Student ID: 010388223

Lngu242@wgu.edu

C950 - Data Structures and Algorithms II

WGUPS ROUTING PROGRAM

A. IDENTIFY ALGORITHM

In order to meet the program expectations, first we load any packages with special instructions onto the trucks. Then, by using the Nearest Neighbor Algorithm, we can sort through packages with similar distance to each other and load them onto the best truck accordingly. When all packages are fully loaded, the Greedy Algorithm is used to determine the best possible route for each package.

B1. ALGORITHM LOGIC

This program attempt to solve the problem with 3 steps. In the first step, we manually load by hand any packages with special instructions onto a truck without using any algorithm. In the second step, Nearest Neighbor Algorithm is used to load any remaining packages that have no special instruction onto the truck. The algorithm first looks for a truck with enough room for another package, then it checks whether the new package has similar delivery distance with packages that already on the truck. If the difference between distance of each package is minimal, it will load the new package onto the truck. When all packages are loaded, we are ready for the next step. In the third step, Greedy Algorithm is used to sort and organized all the packages in each truck for delivery in order to reach the most minimal distance by end of day. The algorithm first check through all available trucks and see if there is a package with the most minimal distance from current location of the truck. When it finds the best package, it will make that package prioritize and be first on the list for delivery. Then it will keep looping over and over again until all of the packages are sorted and organized for best possible outcome.

Second step – Nearest Neighbor Algorithm Pseudocode – Truck.py start from line 48

For any remaining packages that has not been loaded onto a truck yet

Declare variables: minimal distance, preferred truck for package to be loaded

Remove truck from sorting if it is full of 16 packages

For any truck with remaining space

For all packages currently on the truck with remaining space

Lookup that particular package's address ID currently on the truck

Compare and check the difference in distance between new and loaded

package

If the new distance is smaller than current most minimal distance, save its value

Load all the package onto the truck with the shortest distance in increasing order when all packages are checked on every truck accordingly

Third step – Greedy Algorithm Pseudocode – Truck.py start from line 75

For all trucks

Declare variables for most minimal distance, preferred package id, preferred address id

For all packages currently loaded on a truck

For all packages that have not been sorted through yet

Lookup that particular package's address ID

Check and compare the difference in distance of the current package

with other package that have not been sorted through yet

If the new distance is smaller than current most minimal distance, save

the new value

Sort through the package when all packages are checked by putting the package with the shortest distance to next stop first on the list

B2. PROGRAMMING ENVIRONMENT

This program was written with Python as the main programming language. The program was being developed within PyCharm 2022.2.3 (Community Edition), using Python version 3.11, in Windows 11 Version 22H2 (OS Build 22621.755). The data that the program used are 3 CSV file format and can be found in CSVFiles folder. The program can be run and accessed with the main entry point being the Main.py file.

B3. SPACE-TIME COMPLEXITY

Space-time complexity of major segment of the program and the entire program at worst case mainly is from the 2 named self-adjusting algorithm mentioned before. For the nearest neighbor algorithm, it is $O(n\log(n))$ and for the greedy algorithm, it is $O(n) + O(n^3)$. The algorithms was built with nested loops, which is the main reason why its space-time complexity is this way. Nested loops are needed for the ability to scale the application, increasing the number of trucks and packages further down the road is an option if we want to scale.

B4. CAPABILITY TO SCALE

Like mentioned above, the program is very flexible, and is adapted to scale if we want to increase the number of packages or the number of trucks further down the road. By using nested loops in the aforementioned algorithms, we can check every new package that loaded onto the truck and sort through all of them for the best possible outcome. When the data CSV files are updated accordingly with new packages, the program will still able to run normally without any problems. The hash table will be increased automatically if new packages are added, and since each and every package has its unique package ID, there should be no errors between all the packages.

