

CE2001 : Algorithms

Project 2

SEP2

Group 6

Done By:

(U1921023K) Mervyn Chiong Jia Rong

(U1921062L) Ang Kai Jun

(U1920248G) Lee Zhe Ren

(U1923071B) Harsh Doddannavar

**Introduction**

Our group debated between using either BFS(Breadth-First Search) or DFS(Depth-First Search). Ultimately deciding to use BFS. This is purely because the problem posed is to locate the shortest path to a certain location(hospital) in an undirected unweighted graph G. DFS is not applicable in this case because it would traverse more edges than it is required to do from destination to endpoint.

**Analysis of Breadth First Search Method**

**Assumptions:**

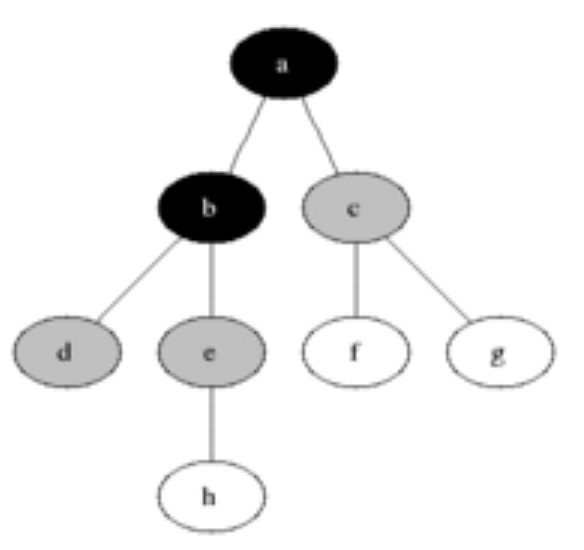
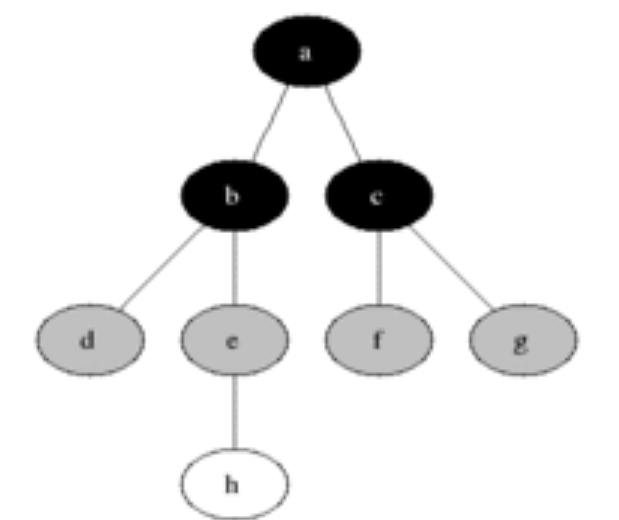
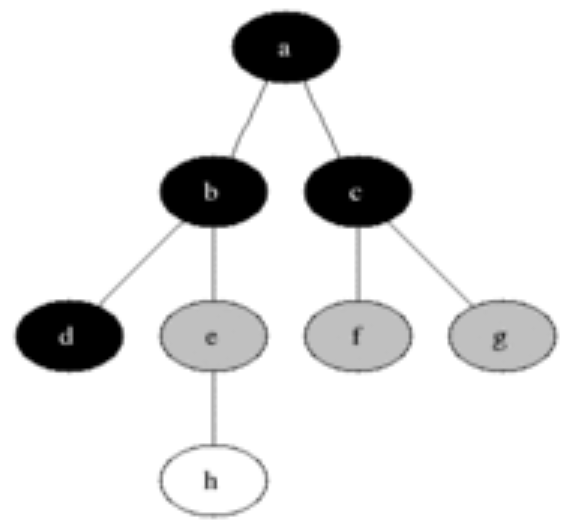
**Graph is G**

**Every edge is taken as a single unit(unweighted graph), the total sum would be E**

**Any other node, S**

**Source(root) node is the vertex,V**

BFS starts off from a root node(something like the top of a large tree diagram) and move to its adjacent nodes, then to all nodes that are connected to the root node and it will then move on to the next set but will go back to do the other nodes from the original source executing the same logic.

