positive irrational weight).  By CLRS 26.2, footnote 2, FF can fail to terminate when given irrational edge capacities.  Thus we have found a network with capacities in (since ) that causes FF to run indefinitely. |
| Undirected graph is called -regular if degree of every vertex is exactly . Girth of the graph is the length of a shortest cycle in G. For example, the simple cycle with 5 vertices is a 2-regular graph of girth 5. We claim that there is no 3-regular graph of girth 5 on 10 vertices. | False | This graph is 3-regular (each node is connected to 3 other nodes), has 10 nodes and has girth 5. |

1. Number of Distinct Shortest Paths
   1. English Description of Algorithm
      1. We will use a modified breadth first search, where we will also maintain (as a node attributes ) the number of ways to get though the node, and the distance from the starting node (number of edges)
      2. We initialize this modified BFS at , setting , and for all other nodes
      3. Every time we encounter a node that is currently on the priority queue, we increment by , where is the node that has “discovered” (i.e. a node with an edge leading to )
      4. When we finish the BFS, we return as the answer
   2. Pseudocode

def algorithm(graph, s, t):

# Set num\_ways and dist\_from\_start of all nodes except s to 0

for node in graph.nodes:

node.num\_ways = 0

s.num\_ways = 1

# Maintain bit-vectors that keep track of nodes:

future = [1] \* len(graph.nodes) # To be explored

present = [0] \* len(graph.nodes) # Currently getting explored

past = [0] \* len(graph.nodes) # Explored already

# Helpers for setting the bitvectors appropriately

def make\_future(node):

future[node], present[node], past[node] = 1, 0, 0

def make\_present(node):

present[node], future [node], past[node] = 1, 0, 0

def make\_past(node):

past[node], present[node], future[node] = 1, 0, 0

# Set s’s values in the bitvector appropriately (we’re starting from here)

make\_present(s)

# Maintain a priority queue of nodes for exploration

queue = PriorityQueueHeap()

queue.push(s) # s is the first node to explore from

while len(queue) > 0:

cur\_node = queue.pop()

adjacent\_nodes = graph.adjacent\_to(cur\_node)

for nxt\_node in adjacent\_nodes:

if present[nxt\_node]:

nxt\_node.num\_ways += cur\_node.num\_ways

elif future[nxt\_node]:

nxt\_node.num\_ways = cur\_node.num\_ways

queue.push(nxt\_node)

make\_present(nxt\_node)

make\_past(nxt\_node)

return t.num\_ways

* 1. Proof of Correctness
     1. We note that the number of ways of going through a node is equal to the sum of the number of ways of going through its parent nodes
     2. Our algorithm modifies BFS such that:
        1. When we see a completely novel node, we set its value to the of its “discoverer” node – since we know that:

(as we can just concatenate onto to get )

* + - 1. When we see a node that has been explored by another node, we just increment its value by the value of the current (parent) node. This is because:

Thus:

* + - 1. Since we are correctly modifying the attribute, the algorithm must be returning the correct result when we query for it’s value
  1. Runtime Justification
     1. This algorithm should have the same runtime as BFS: (CLRS 22.2), as the code is very similar, just with the added activity of manipulating a node attribute, which is done in constant time

1. Server Movement Challenge

*Let be the length of the given cable (I can’t figure out how to do curly L’s in Word)*

* 1. English Description of Algorithm
     1. Construct a *complete* (every node has an edge to every other node), weighted graph of the following form:
        1. Nodes: electrical outlets
        2. Weights: distance from one electrical outlet to the other
     2. Find the minimum spanning tree of , call it
     3. Find the **shortest path in**  from the starting outlet to the target outlet. Call this
     4. Find the longest distance between outlets in , call this .
     5. Return , which is the length of cable needed to move the server without outages
  2. Pseudocode

def dist(pt1, pt2):  
 x\_diff = pt1.x – pt2.x

y\_diff = pt1.y – pt2.y

return ((x\_diff \*\* 2) + (y\_diff \*\* 2)) \*\* 0.5

def build\_graph(outlets):

# Make a graph where the nodes are the outlets

graph = Graph()

for outlet in outlets:  
 graph.add\_node(outlet)