B5. SOFTWARE EFFIENCY

Because the software is very flexible and easily to scale like mentioned in the previous section, it is very easy to maintain the software if we need to. The program is also filled with comments before each major parts. Therefore, it is easy to understand how the program works and for new users to adapt and adjust the codes if more trucks are added. Even though using 2 algorithms in the program may increase its time complexity, it is more efficient and more important to maintain the best possible route with the most minimal distance for the trucks to travel on, as this is the main goal in mind while developing this program. The program also can be able to run without any internet connection, it can be run on a local machine with no issues as all the data the program need is already provided in the CSVFiles folder. If we need to change the data for the program to run on, we can simply edit the CSV data filles in the aforementioned folder without worrying about database confliction or corrupted files over the air.

B6. STRENGTH AND WEAKNESS

A hash table is created when building the program in order to store package information and data. Like mentioned before, each and every package's information such as its ID is unique. And because of this, the program is very flexible and able to scale to adapt to adding more packages down the road. The hash table's size will also be able to increase automatically to adapt for more packages if added. This function can make the program become expensive computationally. However, we should keep in mind this kind of data structure can be proved to be useful when we do not know the number of packages for the program to run upfront.

C1. ORIGINAL PROGRAM AND IDENTIFYING COMMENT

The program's code is included within my submission. And an identifying comment with my information can also be found in the main entry point of the program – Main.py file.

C2. COMMENTS IN MY CODE

The program is already filled with comments in each and every .py files, thus make it easy to understand and maintain

D. DATA STRUCTURE

The program is built with a hash table being the main data structure used. Each and every package data is stored inside a list which is also inside the hash table array. The hash table add

support for package data to be added, updated, removed at any given time. Each package data like package ID is also unique on its own, and they are also used to determine the key for the hash table. And thus, there won't be any conflicts storing package data inside the hash table. The parallel in hash table key and package ID allow for data to be stored in a very structured way with ability to adapt and scale for future use.

E. HASH TABLE

The hash table is developed and can be located inside the Hash.py file. There, we can find its constructor, and also multiple functions to insert data, get data, update data, and remove data from the hash table. After the hash table is built, in Package.py file, we can see it is used alongside with WGUPS-Package-Table.csv file, in order to read the csv file and extract all the package data and information such as the package ID number, the delivery address, the deadline, the city, the zip code, the package weight and the delivery status of the package. All of these data are then added into the hash table with the insertVal functions for future use.

F. LOOK-UP FUNCTION

The look-up function is already built within the program in Main.py using user input and can be used when running the program with Main.py as the main entry point. When asked by the console, user can type "all" twice a row to view status report for all the packages at end of day. Or user can simply choose to loop-up a particular package at any given time just by typing a specific time followed by a particular package ID when asked by the console. For example, this screen shows how the look-up function work when user want to look-up package ID number 2 status at 9:00:00 AM. You can also find the full screenshot in Screenshots folder attached within this submission if needed.

```
C:\Users\Loc\Documents\Repos\C950-Loc_Nguyen\venv\Scripts\python.exe C:\Users\Loc\Documents\Repos\C950-Loc_Nguyen\Main.py
WGUPS Package Delivery System - Loc Nguyen
The estimated total distance of delivery for all trucks is 94.0 miles

Please enter 'quit' whenever you want to exit this program.

Please input a defined time in hh:mm::ss format in order to view it. Or simply type 'all' to view the completed route

Time: Program

Please input a package ID of choice or type 'all' to see all packages.

Package ID: 2

ID: 2 Address: 2530 S 500 E, Salt Lake City 84106 Weight: 44 Deadline: EOD Status: Delivered Delivery Time: 8:50:00

Please enter 'quit' to exit the program or press enter to continue:
```