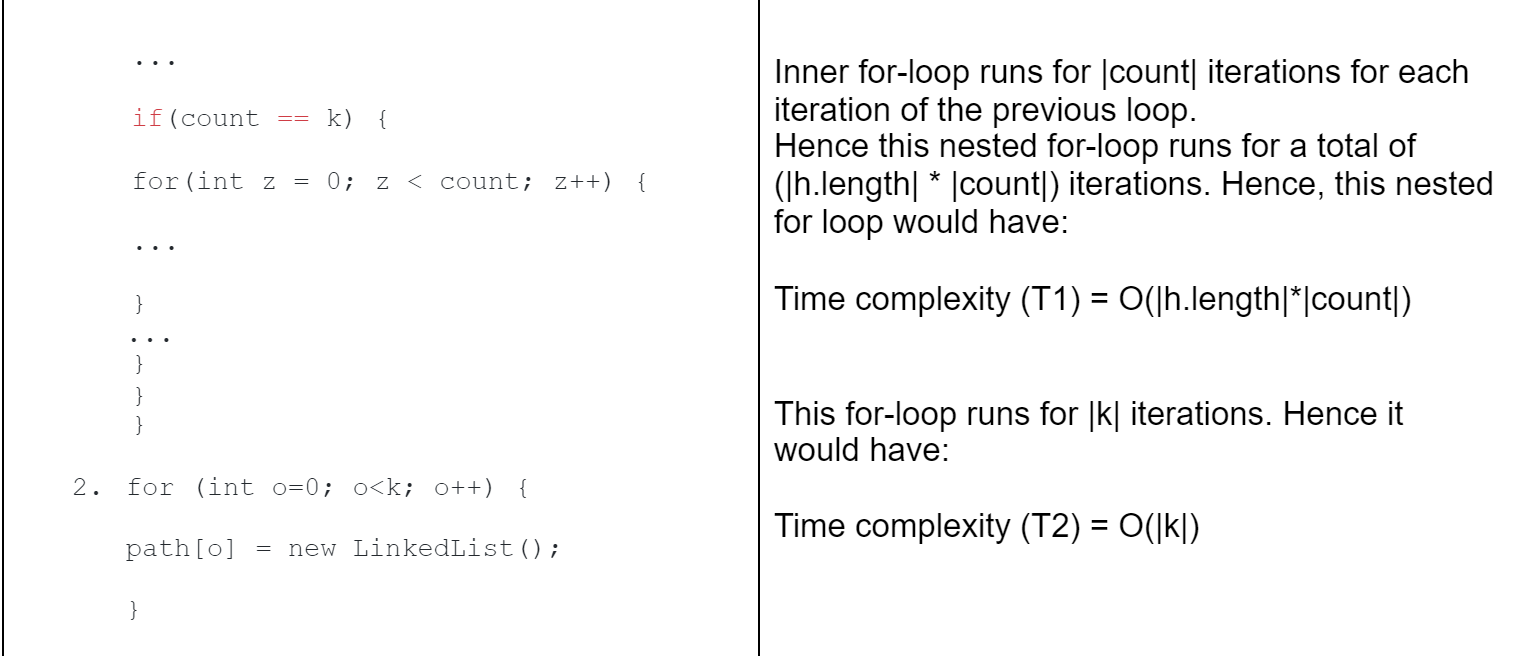
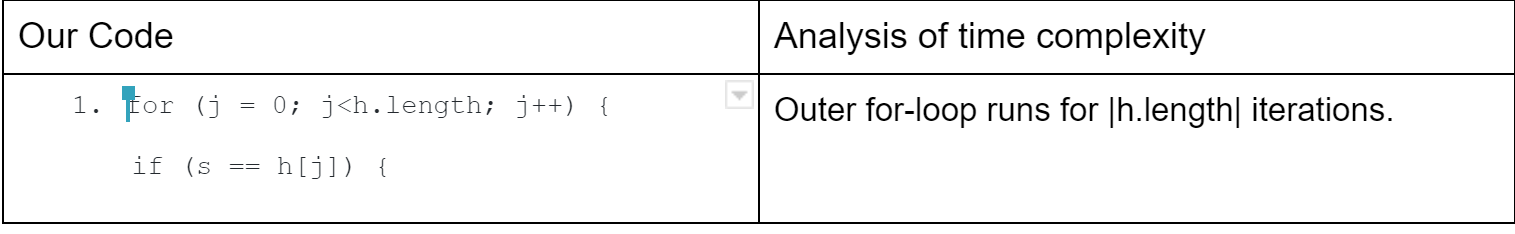
-> ->