# Connect the graph completely (each edge weight is the distance between nodes)

for n1 in graph.nodes:  
 for n2 in graph.nodes:  
 if n1 != n2 and graph.not\_connected(n1, n2): # no self/duplicate edges!

graph.connect(n1, n2, weight=dist(n1, n2))

return graph

def dijkstra(graph, source, target):

return (the shortest path from source to target in list

form:[source, …, target])

def mst\_as\_graph(graph):

# Use Prim to find the MST of graph

mst = Prim(graph)

# Return the MST of graph in graph form

return Graph(nodes = mst.nodes,

edges = mst.edges)

def find\_largest\_gap(points):

# Finds the largest gap between two consecutive points in points

# Setting a high score that can be beaten easily

high\_score = -1 \* infinity

# Look at two nodes at a time (sliding window size 2)

for [p1, p2] in path\_nodes:

# Set a new high-score if it is beaten

if dist(p1, p2) > high\_score:

high\_score = dist(p1, p2)

# Return the distance between the two furthest consecutive pts in points

return high\_score

def algorithm(outlets, given\_cable\_length, src\_outlet, dst\_outlet):

graph = build\_graph(outlets) # Rep as graph

mst\_graph = mst\_as\_graph(graph) # Find MST

short\_path = dijkstra(mst\_graph, src\_outlet, dst\_outlet) # Find min path in MST

largest\_gap = find\_largest\_gap(short\_path) # Find longest edge

return largest\_gap – given\_cable\_length # Return what we want

* 1. Proof of Correctness
     1. Proof of Helper Function Correctness
        + 1. This is just a simple calculation and it is trivially correct
          2. This algorithm is supposed to convert the list of outlet locations into a connected graph of the following form:

Nodes: locations (Cartesian coordinates)

Weights: distances between locations

* + - * 1. We can clearly see that the semantics of the first part clearly add all the locations as nodes in the graph, and the semantics of the second part clearly connect distinct pairs of nodes together with edges with weights equal to the distance between the Cartesian coordinates of the nodes
        2. This just returns the results of Dijkstra’s algorithm in a list format – CLRS 24.3 proves Dijkstra’s algorithm to be correct
        3. This just returns the result of Prim’s algorithm in graph format – CLRS 23.2 proves Prim’s algorithm to be correct
        4. This is just a simple 2-window, linear search for the pair of consecutive nodes with the largest distance between them – proving the correctness of a linear search is trivial
    1. Proof of Main Algorithm:
       1. Definitions / Declarations
          1. Let be the edges, vertices/nodes and cost function of graph , respectively
          2. Let be the minimum spanning tree of graph
          3. Let the *mass* of a path be the **maximum weight** of an edge in

, i.e.

* + - * 1. Let be the path of **minimum mass** between nodes in graph .
        2. Let be the **shortest path** between nodes in – note that since is a tree, must be unique.
        3. Lemma A: . Proof:

