

C950 Data Structures and Algorithms II

#### Introduction

This document will go over the algorithm and code I wrote to fulfill the task of delivering all given packages within the provided constraints.

## A. Algorithm Identification

The algorithm I chose is the Nearest Neighbor algorithm. This algorithm searches through a list of distances from a starting point and chooses the closest one to go to next. It then removes the selected destination from the list and repeats the search from the new location. It does this until all locations in the original list have been visited. I modified this algorithm slightly by multiplying the measured distance from a point to any destination that did not have a deadline attributed to it. I did this to create a rudimentary prioritization of time sensitive deliveries that would not accidentally skip over "low priority" packages at the same address.

# **B1.** Logic Comments

This pseudocode represents the code that determines the closest package in a list using a table with distances between any two addresses, weighted negatively if the package does not have a delivery deadline. By multiplying the distance of packages without deadlines by 5, packages with deadlines will appear relatively closer to the algorithm. Weighted distances will still be zero for a second package going to the address the algorithm is comparing from, and very close weighted destinations will similarly tend to outrank distant unweighted destinations, but in general, packages with deadlines are heavily favored. The algorithm has more variables so that it can keep track of both actual and weighted distances.

## **FUNCTION** deliver(truck):

SET initial distance to 1000
SET start location to Hub
CREATE next\_package variable
CREATE traveled variable

WHILE list of packages on truck has items

FOR each package ID in truck's package list

SET actual\_distance to the distance from start

SET weighted\_distance to comp\_distance

IF package deadline is EOD

MULTIPLY weighted\_distance by 5

IF weighted\_distance is less than distance

SET distance to weighted distance

SET travel\_distance to actual\_distance

SET next\_package to package ID

**THEN** 

ADD travel\_distance to traveled variable

SET next\_package departure time to truck departure time

SET next\_package delivery time to departure time PLUS minutes in transit

REMOVE next\_package from truck's package list

UPDATE start location to next\_package location

RESET distance to 1000

**THEN** 

SET truck mileage to traveled PLUS distance back to Hub
SET truck return time to departure time PLUS minutes in transit

# B2. Development Environment

PyCharm Community Edition 2023.1

Python 3.11.2

Windows 11 Desktop PC

I used PyCharm to develop this application on my PC. I did not download or use any extensions for it.

# B3. Space-Time and Big-O

The hash table has a space complexity of O(n) and the functions in it have O(1) time complexity. The deliver(truck) function has a time complexity of  $O(n^2)$  because the function has a for loop nested inside of a while loop. The function that prints has a time complexity of O(1) if printing the information of a single package, and O(n) if printing all because print all utilizes a for loop.

The program has an overall time complexity of O(n²) based on the deliver function.

## B4. Scalability and Adaptability

This program would have to be expanded a bit to scale up. I loaded the trucks manually, and while it would be possible to manually load more packages into more trucks, that is not a scalable solution. I would need to write code to truly interpret delivery deadlines, priority levels, special notes, manage drivers vs trucks, and truck reloads before it could be considered scalable. Also, because the trucks were manually loaded, it does not adapt if the package list is changed, the above mentioned functions would provide that adaptability.

## B5. Software Efficiency and Maintainability

My software is efficient and easy to maintain because it is minimal, simple, and has comments explaining what the code does. The code is around 250 lines including all whitespace and comments, and is contained across just 4 files. If it has to be changed for some reason, it will be

easy to do so. It runs relatively low in Big-O time for the most part, the deliver function is  $O(n^2)$ , but n is limited to 16.

# B6. Self-Adjusting Data Structures

The strengths of the hash table are how quickly it deals with the data in it and how easily it handles collisions. This means it is well suited to being scaled up. The minor downside is that its size needs to be initiated correctly to balance performance versus wasted space. A larger capacity reduces collisions and consequently speeds performance, but unused buckets do waste some space.

# C. Original Code

Code in attached files.

#### C1. Identification Information

See code.