G. INTERFACE

The interface for the program can be found when running the Main.py file. It first asks the user to input a defined time of choice. It then asks the user for a package ID to view that package's status report. When both inputs have been collected, a status report is generated showing the mentioned package's data and whether it is delivered or not. This is how the interface looks like when first run the program and before inputting anything. This image can also be found in the Screenshots folder within this submission

```
C:\Users\Loc\Documents\Repos\C950-Loc_Nguyen\venv\Scripts\python.exe C:\Users\Loc\Documents\Repos\C950-Loc_Nguyen\Main.py
WGUPS Package Delivery System - Loc Nguyen
The estimated total distance of delivery for all trucks is 94.0 miles

Please enter 'quit' whenever you want to exit this program.

Please input a defined time in hh:mm::ss format in order to view it. Or simply type 'all' to view the completed route
Time: |
```

G1. SCREENSHOT FOR STATUS CHECK NUMBER 1

Status report at 9AM for all packages

G2. SCREENSHOT FOR STATUS CHECK NUMBER 2

Status report at 10AM for all packages

```
380 W 2880 S,
         410 S State St, Salt Lake City 84111 Weight: 5
          1330 2100 S, Salt Lake City 84106 Weight: 8
300 State St, Salt Lake City 84103 Weight: 9
                                                                    Deadline: EOD
Deadline: EOD
                             Salt Lake City 84111 Weight: 2
       600 E 900 South,
2600 Taylorsville Blvd,
                            Salt Lake City 84118 Weight: 1
                                   Holladay 84117 Weight: 4
           1488 4800 S,
                                                                    Deadline: EOD
Deadline: EOD
                             Salt Lake City 84115 Weight: 3
                               Murray 84107 Weight: 7
                             Salt Lake City 84104 Weight: 5
            1330 2100 S,
                             Salt Lake City 84106 Weight: 2
          300 State St,
                             Salt Lake City 84103 Weight: 1
    1860 Dalton Ave S, Salt Lake City 84104 Weight: 88 Deadline: EOD 2300 Parkway Blvd, West Valley City 84119 Weight: 88 Deadline: EOD
    410 S State St, Salt Lake City 84119 Weight: 8
                             Salt Lake City 84111 Weight: 9
           2010 W 500 S, Salt Lake City 84104 Weight: 9 Deadline: EOD 380 W 2880 S, Salt Lake City 84115 Weight: 45 Deadline: 10:30:00
```