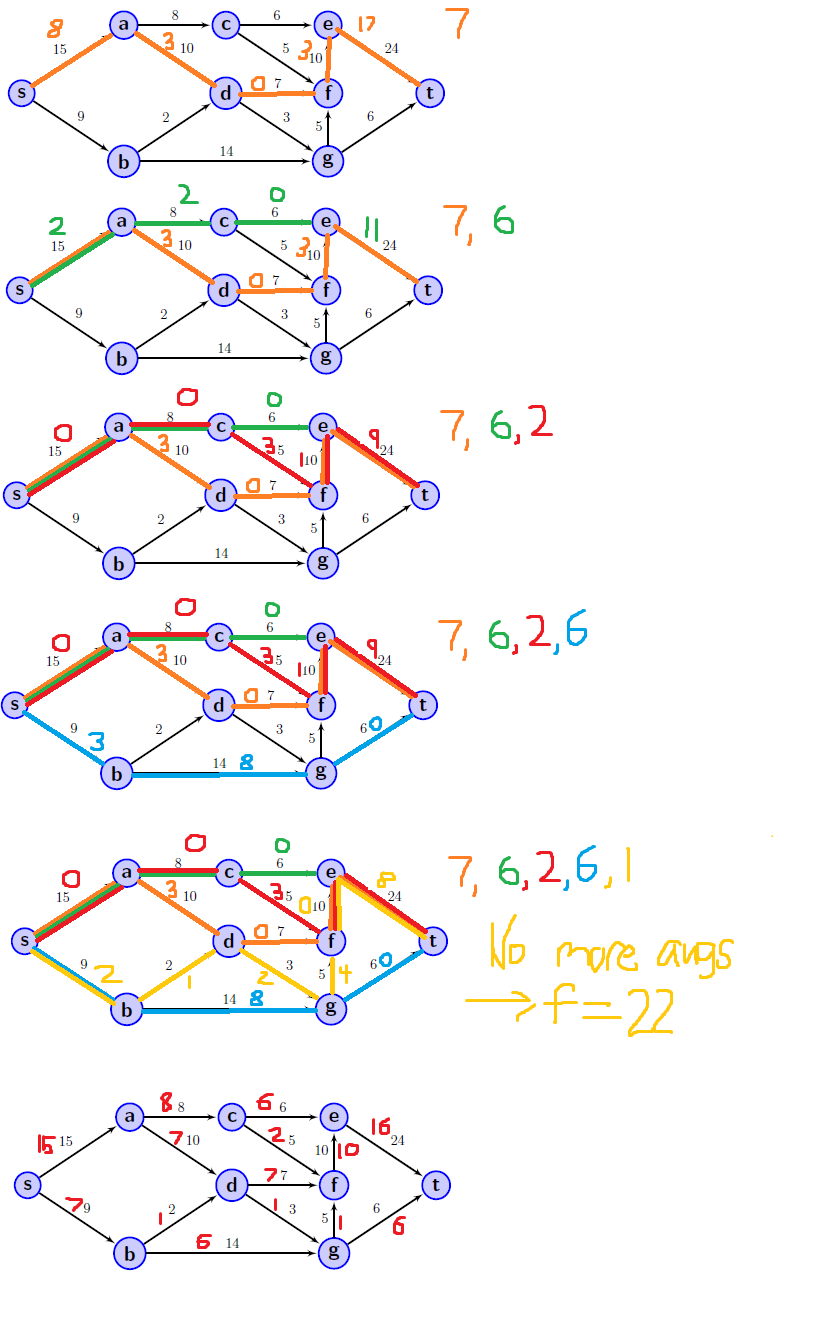
Contained in page 3 of the following document – for maximum spanning trees, but logic can be adapted for minimum spanning trees: http://suraj.lums.edu.pk/~te/cspf/cspf34.pdf

* + - 1. Proof of Correctness:
         1. The goal of our algorithm is to find the minimum length of cable needed to get the cart from starting outlet to destination outlet .
         2. In order to find this length of cable, we first need to find a path where the largest gap between nodes is minimized (because this will minimize the amount of cable that we need to “leapfrog” the cart)
         3. By A, we know that the shortest path between nodes in an MST also happens to be the path with the lowest maximum weight – this is what we need to satisfy the above requirement
         4. Since our algorithm correctly finds an MST and then correctly finds a shortest path on that MST, it is also correct in finding the maximum distance between two-nodes on this minimum-mass path.
         5. Since the algorithm is correct in finding the total distance of cable needed, it is also correct in returning the amount of cable needed to compensate for (since )
  1. Runtime Justification
     1. Assume we use the following implementations (runtimes sourced from CLRS):