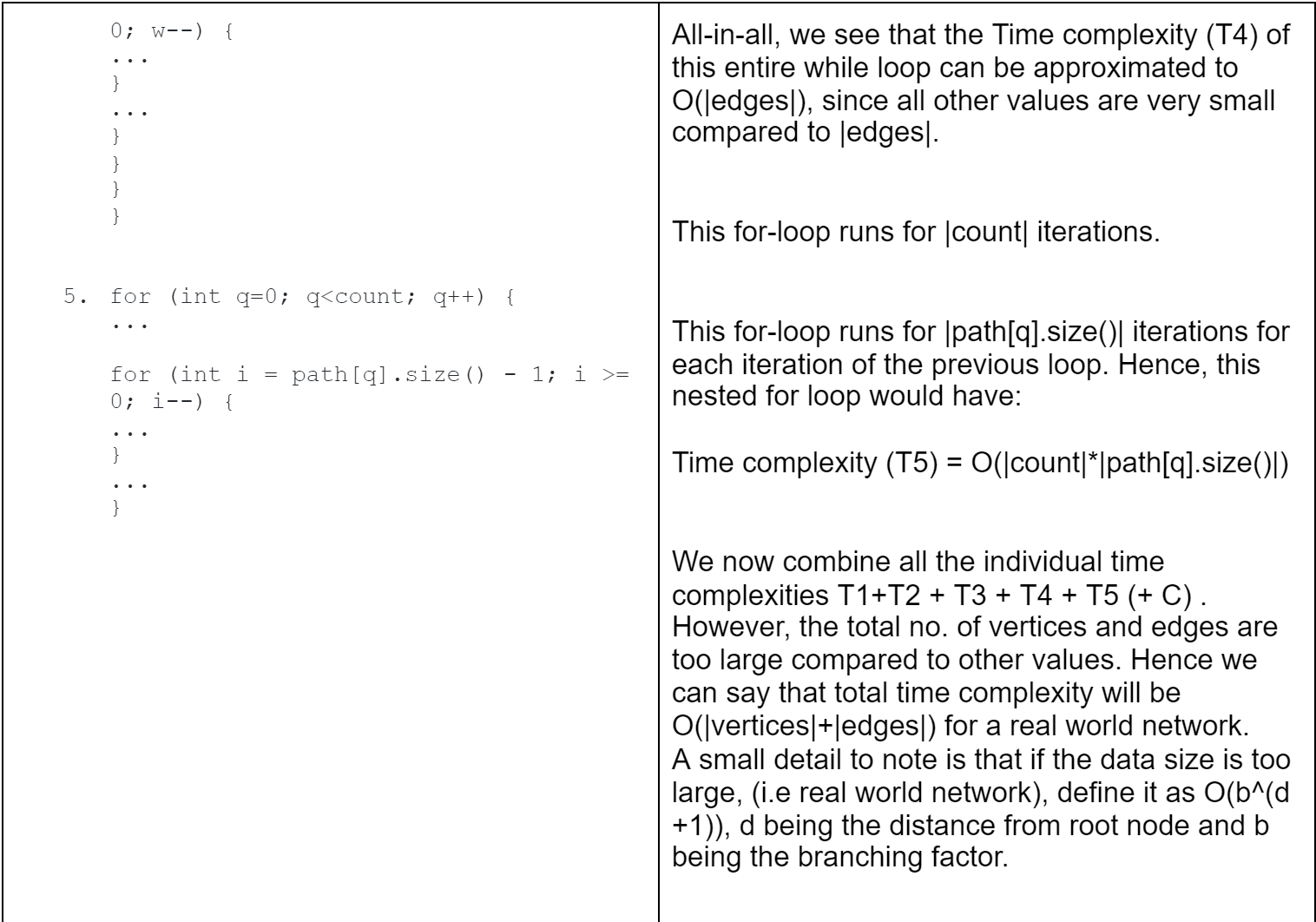
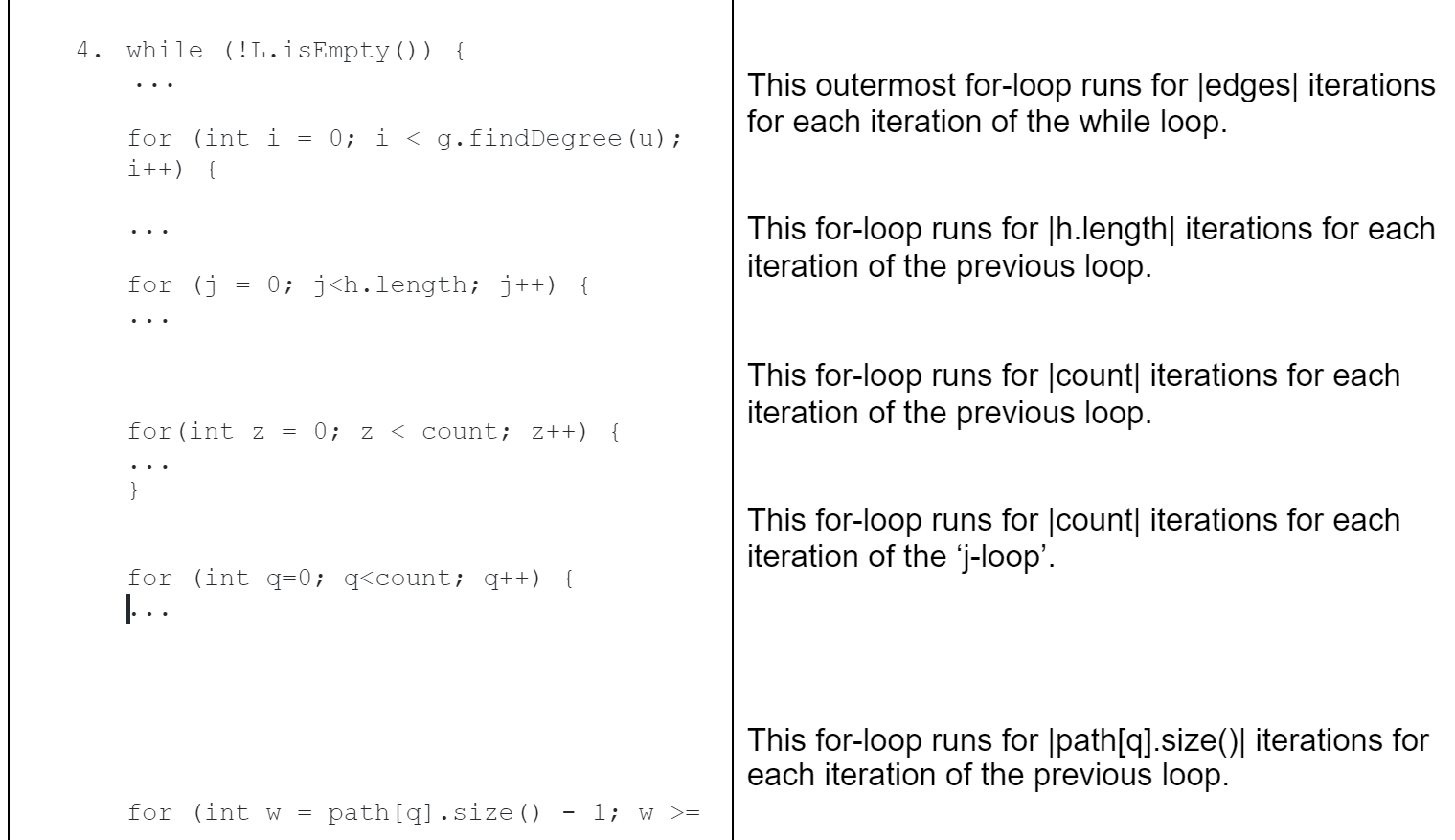
As such G=(V,E) , this will compute the shortest distance to any vertex as any path would corresponds to shortest path from root,v, to any other node S

**Our code**

Our code will similarly do the same concept just that the entire path to each node is stored and outputted onto the output.txt file and does not rely on the number of hospitals for the time complexity as a result. [Considering that the total no. of nodes and edges is much larger than the total number of hospitals]. Instead it relies on the total number of nodes inputted.(real world network estimate takes about 3 hours from theoretical calculation)

**Time Complexity**

Ultimately, the time complexity will be O(|V|+|E|) as the function will explore every single edge. |V| being the total number of vertices and |E| being the total number of edges in the graph. Hence it may vary if you have a lot of nodes between O(1) and O(|V^2|). If the files inputted are very large(real world network), the time & space complexity will result to *O*(*bd* + 1) ; b being the branching factor and d being the distance from the source node. This is merely an AOD(Average Out of Degree).