G3. SCREENSHOT FOR STATUS CHECK NUMBER 3

Status Report at 1PM for all packages

```
Please input a defined time in hh:mm::ss format in order to view it. Or simply type 'all' to view the completed route
Please input a package ID of choice or type 'all' to see all packages.
Package ID:
                                                                                      233 Canyon Rd, Salt Lake City 84103 Weight: 2
380 W 2880 S, Salt Lake City 84115 Weight: 4
                                                                                                                                                                                                                                 EOD Status: Delivered Delivery Time: 10:23.00
                                                                                        1330 2100 S, Salt Lake City 84106 Weight: 8 Deadline:
300 State St, Salt Lake City 84103 Weight: 9 Deadline:
                                                                                   600 E 900 South, Salt Lake City 84105 Weight: 1 Deadline: EOD Status: Delivered Delivery Time: 10:14:00
                                                                                                                                                                                                                                            Status: Delivered Delivery Time: 8:06:00
Status: Delivered Delivery Time: 8:13:00
                                                                                  4580 S 2300 E, Holladay 84117 Weight: 88
3148 S 1100 W, Salt Lake City 84119 Weight: 2
                                                                                                                                                                                                                                            Status: Delivered Delivery Time: 8:13:00
Status: Delivered Delivery Time: 9:43:00
                                                                         1488 4800 S, Salt Lake City 84123 Weight: 6 Deadline: EOD
177 W Price Ave, Salt Lake City 84115 Weight: 37 Deadline: EOD
3595 Main St, Salt Lake City 84115 Weight: 37 Deadline: 10:30:00
3595 Main St, Salt Lake City 84115 Weight: 3 Deadline: EOD
                                                                                                                                                                                                                                 EDD Status: Delivered Delivery Time: 11:25:80
EDD Status: Delivered Delivery Time: 9:19:80
                                                                                                                                                                                                                                            Status: Delivered Delivery Time: 8:30:00
Status: Delivered Delivery Time: 9:17:00
ID: 20
                                                             3595 Main St, See. Murray 84121 Weign.
6351 South 900 East, Murray 84121 Weight: 5
5100 South 2700 West, Salt Lake City 84118 Weight: 7
FA35 State St, Murray 84107 Weight: 7
                                                                                                                                        Murray 84121 Weight: 2 Deadline:
ake City 84118 Weight: 5 Deadline:
                                                                                                                                                                                                                                EOD Status: Delivered Delivery Time: 10:59:00
EOD Status: Delivered Delivery Time: 11:43:00
                                                        5025 State St, Murray 84107 Weight: 7
5383 South 900 East #104, Salt Lake City 84117 Weight: 7
                                                            5383 South 900 East #104, Salt Lake City 84117 Weight: 25 Deadline: EOD
1060 Dalton Ave S, Salt Lake City 84104 Weight: 5 Deadline: EOD
                                                                                    2835 Main St, Salt Lake City 84115 Weight: 7 Deadline: EOD StatUS: Delivered Delivery Time: 8:55:00
1330 2100 S, Salt Lake City 84106 Weight: 2 Deadline: 10:30:00 Status: Delivered Delivery Time: 9:12:00
                                                                                     300 State St, Salt Lake City 84103 Weight: 1 Deadline: 10:30:00 Status: Decivered Science,
3365 S 900 W, Salt Lake City 84119 Weight: 1 Deadline: 10:30:00 Status: Delivered Delivery Time: 8:35:00
Status: Delivered Delivery Time: 9:45:00
                                                                        3365 S 900 W, Salt Lake City 84114 Weight: 1 Deadline: EOD Status: Delivered Delivery Time: 9:45:80  
2530 S 500 E, Salt Lake City 84106 Weight: 1 Deadline: EOD Status: Delivered Delivery Time: 10:44:00  
4580 S 2300 E, Holladay 84117 Weight: 2 Deadline: 10:30:00 Status: Delivered Delivery Time: 8:15:00  
1060 Dalton Ave S, Salt Lake City 84104 Weight: 88 Deadline: EOD Status: Delivered Delivery Time: 11:20:00  
2300 Parkway Blvd, West Valley City 84119 Weight: 88 Deadline: EOD Status: Delivered Delivery Time: 9:34:00  
410 S State St, Salt Lake City 84111 Weight: 2 Deadline: EOD Status: Delivered Delivery Time: 9:09:00  
410 S State St, Salt Lake City 84111 Weight: 9 Deadline: EOD Status: Delivered Delivery Time: 10:20:00  
Add S State St, Salt Lake City 84111 Weight: 9 Deadline: EOD Status: Delivered Delivery Time: 11:15:00
```

H. SCREENSHOT FOR SUCCESSFUL COMPLETION OF CODE

Screenshot when executing the program and free from runtime errors or warning. Total milage traveled by all trucks are included

```
Run:

| Main | M
```

11. STRENGHTS OF ALGORITHM

By using both the nearest neighbor algorithm and the greedy algorithm on top of each other, it enables the program to find the most minimal distance and most optimized route even when more packages are added, and it can deal with packages with weird special instruction. As we can see from the status report, using those algorithms allow the trucks to delivered all required packages way before their deadline and in a very optimized route with minimal distance (less than 140 miles).

With the use of these two algorithms, we can also scale the program and increase the number of packages or trucks with ease of access as well as maintainability. Both algorithms are very simple and two of the most used techniques in sorting and organizing, and they do not require much hardware resource. Simplicity is key when we want to focus on efficiency and scalability down the road.