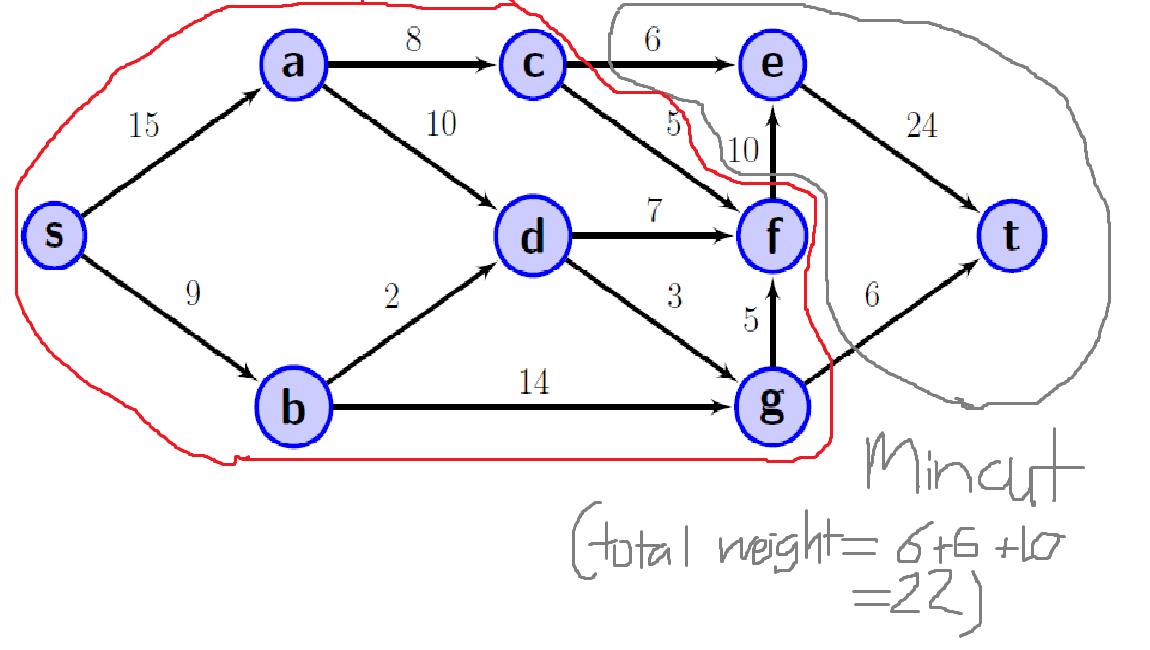
|  |  |  |
| --- | --- | --- |
| **Data Structure / Algorithm** | **Implementation** | **Runtime(s)** |
| Graph | Adjacency List (hash-map of nodes to linked lists of edges) | Adding nodes (hash-map insertion worst case)  (CLRS 11.2)  Connecting nodes (if we maintain a pointer to the end of each linked list) |
| Prim’s Algorithm | Binary Heap (priority queue)  Adjacency List (graph) | (CLRS 23.2) |
| Dijkstra’s Algorithm | Binary Heap (priority queue)  Adjacency List (graph) | (if all nodes are reachable from source, CLRS 24.3) |

* + 1. These are the non-trivial steps to the algorithm:
       1. Building the graph –
          1. First step is to add all the outlets as nodes to the graph (can be done with a simple loop over all outlets, will take time for each insertion in the worst case, so total)
          2. Second step is to connect all the nodes together according to Euclidean distance (since there are many nodes, there are many pairs of different nodes, thus the double-loop takes time (if node connection is time)
          3. The dominant term of the above is , thus this is the upper-bound for this step’s runtime
       2. Finding the MST of the graph -
          1. We note that for the initial graph representation of the outlets,
          2. Thus, finding the MST using Prim’s algorithm will take time
          3. We also note that the MST has nodes and edges (by virtue of being a tree with nodes)
          4. Converting this MST to a graph will take time to add the vertices and time to add the edges
          5. The dominant term of the above is
       3. Finding the shortest path in the MST to get from the start to the end –
          1. We note that the graph we are running Dijkstra’s on has , so the runtime of this step is
       4. Finding the longest gap in the shortest path -
          1. This is just a simple linear search for the longest length between points, with a sliding window of size 2
          2. With a window size of 2, this step will take iterations
    2. We note that the dominant term out of all of the above steps is , and this is the upper-bound for the runtime of our algorithm