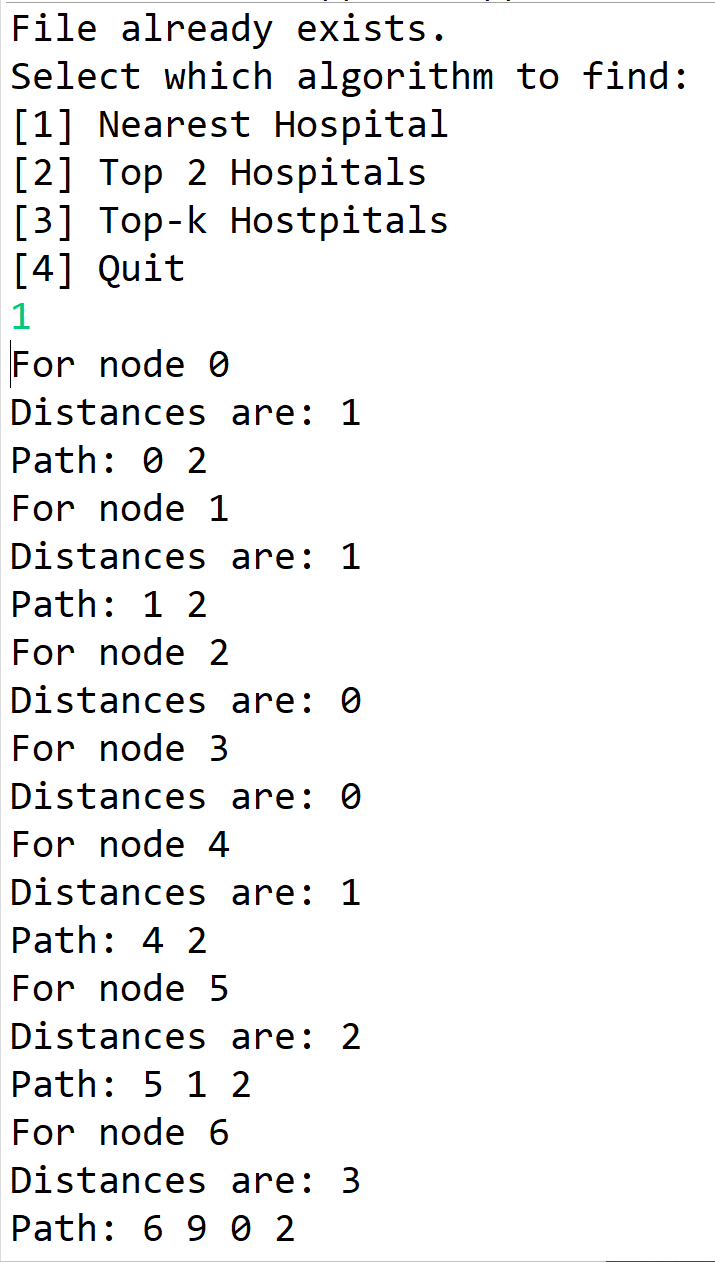
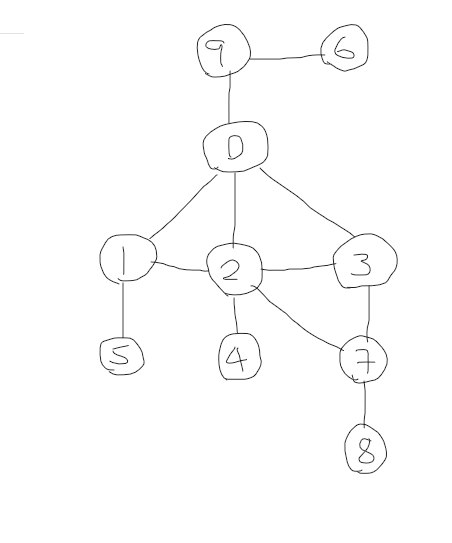