12. VERIFICATION OF ALGORITHM

Inside the nearest neighbor algorithm in Truck.py line 54, we can see the implementation to remove truck from sorting if it is full of 16 packages. In Truck.py line 125, we take into account that the truck can only travel at an average speed of 18 miles per hour, and use it to calculate the distance that the trucks can travel. The program is also built with no collision and does not take any collision into account. At any given time, there is also a maximum of 2 trucks being on the road. The truck does not start delivering package until 8:00:00 AM. The program also does not take into account any delivery and loading time, and assume delivery and loading time are instantaneous. All packages with special instructions are also manually dealt with. The day also end when the last package is delivered at 11:43:00 AM

13. OTHER NAMED ALGORITHMS

Two other named algorithms that can be used to solve this problem is Brute-Force and using Dijkstra algorithm. With Brute-Force algorithm, we can calculate and find every possible combinations of paths and route between each and every packages inside the program. The total amount of possible paths is calculated to be n! . Then, we just have to choose the most optimized path with the most minimal distance. The space-time complexity of using Brute-Force algorithm would be O(n!), which is vastly different from Nearest Neighbor + Greedy Algorithm that we have chosen. It is quite straight forward and can sometimes be seen as the naïve approach. The main disadvantage of this is it is quite extensive to look through all possible paths, especially when the number of trucks or packages will be increased if our main goal is to scale the application.

One other possible algorithm that can be used is Dijkstra, like mentioned before. It can be used to calculate and find the most optimized path by loading the delivery address into a weight graph, then calculate the most minimal distance by using points that are not at delivery address. Whereas in Greedy Algorithm, we calculate the path by using routes in between delivery address.

J. WHAT TO DO DIFFERENTLY IF REDO PROJECT

One possible thing to do things differently if we can redo this project is adding a SQL database to store all the packages data with delivery address and truck information. And after running the program, a report will be uploaded onto the cloud. That way, the data can be easily modified externally, and other users can view the status report without having the program on their local machine. Another modification that can be made is creating a better user interface that does not output the report on the console, but rather in CSV or similar format that can then be uploaded and saved automatically onto the cloud like mentioned before.

K1. VERTIFICATION OF DATA STRUCUTRE

The data structures used for the program is creating a hash table to store package data, it is built without using any external libraries. And it is very flexible and able to adapt and scale to store any number of packages. The look-up function of the program is directly affected by the hash table as it uses the hash table getVal function in order to retrieve package data. However, the time needed to complete the look-up function is always constant and does not depends on how many packages are stored inside the hash table. As for data structure space usage, as more packages are stored, the hash table size will also be increased automatically to adapt any new packages. And thus, more packages added also means more space will be taken, and will increase the total amount of memory consumption. As the program is very flexible, adding more trucks would have very minimal or no effect on the look-up time and space usage of the data structure. Adding more cities would also have minimal or zero effect to look-up time and space usage as well, as city and address data are only used to calculate distance between each packages. Comparing to increasing the total amount of packages, it would be safe to say that increasing number of trucks or cities does not have that much significant of an impact like the total amount package would.

K2. OTHER DATA STRUCTURES

Two other possible data structure that can be used within this program is circular linked list and weighted graph. With weighted graph, it is possible for even better time-space complexity as well as smaller estimated distance traveled by the trucks. Instead of just using direct path like in our current application, we can just find shortcuts between package destinations.

In real world situation, where we would also take into account the distance for the trucks to travel back to the hub, circular linked list would be the perfect ideal data structure to be used as we can simply implement each address of the program into a circular linked list. And after delivering the last package for the day, the truck will need to be back at its starting position which is the hub. However, in order to implement the same "get-update-delete-insert" functions like we are doing with hash table, we would also need an indexing system for a circular linked list.

L. SOURCES

WGU. (n.d.). *C950: Data Structures and Algorithms II*. Zybooks. Retrieved November 5, 2022, from https://learn.zybooks.com/zybook/WGUC950AY20182019