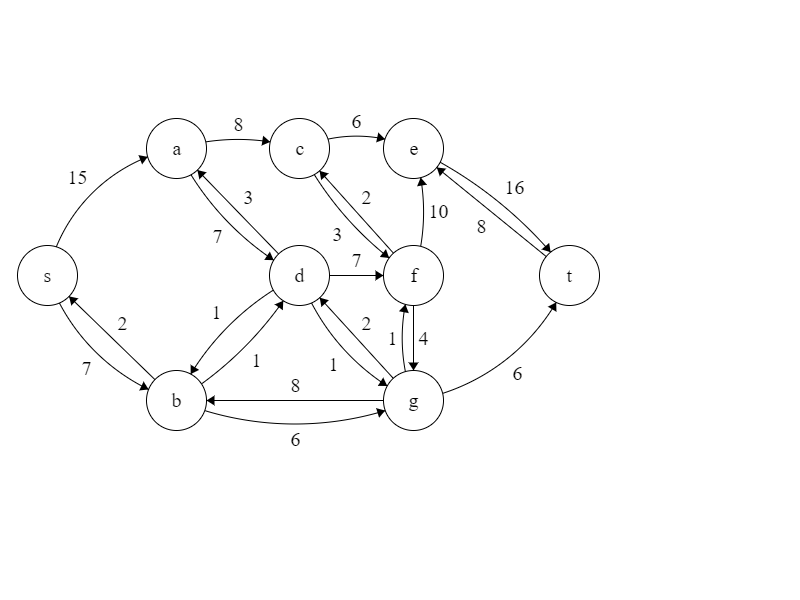
1. Network Flow
   1. Computations
      1. Max Flow: 22 (16 + 6 or 15+7)



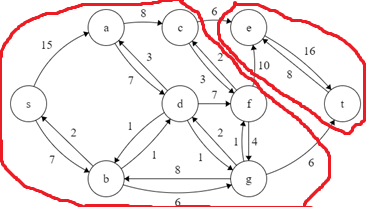
* + 1. Min Cut



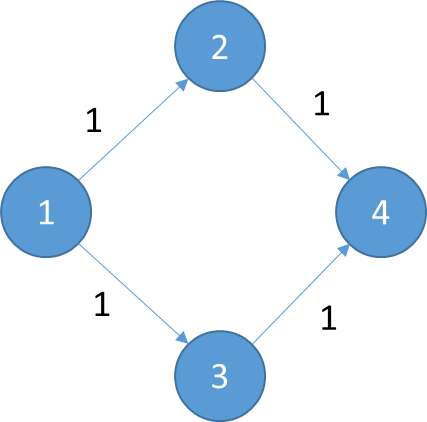
* 1. Residual Graph
     1. Diagram



* + 1. Min Cut



* 1. List of Constrictors
  2. Small Graph with No Constrictors



* 1. Constrictor-Finding Algorithm
     1. English Description:
        1. We are going to find the edges that belong on the min-cut – these are constrictors
        2. Construct the residual graph using Edmonds-Karp, a variant of FF
        3. Find all nodes that can be reached from the source in the residual graph using DFS, mark the rest of the nodes as unreachable
        4. In the original flow network, look at the edges leaving the reachable set of nodes – filter these by the edges that connect reachable nodes to unreachable nodes
        5. Return the set of edges that link the nodes marked as reachable and unreachable together
     2. Proof of Correctness
        1. Note that edge is a constrictor increasing results in the creation of a new
        2. If increasing results in the creation of a new then more flow can be pushed along this path to increase max-flow, which means that must be a constrictor – if this increase in capacity does not result in the creation of a new augmenting path, then is not a constrictor
        3. Thus, the edges that cross between the reachable and unreachable sets of nodes are constrictors – since the max-flow of a network is equal to its minimum cut, increasing the capacity of any of the edges on the minimum cut must also increase the max-flow
     3. Runtime Justification
        1. Step Analysis:
           1. We know that Edmonds-Karp runs in time (CLRS 26.2)
           2. Finding the set of reachable and unreachable nodes can be done using DFS and looping (runtime )
           3. Finding the set of edges that link the reachable and unreachable node sets together takes time (to loop over all the edges)
        2. The dominant term of the above is , which means it is the upper-bound for the running time of this algorithm