## C2. Process and Flow Comments

See code.

#### D. Data Structure

The self adjusting data structure is the chaining hash table on HashTable.py.

## D1. Explanation of Data Structure

The hash table is based on the chaining hash table outlined in the zyBooks material (Lysecky, section 7) provided. It has been initialized to a size of 17 buckets, fewer than the number of packages, to guarantee it is capable of handling collisions. The unique key is the package ID number, which is linked to a bucket based on the key modulo 17. For insert, search, and remove functions, the bucket is selected with key modulo 17, then a for loop checks the bucket for a key match. A list function compiles a list of all the key values in all the buckets.

The hash table allows either an entire package or any desired element of a package to be accessed with the HashTable.search(key) or HashTable.search(key).element commands.

#### E. Hash Table

HashTable.py

## F. Look-Up Function

Main.py nice\_output()

## G. Interface

## Main.py

#### G1. First Status Check

All packages at a time between 8:35 a.m. and 9:25 a.m.

```
All packages at a time to energy special (norms augho) or type "quit" the unit.

There is passing to be uniter of the packages at a time to lower or "Lift" to be interested to the seasons.

**ROTEL M package II, suppling I kilos, destined for 10% & Soluted Ave, Salt Lawe City II, Solid by 10% of the yet cost for delivery, But is scheduled to service at 10% M at 60% II M package II, suppling I kilos, destined for 10% & Soluted Ave, Salt Lawe City III, Solid by 10%, Is out for elivery and is scheduled to service at 10% M at 60% II M package II, suppling I kilos, destined for 20% Casyon Re, Salt Lawe City III, Solid by 10%, Is out for elivery, But is scheduled to service at 10% M at 60% II M package II, suppling I kilos, destined for 20% Casyon Re, Salt Lawe City III, Solid by 10%, Is out for elivery and is scheduled to service at 10% M at 60% II M package II, suppling I kilos, destined for 20% Casyon Re, Salt Lawe City III, Solid by 10%, Is out for elivery and is scheduled to service at 10% M at 60% II M package II, suppling I kilos, destined for 20% Casyon Re, Salt Lawe City III, Solid by 10%, Is out for elivery and is scheduled to service at 10% M at 60% II M package II, suppling I kilos, destined for 20% Casyon Re, Salt Lawe City III, Solid by 10%, Is out for elivery and is scheduled to service at 10% M at 60% II M package II, suppling I kilos, destined for 10% Casyon Casyon Re, Salt Lawe City III, Solid by 10%, Is out for elivery and is scheduled to service at 10% M at 60% II M package III, suppling I kilos, destined for 10% Casyon Casyon Re, Salt Lawe City III, Solid by 10%, Is out for elivery at its scheduled to service at 10% M at 60% II M package III, suppling I kilos, destined for 10% Casyon Salt Lawe City III, Solid by 10%, III, Solid by 10% III, Solid by
```

#### G2. Second Status Check

All packages at a time between 9:35 a.m. and 10:25 a.m.

#### G3. Third Status Check

## All packages at a time between 12:03 p.m. and 1:12 p.m.

## H. Screenshots of Code Execution

```
C:\Users\Aaron\PycharmProjects\C950\venv\Scripts\python.exe C:\Users\Aaron\PycharmProjects\C950\main.py
All packages delivered and trucks returned at 12:25 PM , with a total of 116.9 miles.
Input a time to check package status (hh:mm am/pm) or type 'quit' to exit. and
Enter a package ID number or 'all' to see information for those packages. 37
At 12:25 PM package #37, weighing 2 kilos, destined for 410 S State St, Salt Lake City UT, 84111 by 10:30 AM, was successfully delivered at 10:04 AM
Input a time to check package status (hh:mm am/pm) or type 'quit' to exit. and input
That wasn't what I was expecting, try it again.
Input a time to check package status (hh:mm am/pm) or type 'quit' to exit. 10:10 and
Enter a package ID number or 'all' to see information for those packages. 77
There doesn't seem to be a package with that ID number.
Input a time to check package status (hh:mm am/pm) or type 'quit' to exit. 10:10 and
Enter a package ID number or 'all' to see information for those packages. 7
At 10:10 AM package #7, weighing 8 kilos, destined for 1330 2100 S, Salt Lake City UT, 84106 by E00, was successfully delivered at 08:48 AM
Input a time to check package status (hh:mm am/pm) or type 'quit' to exit. 10:15
```