**Number of hospitals and locations:**

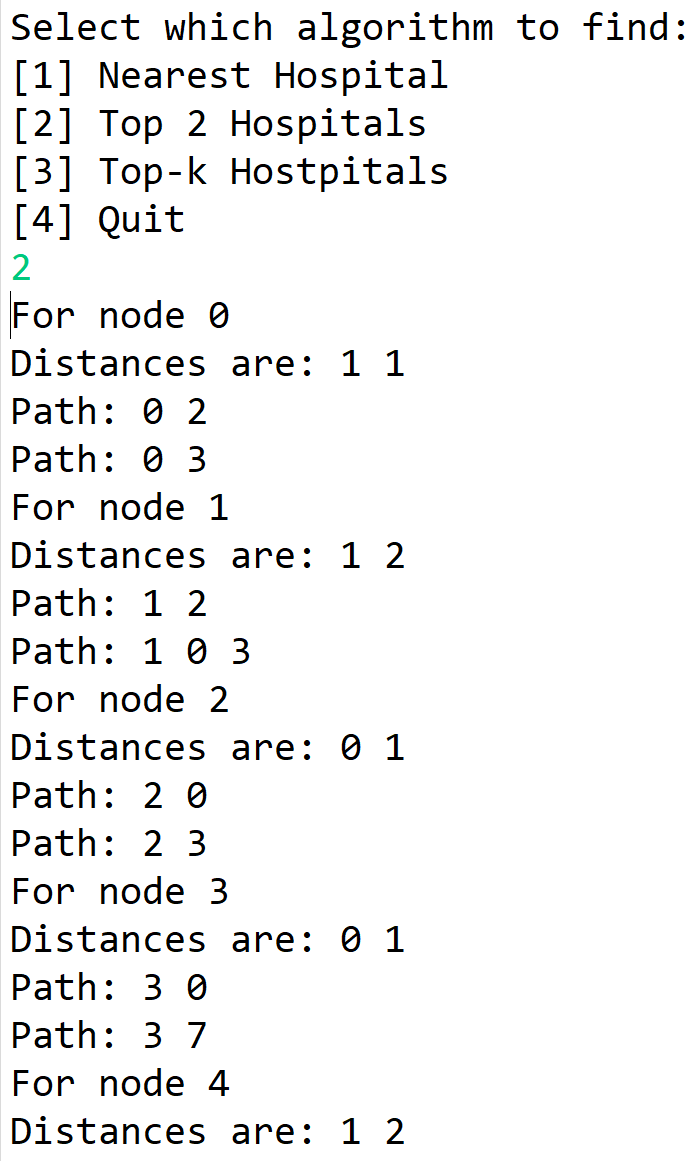
**Using the same set,for all the cases.**

**Assuming nodes 3,7,8,2 are hospitals**

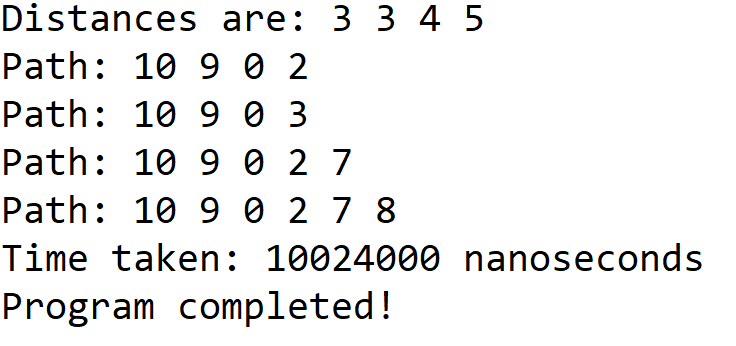
**Case(1a): (in example)**

Table output, code still goes onvisual representation(theoretical)

**Case(1c):**

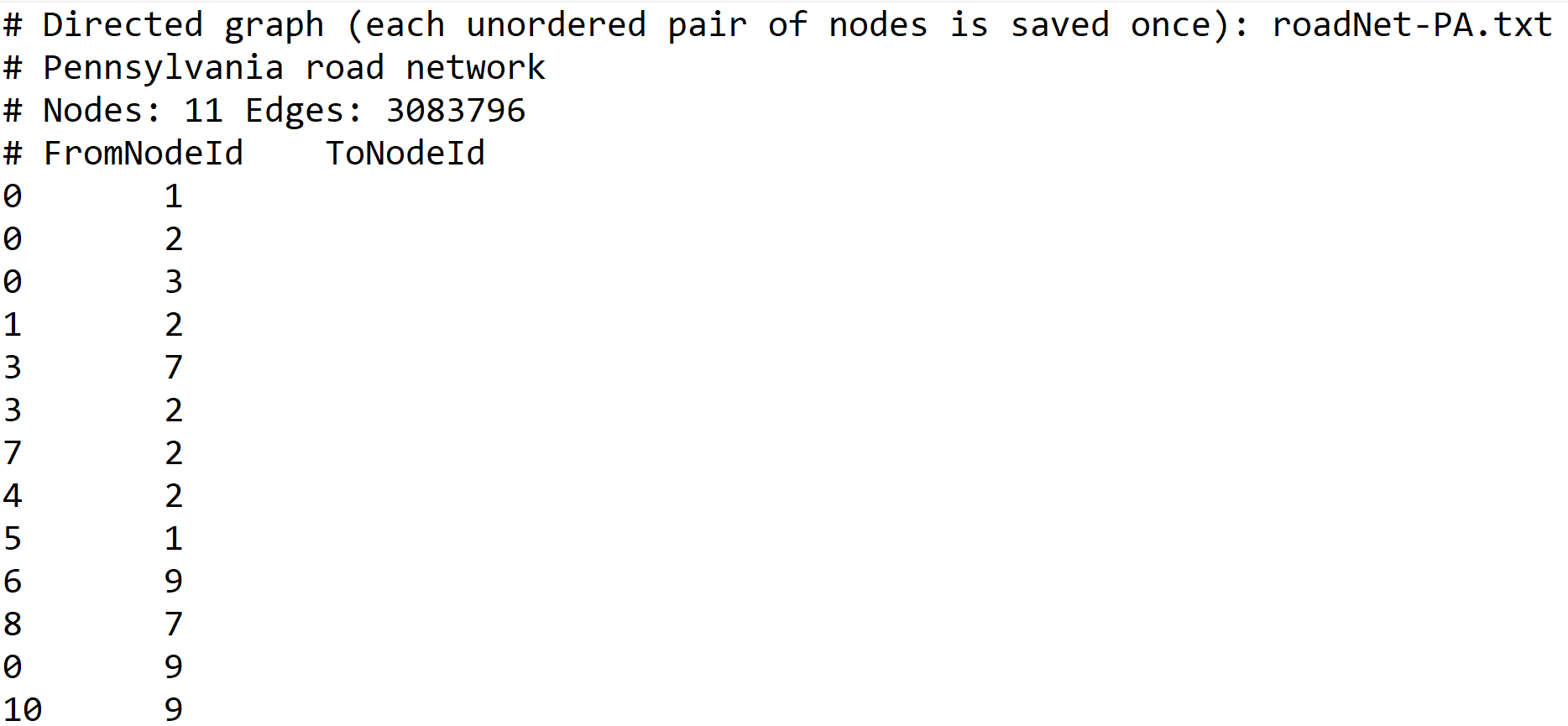
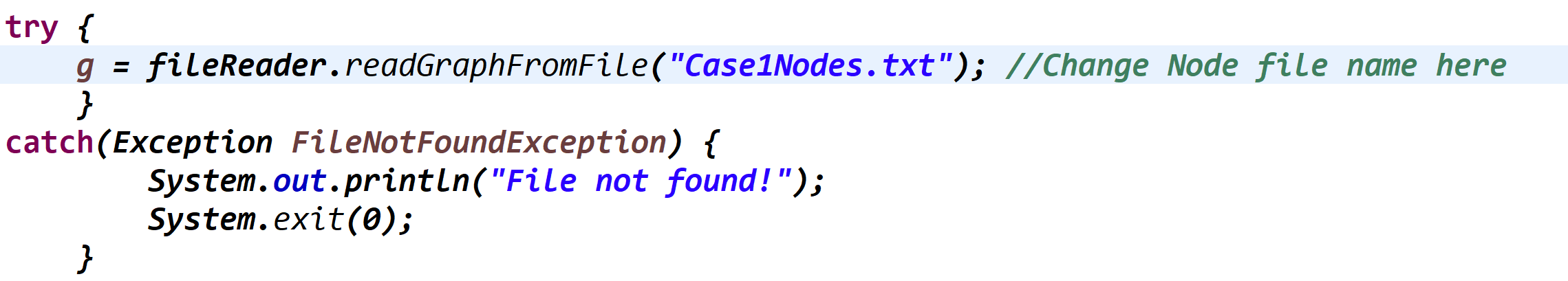
Top 2 nearest hospitals distance, this is a snippet

