INFO 6205 Spring 2023 Project Traveling Salesman

Team Members:

Ramya Devie Saravana Bhava <u>saravanabhava.r@northeastern.edu</u> 002770163 Varun Shetty <u>shetty.varu@northeastern.edu</u> 002778981

Introduction

- **Aim:** Implement the TSP (Travelling Salesman Problem) using Christofides algorithm and optimize the output using any two of the tactical (2-opt, 3-opt, random swapping algorithm) and strategic optimization (simulated annealing, ant colony, genetic algorithm).
- **Approach:** The Travelling Salesman Problem has been deemed as an NP-Hard Problem ever since its mathematical formulation in the 19th century. A brute force approach to solve this problem would be to consider each vertex, calculate the distance between this node and every other node, and then pick a route by analyzing the distances between all the available nodes using these same steps repeatedly.

This approach would end up giving us the best route possible. However, this approach would only work if the number of vertexes is low. The number of possible routes this method would have to analyze would be equal to (N-1)! /2 which signifies a substantial number for even a relatively small value of N like 12. Therefore, since finding an optimal solution for this problem is practically impossible, we use a different approach which would provide us with a reasonably good solution.

Our approach in this project is to utilize the Christofides Algorithm to find a reasonable solution and then later optimize this solution further using various tactical and strategic optimizing techniques. The solutions that will be provided would not be optimal, however, the length of the routes will be within 50% of the MST (Minimum Spanning Tree) Length which would be an acceptable solution for all practical purposes.

To solve the TSP, our first step would be to implement the Christofides Algorithm to find a route. To achieve this, we begin with creating an Adjacency Matrix which would have a 2-D Array containing the distance between each node. Using this 2D graph, we would proceed to find the Minimum Spanning Tree between the Nodes using the Prim's algorithm. Once we have obtained the MST, we would proceed to find the number of Odd Vertices in this MST along with the Minimum weight Matching Techniques to make sure that each node has an even degree of freedom. Finally, we would find the Eulerian tour of this Graph and find the route which would be an acceptable solution to this problem.

This tour will then further be optimized using various optimizing techniques such as random Swapping algorithms like 2-Opt, 3-Opt along with some strategic Algorithms such as Ant colony Optimization and Simulated Annealing. These techniques will be further discussed below.

Program

- **Data Structures:** Array list, Array, Custom data class (Location)
- Classes:
 - 1. Edge
 - 2. Location
 - 3. Christofides
 - 4. TwoOpt
 - 5. ThreeOpt
 - 6. AntColonyOptimization
 - 7. SimulatedAnnealing
 - 8. Prims
 - 9. EdgeWeightedGraph
 - 10. Bag
 - 11. TSP

• Algorithm:

The algorithm used to implement TSP here is Christofides algorithm with two different tactical optimization such as 2-opt and 3-opt and with two different strategic optimization such as simulated annealing and ant colony optimization. Our objective here is to observe and compare the different optimizations applied to the TSP (Travelling Salesman Problem) using Christofides algorithm. We find MST of the given graph and then create a subgraph that includes only vertices of odd degrees. After finding minimum weight perfect matching, their edges and MST edges are combined to form a multigraph. Then we find eulerian circuit and convert it into Hamiltonian circuit by skipping repeated vertices.

Now, to optimize the output of the Christofides algorithm, we can use two types of optimization algorithms: tactical and strategic. Tactical optimization algorithms focus on improving the current solution by making local changes, while strategic optimization algorithms explore the search space to find better solutions.

Two common tactical optimization algorithms for the TSP are the 2-opt and 3-opt algorithms. The 2-opt algorithm works by removing two edges from the current solution and reconnecting them in a different way if doing so will shorten the total distance. The 3-opt algorithm works similarly, but it removes three edges and reconnects them in different ways. Simulated annealing works by randomly perturbing the current solution and accepting the perturbation if it improves the solution, but also sometimes accepting perturbations that worsen the solution to escape local optima. Ant colony optimization is inspired by the behavior of ants searching for food. In this algorithm, artificial ants deposit pheromones on the edges they travel, and the probability of choosing an edge is based on the amount of pheromone on the edge.