#### Demonstration of:

- 1. Initial summary and input prompt
- 2. Status of package 37 at EOD
- 3. Return to input prompt
- 4. Handling of bad input
- 5. Handling of incorrectly queried package (package ID 77)
- 6. Status of package 7 at 10:10 AM
- 7. Successful quit

# I1. Strengths of Chosen Algorithm

The primary strength of the nearest neighbor algorithm in this project is that it is easy to understand. I was able to conceptualize and write it quickly in a way that is simple to read and understand. If this was an application that was being utilized by others, the simple straightforwardness of it would benefit other users when they examine the code. Another strength is that it can easily handle more packages if the capacity of the trucks is ever increased. This is a positive for scalability.

## I2. Verification of Algorithm

The itself algorithm meets the requirements that all packages are delivered before deadlines and total truck mileage is less than 140. In loading the trucks and setting their depart times manually, I was able to account for all special notes, 16 package max capacity per truck, only 2 trucks being driven at a time, and trucks leaving no earlier than 8:00 AM.

## I3. Other possible Algorithms

I could have alternatively used either Dijkstra's algorithm or brute force to find the shortest route, though brute force is not a reasonable option.

## I3A. Algorithm Differences

Dijkstra's algorithm is different from the nearest neighbor because uses nodes on a graph and it calculates the shortest route through all the nodes from a starting node to an ending node, and the

nearest neighbor is a greedy algorithm that only sees two locations at a time. Dijkstra would probably deliver all the packages in fewer miles than I did.

Brute force would guarantee the optimal route by checking every option, but it would be very resource intensive and slow having a time complexity of O(n!). It is technically optimal but not a reasonable option.

# J. Different Approach

If I were to do this project again, I would invest the time to create an automatic truck loader that could load the trucks in an efficient way. This would give the program much more scalability.

#### K1. Verification of Data Structure

The chaining hash table was very well suited to the needs of this program.

## K1A. Efficiency

The time needed to look up data in the chaining hash table does not change much as the number of packages increases. I intentionally set the number of buckets small so that I could have and handle collisions for my own edification, but in a more realistic implementation, the number of buckets would be much larger, and searches would average O(1) time.

#### K1B. Overhead

The space complexity of the chaining hash table is O(n), so space usage scales proportionately to the number of packages.

## K1C. Implications

Adding or removing trucks would not have an effect on the hash table space or look up time, but would have a linear impact on the space usage of trucks.

Adding more destinations would also not have an effect on the hash table space or look up time, but would increase the space usage of the distance map exponentially.

#### K2. Other Data Structures

A linked list or an AVL tree could have been used instead of the hash table.

## K2a. Data Structure Differences

A linked list would always perform searches in O(n) time and insertions or deletions in O(1)time, compared to a hash table that does all in an average of O(1) time but in O(n) time at worst. Considering that the function I used most in my program was the search function, the fast insertion and deletion times do not make up for the search time being equivalent to a hash table's worst case time.

An AVL tree would perform all the operations in  $O(\log n)$  time, which is worse than the O(1) time hash table functions usually take, but is better than the worst case O(n) time of the hash table. If the hash table has an appropriate number of buckets, it should outperform an AVL tree.

## M. Professional Communication

# L. Sources - Works Cited

Lysecky, R., & Vahid, F. (2018, June). C950: Data Structures and Algorithms II. zyBooks.

Retrieved March 22, 2021, from

https://learn.zybooks.com/zybook/WGUC950AY20182019/