**Case(1d):**

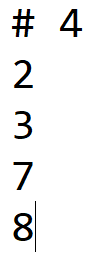
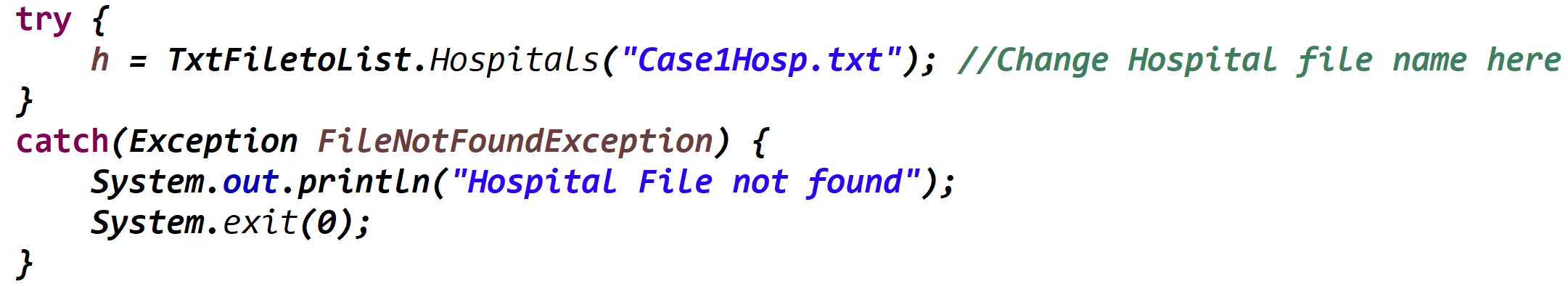
Top-K Hospitals all h nodes detected(output is much longer, this is the end result)

**Case(2)+(3):**

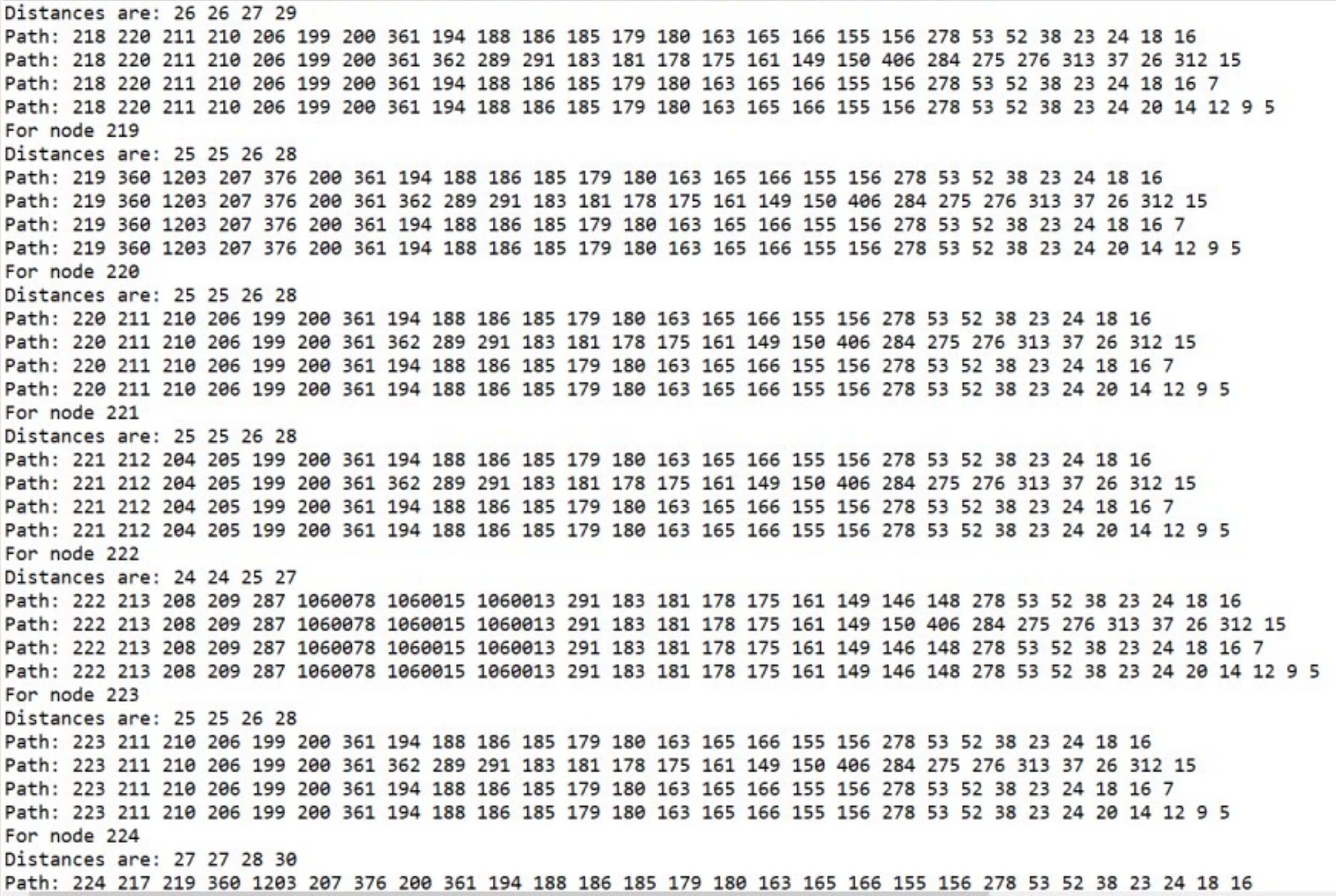
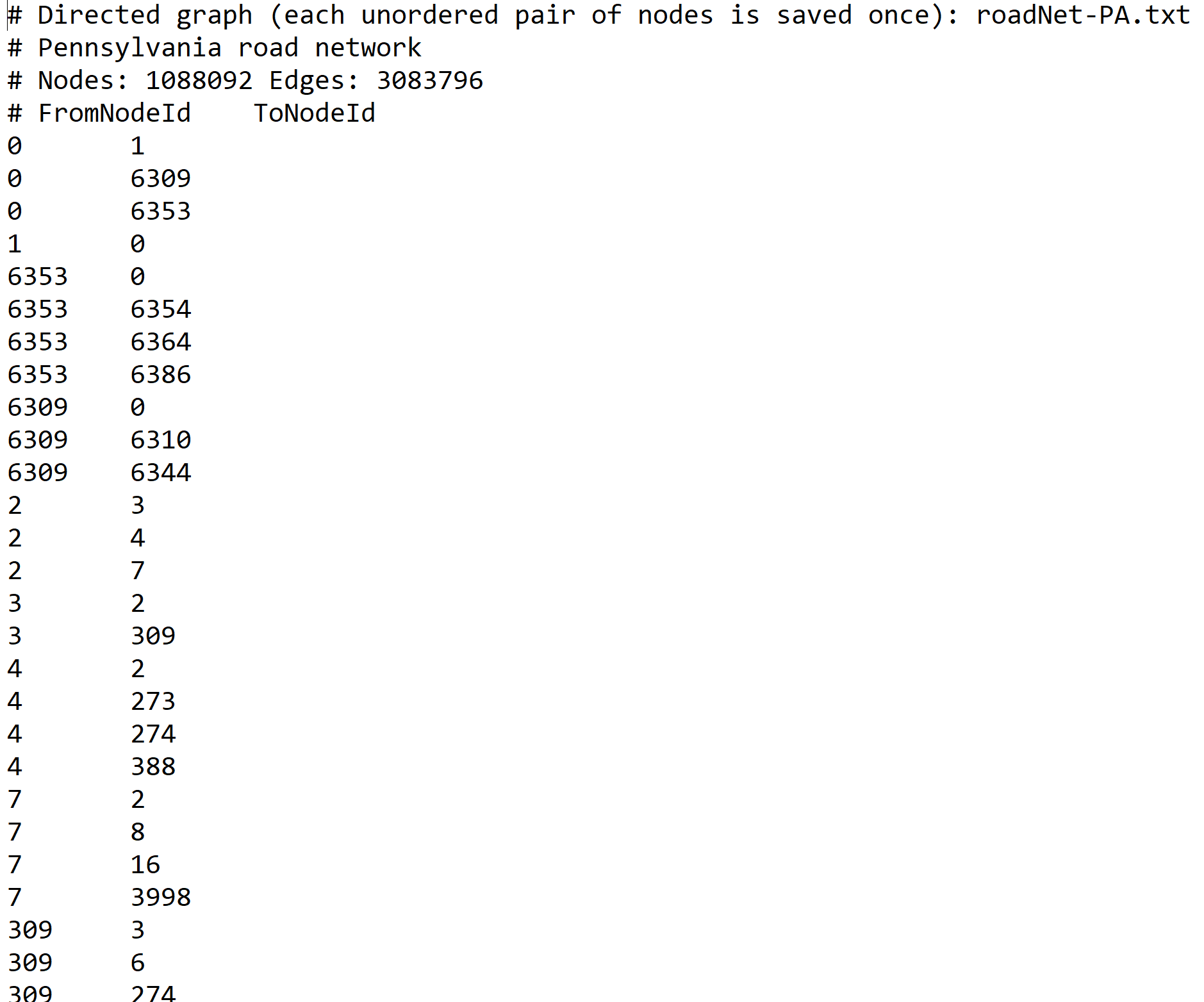
using a small input graph, the below image is for Case1Nodes.txt



The below images are the other file, fulfilling 2 file input(Case1Hosp.txt)



**Experiment(case 2):Testing with real world network:**

 Output will be very large and long This is merely a small snippet of the large real road network + what is inside the txt file as input

**Empirical Findings:**

**The question :** see the effects of h and k(input size) on the performance of various algorithms and conclude if the empirical findings are consistent with theoretic analysis.

**Assumptions:**

We take the same root node for all cases

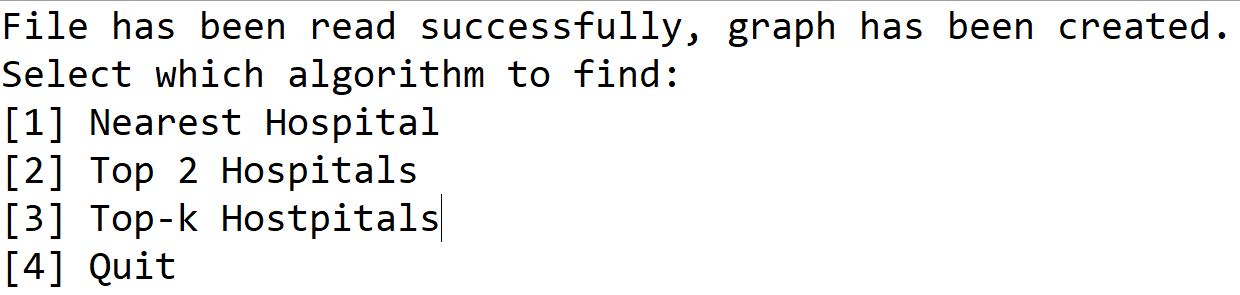
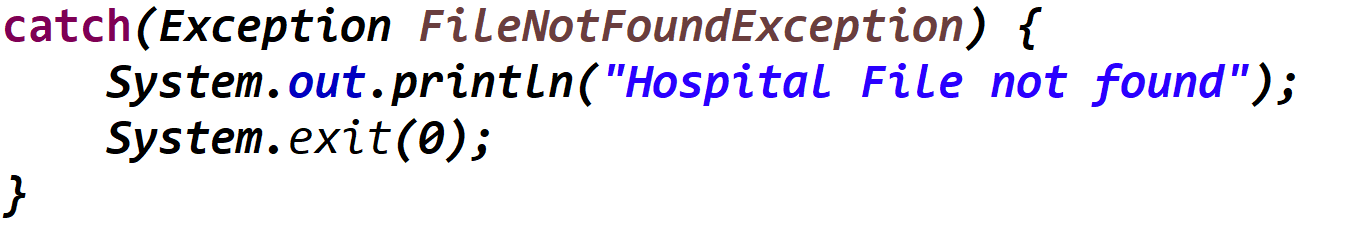
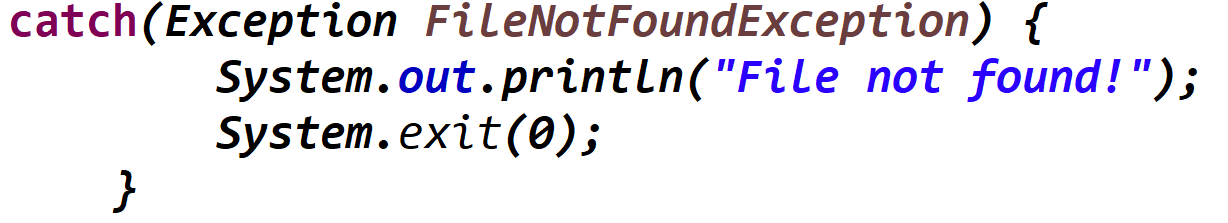
And the total number of nodes increases

**Found result:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Cases** | **h(hospitals)** | **K(top nearest number of hospital)** | **Empirical study BFS(total distance to h)** | **Various Algorithms(i.e BFS)(taken from online)** | **Theoretical Result** |
| **1** | **5** | **1** | **53** | **50** | **53** |
| **2** | **5** | **2** | **53, 54** | **50,51** | **53, 54** |
| **3** | **5** | **3** | **53, 54, 55** | **50,51,52** | **53, 54, 55** |
| **4** | **5** | **4** | **53, 54, 55, 55** | **50,51,52,53** | **53, 54, 55, 55** |
| **5** | **5** | **5** | **53, 54, 55, 55, 55** | **50,51,52,53,54** | **53, 54, 55, 55, 55** |

**Example of user-friendly interface:**

lab supervisor can test on arbitrary input graphs.



**References:**

<https://www.researchgate.net/publication/221302072_BFS_Solution_for_Disjoint_Paths_in_P_Systems/download>

<https://www.geeksforgeeks.org/depth-first-search-or-dfs-for-a-graph/>

<https://www.geeksforgeeks.org/shortest-path-unweighted-graph/>

Contributions:

Harsh worked on the code to read data from the hospital file. Kai Jun worked on the graph class and the code to read and create the graph from the node file. Zhe Ren and Kai Jun mainly implemented the BFS Algorithm, while everyone worked together on debugging it. Time complexity of the algorithm was derived by Harsh. Testing and documentation of the results were carried out by Mervyn.