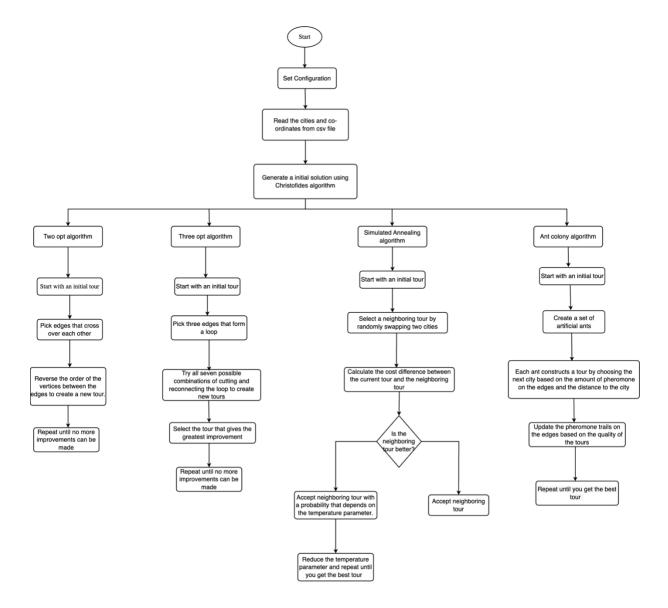
- **Invariants:** Invariant is a function, quantity, or property which remains unchanged when a specified transformation is applied.
 - **TSP:** The Travelling Salesman Project asks for the shortest possible routes that visits a given set of cities and returns to the starting city. The problem is NP-hard, meaning that no polynomial-time algorithm is known that can solve all instances of the problem. Some of the invariants of TSP are as follows.
 - i. Symmetry: The distance between two cities is the same in both directions.
 - ii. Triangle inequality: The distance between two cities is always less than or equal to the sum of the distances between those cities and any other city in the problem. Mathematically, this means that $d(i, j) \le d(i, k) + d(k, j)$ for all cities i, j, and k.
 - iii. Non-negativity: The distance between any two cities is non-negative.
 - iv. Unique solution: There is only one optimal solution to the TSP.
 - Christofides algorithm: The Christofides algorithm is a heuristic algorithm that provides a solution to the TSP that is guaranteed to be within a factor of 3/2 of the optimal solution. The algorithm has several invariants that are used to ensure the quality of the solution:
 - i. Minimum spanning tree: The first step of the algorithm is to find the minimum spanning tree (MST) of the complete graph of the cities. It is done using Prim's algorithm here.
 - ii. Odd-degree vertices: In the MST, some vertices will have odd degree (i.e., an odd number of edges incident to them). The second step of the algorithm is to find a minimum-weight perfect matching among the vertices with odd degree.
 - iii. Eulerian tour: The third step of the algorithm is to combine the edges of the MST and the minimum-weight perfect matching to form a connected graph that contains each vertex twice (once for the outgoing edge and once for the incoming edge). This graph can be shown to have a Eulerian tour, which is a path that visits each edge exactly once.
 - iv. Shortest path: The fourth and final step of the algorithm is to traverse the Eulerian tour, skipping any vertices that have already been visited, and computing the shortest path between adjacent unvisited vertices. The result is a Hamiltonian cycle that visits each vertex exactly once, and whose length is within a factor of 3/2 of the optimal solution.

All the optimization algorithms have the below invariants in common.

- i. Tour representation: The solution is represented as a tour that visits each vertex exactly once and returns to the starting vertex.
- ii. Improvement: The 2-opt algorithm only makes moves that reduce the total length of the tour, ensuring that the solution is continuously improving.
- **Two-opt algorithm:** The 2-opt algorithm is then used to optimize the solution obtained from the Christofides algorithm by iteratively removing two edges from the tour and reconnecting the resulting subpaths in a way that reduces the total length of the tour. The invariants used in the 2-opt algorithm are:
 - i. Validity: At any point during the optimization process, the solution must still satisfy the invariants of the TSP.

- Three-opt algorithm: The 3-opt algorithm is then used to optimize the solution obtained from the Christofides algorithm by iteratively removing three edges from the tour and reconnecting the resulting subpaths in a way that reduces the total length of the tour. The invariants used in the 3-opt algorithm are:
 - i. Validity: At any point during the optimization process, the solution must still satisfy the invariants of the TSP.
- ii. Feasibility: The 3-opt algorithm only makes moves that result in a feasible solution, meaning that the resulting tour is still valid.
- **Simulated Annealing algorithm:** The SA algorithm is then used to optimize the solution obtained from the Christofides algorithm by iteratively changing the tour using random moves, and accepting moves that reduce the total length of the tour or moves that increase the total length of the tour with a certain probability. The invariants used in the SA algorithm are:
 - i. Validity: At any point during the optimization process, the solution must still satisfy the invariants of the TSP.
- ii. Temperature: The SA algorithm uses a temperature parameter that controls the probability of accepting moves that increase the total length of the tour. The temperature decreases over time, which allows the algorithm to focus on improving the solution as it converges towards the optimal solution.
- Ant colony optimization: The ACO algorithm is then used to optimize the solution obtained from the Christofides algorithm by iteratively constructing and improving tours using a population of artificial ants. The invariants used in the ACO algorithm are:
 - i. Pheromone trail: The artificial ants deposit pheromone along the edges they traverse, which creates a trail that influences the behavior of subsequent ants. This allows the algorithm to explore promising areas of the search space and converge towards the optimal solution.
 - ii. Local search: The ACO algorithm can also include a local search heuristic, which allows ants to improve the tours by making small adjustments to the current solution. This can help the algorithm converge towards the optimal solution faster.
- iii. Parameter tuning: The ACO algorithm uses several parameters that need to be carefully tuned to ensure good performance. These parameters include the pheromone evaporation rate, the pheromone deposit amount, and the balance between exploration and exploitation.

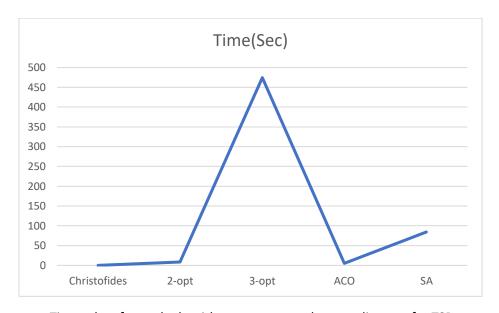
• Flow Chart:



• Observations & Graphical Analysis:

The time taken by three-opt to calculate shortest distance is higher than the other algorithms such as Christofides, two-opt, simulated annealing, ant colony optimization. This happens because worst case time complexity of three-opt is $O(n^3)$ compared to two-opt which has worst case time complexity of $O(n^2)$, so for large values of n (number of cities) three-opt takes more time but provides shortest tour distance compared to two-opt algorithm. When we tried giving the output of two-opt to the three-opt algorithm we were able to reduce the running time.

The 3-opt algorithm considers changes that involve removing and reconnecting three edges at a time. Specifically, the algorithm considers all possible combinations of three edges that can be removed from the tour and then reconnected in a different order. The resulting tour is then evaluated, and if it improves the total tour distance, it is accepted. By considering more possible changes to the tour, the 3-opt algorithm has a greater chance of finding a better solution than the 2-opt algorithm.



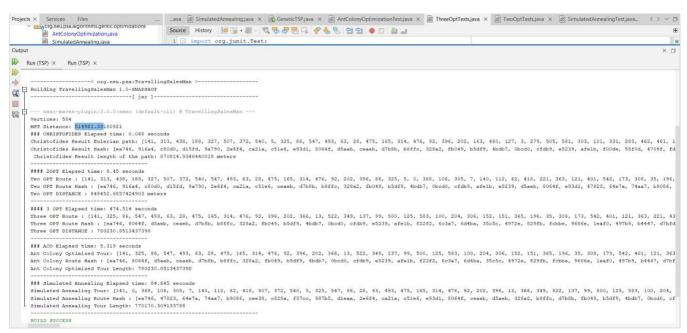
Time taken for each algorithm to generate shortest distance for TSP

The total tour distance calculated by each optimization algorithm is compared here. The Christofides algorithm can provide a better approximation than the nearest neighbor heuristic, and then the tactical optimization techniques of 2-opt and 3-opt can further improve the solution obtained from the Christofides algorithm. Similarly, the strategic optimization techniques of simulated annealing and ant colony algorithm can further refine the tour and potentially provide better solutions than the tactical optimization techniques.



Total Tour distance calculated by each algorithm to generate shortest distance for TSP

• Results & Mathematical Analysis:



To compute the MST of connected edge-weighted graph, the Prim's algorithm has a time complexity proportional to $E \log V$, where E is Edges and V is Vertices. In terms of space complexity, the Christofides algorithm requires O (E + V) space to store the graph and the minimum spanning tree.

The 2-opt algorithm has a worst-case time complexity of $O(n^2)$, where n is the number of cities, while the 3-opt algorithm has a worst-case time complexity of $O(n^3)$. The 2-opt and 3-opt algorithms require O(n) space to store the current solution.

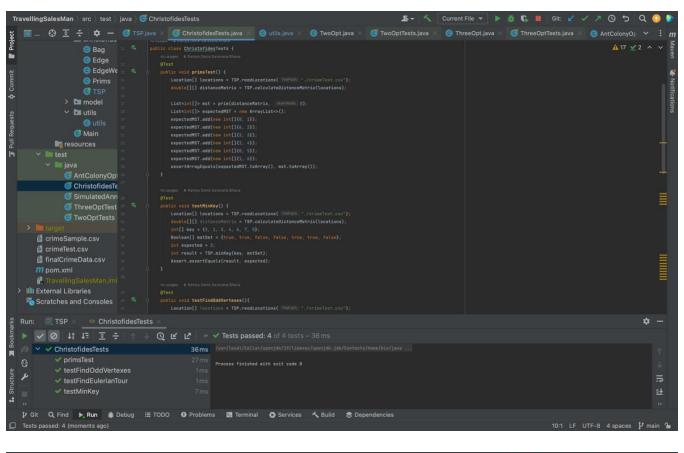
Simulated annealing is a metaheuristic algorithm where the time complexity depends on the number of iterations and the size of the neighborhood. The time complexity can be expressed as $O(kn^2)$, where k is the number of iterations and n is the number of cities. The space complexity of the algorithm is O(n), which is required for storing the tour and O(1) space to store the temperature.

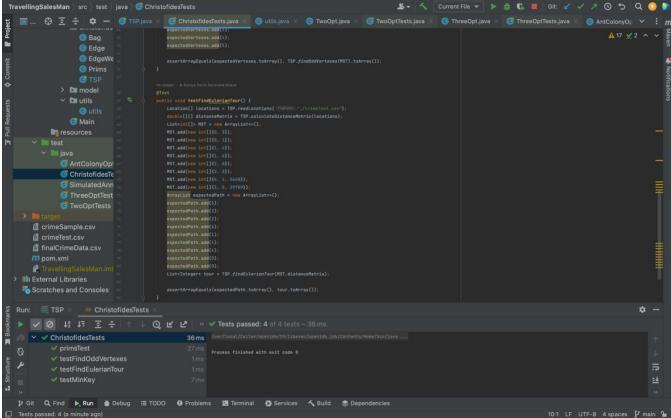
The worst-case time complexity of the ACO (ant colony optimization) algorithm for the TSP is $O(n^2m)$, where n is the number of cities and m is the number of iterations. The space complexity is also $O(n^2)$ since the pheromone matrix requires n^2 memory locations. The number of ants used in the algorithm can also impact its running time, as well as the termination criterion for the algorithm. Thus, the actual running time can be affected by several factors and may differ from the theoretical worst-case estimate.

• Unit tests:

ChristofidesTests:

- **primsTest:** Loads few samples (7 data points) from crimeTest file and validates the MST generated by prims algorithm that is to be used as a base to find the TPS.
- **testMinKey:** Loads few samples (7 data points) from crimeTest file and validates the minKey method that is used to find MST.
- **testFindEulerianTour**: Loads few samples (7 data points) from crimeTest file and asserts the eulerian tour generated and returned for TSP class.
- **testFindOddVertexes**: Loads few samples (7 data points) from crimeTest file and tests the odd vertices that is generated by the FindOddVertexes method.



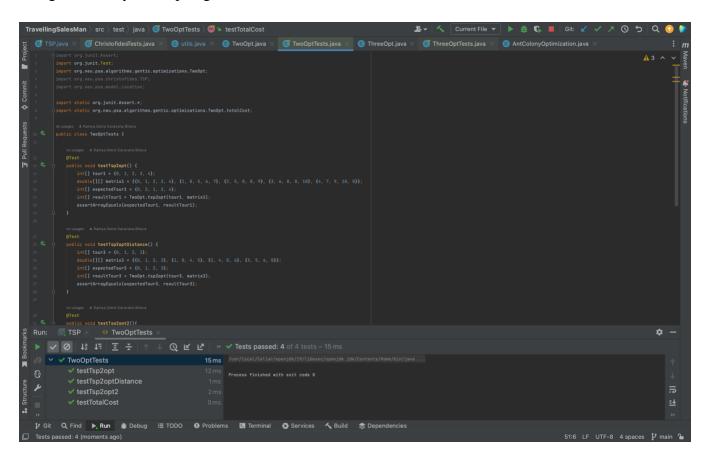


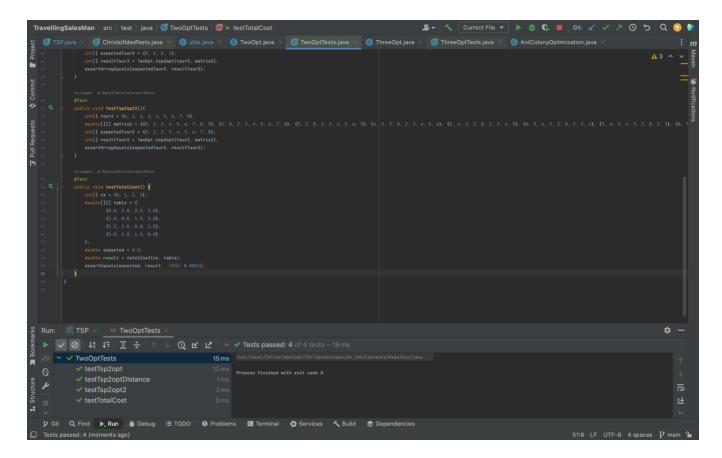
TwoOptTest:

testTsp2Opt: Initializes a small tour in form of an array along with adjacency matrix and asserts that tour generated by TwoOpt algorithm class.

testTotalCost: Initializes a small graph along with adjacency matrix and confirms the total cost i.e., the distance for this graph's TSP.

testTsp2Opt2: Initializes a larger tour in form of an array along with adjacency matrix and asserts that tour generated by TwoOpt algorithm class.



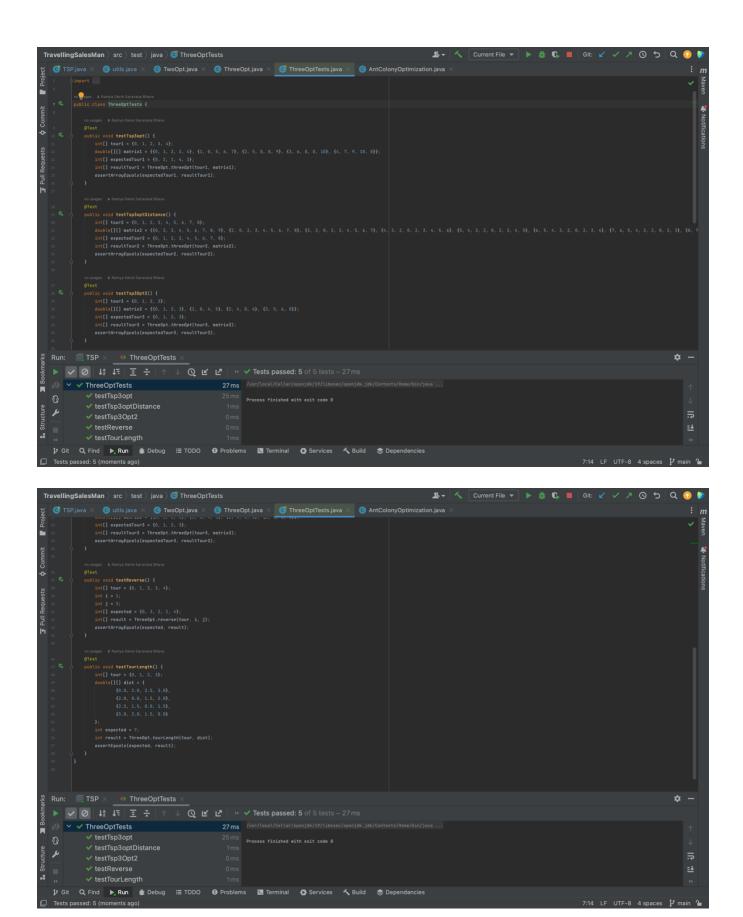


ThreeOptTests:

testTsp3Opt: Initializes a small tour in form of an array along with adjacency matrix and asserts that tour generated by ThreeOpt algorithm class.

testTotalCost: Initializes a small graph along with adjacency matrix and confirms the total cost i.e., the distance for this graph's TSP.

testTsp3Opt2: Initializes a larger tour in form of an array along with adjacency matrix and asserts that tour generated by ThreeOpt algorithm class.