1. Algorithm to Update Max-Flow on Capacity Changes
   1. English Description of Algorithm
      1. Look at the residual graph of the flow network and the flow network itself
      2. Create a new graph with the original nodes with the same edge directions as the original flow network
      3. We mark unsaturated edges with weight 0 (as there is spare room to push flow through), and saturated edges with weight (as we need to upgrade these to push flow through), where is the current edge.
      4. We then find the minimum-weight path needed to get from the source to the terminal nodes – if the total weight is 0, we return the empty list – we don’t need to upgrade any edges
      5. Otherwise, this shortest-path will be the collection of edges that we need to upgrade to increase our max-flow
   2. Pseudocode

def make\_upgrade\_cost\_graph(flow\_graph, original\_network):

# original\_network is the graph of the flow network

# flow\_graph is the “implementation” of the max\_flow

# Graph to be returned at the end of the algorithm

ret\_graph = DirectedGraph()

# Add the nodes to the graph

for node in original\_network:  
 ret\_graph.add\_node(node)

# Assume that edges are identified identically between the two inputs,

# and that weights in flow\_graph do not exceed the weights of original\_network

for edge in original\_network.edges:

# If the current edge is saturated, set weight of this edge to p(e)

if flow\_graph.weight(edge) == original\_network.weight(edge):

ret\_graph.connect\_with\_weight(edge.source, edge.end, p(edge))

# Otherwise set the weight of this edge to 0

else:

ret\_graph.connect\_with\_weight(edge.source, edge.end, 0)

# Return the upgrade-cost graph

return ret\_graph

def dijkstra(graph, source, target):

return (the shortest path from source to target in list

form:[source, …, target])

def consecutive\_nodes\_to\_edges(nodes):

ret\_lst = []

for [n1, n2] in nodes:  
 ret\_lst += WeightedDirectedEdge(n1, n2, weight=p(n1, n2))

return ret\_lst

def algorithm(flow\_graph, original\_network):

# Make the graph of upgrade costs

upgrade\_graph = make\_upgrade\_cost\_graph(flow\_graph, original\_network)

# Find the shortest path along the upgrade-cost graph

cheapest\_path\_to\_upgrade\_nodes = dijkstra(upgrade\_graph,

flow\_graph.source,

flow\_graph.terminal)

cheapest\_path\_to\_upgrade =

consecutive\_nodes\_to\_edges(cheapest\_path\_to\_upgrade\_nodes)

# Return the empty list if we don’t have to upgrade anything

if sum(cheapest\_path\_to\_upgrade, key=lambda(x)->x.weight) == 0:

return []

# Otherwise return the shortest path as the collection of edges to upgrade

else:

return cheapest\_path\_to\_upgrade

* 1. Proof of Correctness

**I DO NOT KNOW HOW TO ANSWER THIS QUESTION**

* 1. Runtime Justification

|  |  |  |
| --- | --- | --- |
| **Data Structure / Algorithm** | **Implementation** | **Runtime(s)** |
| Graph | Adjacency List (hash-map of nodes to linked lists of edges) | Adding nodes (hash-map insertion worst case)  (CLRS 11.2)  Connecting nodes (if we maintain a pointer to the end of each linked list) |
| Prim’s Algorithm | Binary Heap (priority queue)  Adjacency List (graph) | (CLRS 23.2) |
| Dijkstra’s Algorithm | Binary Heap (priority queue)  Adjacency List (graph) | (if all nodes are reachable from source, CLRS 24.3) |

* + 1. Steps of the Algorithm:
       1. First step is to create the upgrade cost graph:
          1. (if hashing is done stupidly)

(if hashing is done competently)

* + - * 1. Adding nodes to the graph takes time in the worst case if hashing is done stupidly.
        2. If the hashing is competent, this takes time to do
        3. Adding edges to the graph takes time
      1. Second step is to find the shortest path along the upgrade-cost graph:
      2. Third step is to convert the shortest path from node form into edge from:
         1. This is just a single sliding window, will obviously run in time proportional to the input size
      3. Fourth step is to check whether the shortest path has total weight of 0:
         1. Summing a collection of numbers using a for-loop will take time proportional to the size of the collection
    1. Total Runtime (dominant term analysis):
       1. Dumb hashing:
       2. Normal hashing: