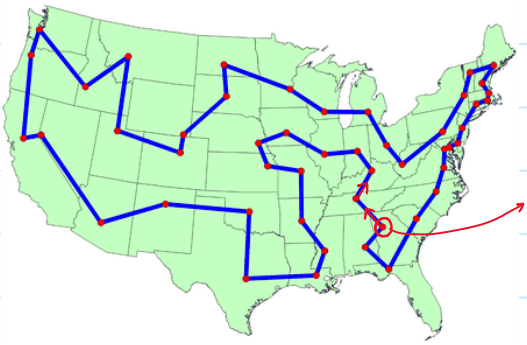
Traveling Salesman Problem (TSP)

Sunday, May 13, 2018 17:38

For the glory of God

What is Traveling Salesman Problem ?

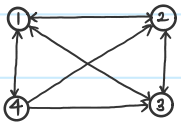
- It can be described in term of a salesman who must travel to a number of cities during one tour.
- The salesman has to visit each city exactly once before returning to his home city.
- The objective is to determine which route will minimize the total distance that the salesman must travel.



let's say Atlanta (city 1) is the salesman's home city.

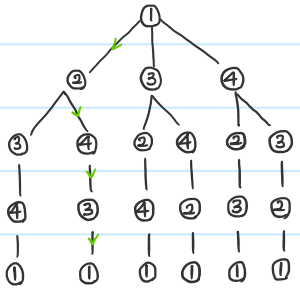
Characteristics of Traveling Salesman Problem

- Let's think about a simple case ; say we have only 4 cities including the salesman's home city.



; 1 = Atlanta (Home city)

- Let's assume that the distance (or cost) between the cities is given.
- Then, we may be able to think about all possible routes as following ;



let's say 1-2-4-3-1 is the shortest distance.

→ Returning to home city

- Finally, we'll be able to choose one of the routes, which is the optimal solution.
- However, the difficulty of traveling salesman problems increases rapidly as the number of cities increases.
  - For a problem with n cities, the number of feasible routes to be considered is (n-1)!
- Because of the enormous difficulty of solving large traveling salesman problems, Heuristic methods continue to be

a popular way of addressing such problems.

What is Metaheuristic and heuristic method ?

- Classical optimization methods (e.g. Linear programming) do not always work.
  - For instance, some problems are so complicated that it may not be possible to find the optimal solution.
    ↳ If it is possible to have a mathematical model for all objective functions and constraints, the method shouldn't be matter. However, it's really hard to get them happen in reality.
- In such situation, Heuristic methods are commonly used to find a good feasible solution.
    ↳ At least reasonably close to being optimal
- A heuristic method can be defined as an empirical search (e.g. based on experience, observation, and so forth) that is likely to discover a very good feasible solution but not necessarily an optimum solution.
  - No guarantee can be given about the quality of the solution obtained ; but a well-designed heuristic method usually can provide a solution that is at least nearly optimal.
  - This method is designed to fit a specific problem type rather than a variety of applications.
    (In other words, this is very specific and problem-dependent)
- A metaheuristic method was introduced to fit heuristic methods into a general way.
  - This is a general solution method that provides both a general structure and strategy guidelines.
  - One of key features of metaheuristic method is its ability to escape from a local minimum.
  - Genetic algorithm has become one of the popular techniques in metaheuristic methods.
    ↳ For more information, refer to Optimization hand-written note.

Genetic Algorithm for Traveling Salesman Problem

- Genetic Algorithm (GA) is inspired by an analogy with natural selection in Darwin's theory of evolution.
- The invention of Genetic Algorithm was credited to John Holland (professor at University of Michigan)
- Steps in Genetic algorithm are as following ;
  Initialization, Selection, Reproduction (crossover and mutation), Replacement
- Let's apply this algorithm to Traveling Salesman problem.

a) Initialization

- In this case, let us assume that we have five cities.

City 2

City 1

City 3

Random Number	Sequence
0.23	2
0.65	3
0.49	4
0.58	5
0.75	6
0.34	7

⇒

Random Number	Sequence
0.23	2
0.34	7
0.49	4
0.58	5
0.65	3
0.75	6

Sort

City 2 City 3 City 1 City 5 City 4

0.23	2
0.65	3
0.49	4
0.58	5
0.75	6
0.34	7

⇒

0.23	2
0.34	7
0.49	4
0.58	5
0.65	3
0.75	6

sort

one of ways is to generate a double array and sort the array.

- With given cities, randomly generate a population of possible routes

Route 1	City 3	City 1	City 5	City 2	City 4
Route 2	City 1	City 2	City 4	City 3	City 5
Route 3	City 2	City 5	City 4	City 1	City 3

let's say this represents Route 1

The first generation

b) Selection

- Calculate the fitness of each route and select a few good solutions to be survived. (Generally, 20%)
- In TSP, the fitness is related to the shortest distance. (Tournament selection would be a good option)
- Let's say we selected two solutions, namely Route 1 and Route 2.

Route 1	City 3	City 1	City 5	City 2	City 4
Route 2	City 1	City 2	City 4	City 3	City 5
Route 3	City 2	City 5	City 4	City 1	City 3

✓ ✓ These are going to be used for reproduction.

The first generation

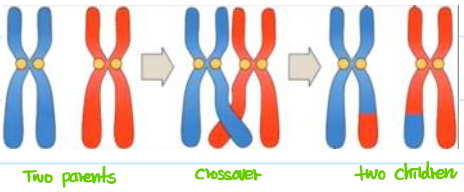
c) Reproduction (Crossover and mutation)

- Crossover is the genetic recombination of two parents.
  - ↳ After selecting two parents, we will consider the percentage determining if crossover occurs.
    - Rule of thumb is 70%; if no crossover occurs, both parents simply become children.
- There are one-point, two-point, and uniform crossover.
- In this case, let's think about the general case for the reproduction.

R1	C3	C1	C5	C2	C4

R2	C1	C2	C4	C3	C5

Child 1	City 4	City 1	City 5	City 3	City 2



Two parents Crossover two children

- Mutation is a process that is randomly flipping a few of the bits in the children. (It is based on a specified probability, e.g. 10%)
  - ↳ It is needed to maintain genetic diversity but it may be too randomly if it is too much.

Child 1	City 4	City 1	City 5	City 3	City 2

↳ City 1 is chosen for the mutation

Child 1	City 4	City 2	City 5	City 3	City 1

↳ City 1 was swapped with city 2 in the route.

; Let's say here, each city has a 5% chance of being swapped with another city.

d) Replacement

- There are two ways to replace them for the next generation ;

  ↳ Kill all parents and only consider childrens as new parents

  Evaluate all parents and childrens, then select the best populations

- Let's say that we selected the children 1 and got rid of parent 1 for the route 1.

Child 1	City 4	City 2	City 5	City 3	City 1

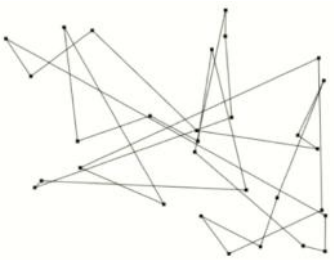
The second generation

↳ As the process is repeated, the algorithm will generate fitter and fitter generations.

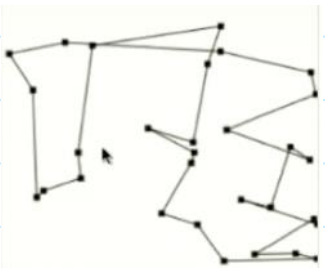
Characteristics of Genetic Algorithm for solving Traveling Salesman Problem

· The feasible solution would be affected by population size, crossover chance, and mutation chance.

· It will not always find the best solution but it can do a decent job and very quickly.



< Initial route >



< Feasible solution >

→ Here, 30 random cities problem

: The feasible solution is definitely not the best solution but it's probably quite close to what the optimum would be.

· There is a key trade off between Exploitation and Exploration.

↳ Let's a little bit more talk about this.

a) Exploration

- It is the process of visiting entirely new regions of a search space.
- It means that the algorithm will attemp to achieve a good coverage of the search space.
- It will not eventually find just any local optimum but ideally be close to the global optimum.
- It is searching the space of possible solutions well.

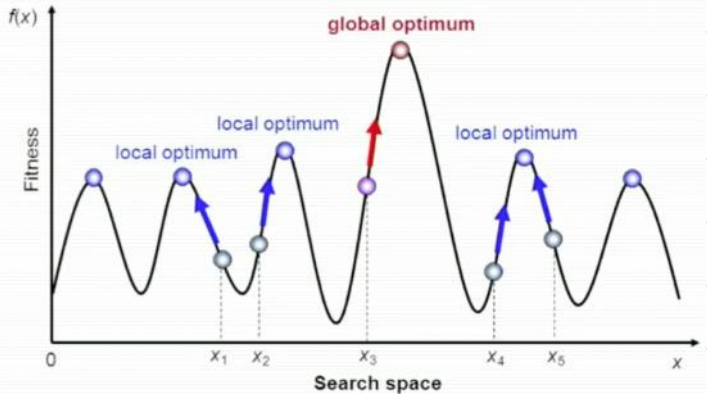
b) Exploitation

- It means that given a reasonable solution, the algorithm will keep refining the solution until it reaches a local optimum.
- It is the process of visiting those regions of a search space within the neighborhood of previously visited points.
- It means using already exist solutions and make refinement to it, so that its fitness will improve.
- It takes advantage of existing best solutions in the population.

c) Note that

- In order to be successful, a search algorithm needs to establish a good ratio between exploration and exploitation.
- For Genetic Algorithm, exploration is typically achieved by crossover and mutation whilst selection is used to promote exploitation.
- we could say that { Larger mutation rate : a kind of exploration
  Smaller mutation rate : a sort of exploitation
- Someone would argue that exploration is the genetic operator while exploitation is the selection process itself.

- Anyhow,



- So earlier, we might be stuck on the solution by hitting local optimum in the global solution space.
- The hope is that by introducing Genetic algorithm we won't be stuck in the local optimum.
- We may be even climbed down or jump across in our solution space.

Python code programming for Genetic Algorithm to solve Traveling Salesman Problem