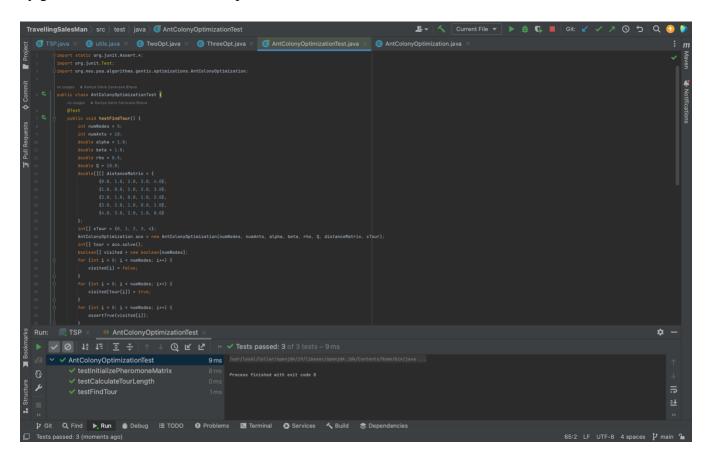


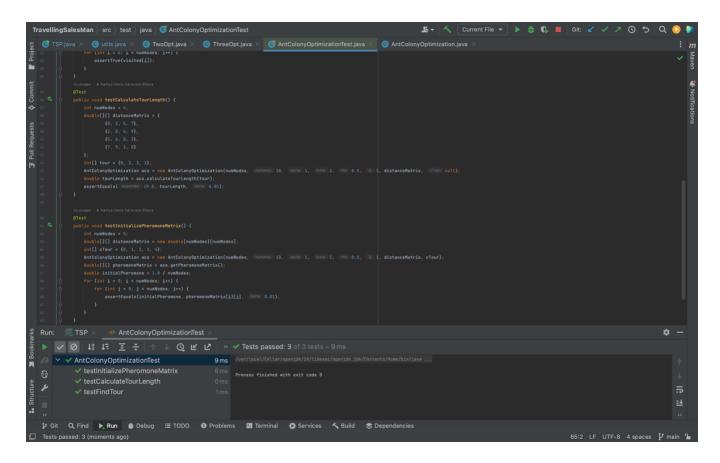
AntColonyOptimizationTest:

testFindTour: Initializes nodes and the adjacency matrix and validates tour generated and that all nodes are visited.

testCalculateTourLength: Initializes number of nodes in the tour along with the adjacency distance matrix and asserts the value of tour length.

testInitializePheromoneMatrix: Initializes number of nodes and verifies if the matrix values generated by getPheromoneMatrix are as expected.





SimulatedAnnealing:

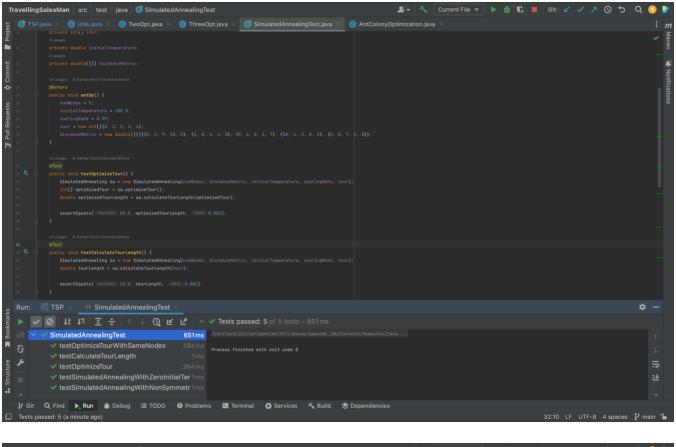
testOptimizeTour: Sets the nodes, distance matrix, along with the temperature and validates the optimal tour length generated by the algorithm is the same.

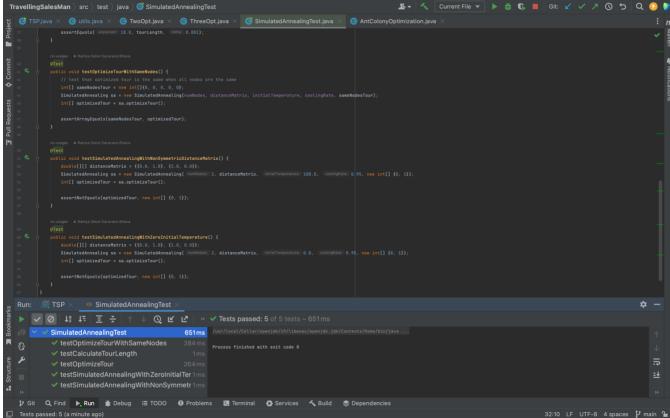
testOptimizedTourWithSameNodes: Sets the tour with just one node and asserts that the optimized tour is same when all nodes are same.

testCalculateTourLength: Sets the nodes, distance matrix, along with the temperature and validates the tour length generated by the algorithm is the same.

testSimulatedAnnealingWithZeroInitialTemperature: Sets the nodes, distance matrix, with initial temperate set to zero and validates the tour generated by the algorithm is as expected.

testSimulatedAnnealingWithNonSymmetricDistanceMatrix: Sets the nodes where the distance between nodes is not symmetric and validates the tour generated by the algorithm is as expected.





Conclusion:

In conclusion, we have implemented the TSP using the Christofides algorithm and optimized the output using various optimization techniques. We started by constructing the minimum spanning tree of the given graph using Prim's algorithm and then used the nearest neighbor heuristic to find an approximate solution for the TSP. Then, we applied the Christofides algorithm to the graph to obtain a better approximation by connecting the odd-degree vertices in the minimum spanning tree using the minimum-weight perfect matching algorithm.

Next, we applied the tactical optimization techniques of two-opt and three-opt to further improve the solution by removing local optima and refining the tour. These techniques were able to significantly reduce the total distance of the tour in a short amount of time. Three-opt algorithm takes more time than two-opt because it considers all possible combinations of three edges that can be removed from the tour and then reconnected in a different order. The resulting tour is then evaluated, and if it improves the total tour distance, it is accepted.

Finally, we implemented the strategic optimization techniques of simulated annealing and ant colony algorithm to refine the tour further and avoid getting trapped in local optima. Simulated annealing is a metaheuristic that uses random search to escape local optima and improve the overall solution, while the ant colony algorithm mimics the behavior of ants in finding the shortest path between their colony and food source by depositing pheromones.

Overall, these results suggest that the choice of algorithm plays a significant role in the solution to the traveling salesman problem. While some algorithms may produce better results in certain situations, others may perform better in different circumstances. Therefore, researchers and practitioners must carefully evaluate different algorithms and select the most appropriate one for their particular use case. Among all the optimization algorithms (two-opt, three-opt, ant colony, simulated annealing) used here, simulated annealing algorithm provides the shortest distance for the data set provided.

References:

- 1. https://www.youtube.com/watch?v=GiDsjIBOVoA&ab_channel=Reducible
- $2. \ \ \, \underline{https://bank.engzenon.com/tmp/5e7f6ee5-d4dc-4aa8-9b0a-42d3c0feb99b/6062caf3-c600-4fc2-b413-4ab8c0feb99b/Algorithms-4th-Edition.pdf}$
- 3. https://en.wikipedia.org/wiki/Christofides_algorithm
- 4. https://en.wikipedia.org/wiki/Ant_colony_optimization_algorithms
- 5. https://en.wikipedia.org/wiki/Simulated annealing
- 6. https://en.wikipedia.org/wiki/2-opt
- 7. https://en.wikipedia.org/wiki/3-opt
- 8. https://ieeexplore.ieee.org/document/5200345
- 9. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=995530
- $10.\ \underline{https://www.researchgate.net/profile/Mauro-}$
 - <u>Birattari/publication/220380876 On the Invariance of Ant Colony Optimization/links/547b4b9b0cf293e2da2d7098/On-the-Invariance-of-Ant-Colony-Optimization.pdf</u>