National University of Computer & Emerging Sciences, Karachi

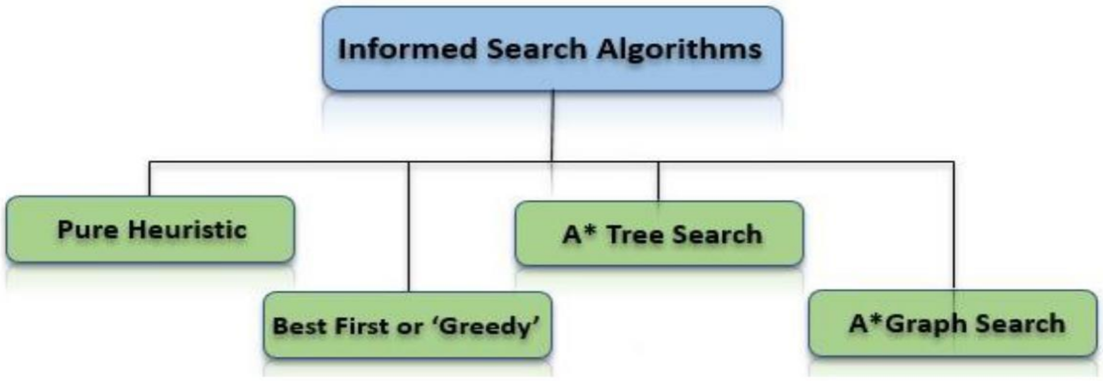
Computer Science Department
Spring 2023, Lab Manual – 05

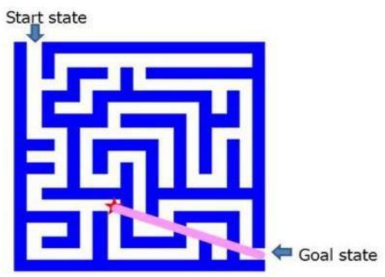
Course Code: AI-2002	Course: Artificial Intelligence Lab
Instructor(s):	**Kariz Kamal, Saeeda Kanwal, Sania Urooj, Shakir Hussain, Omer Qureshi, Muhammad Ali Fatmi.**

Contents:

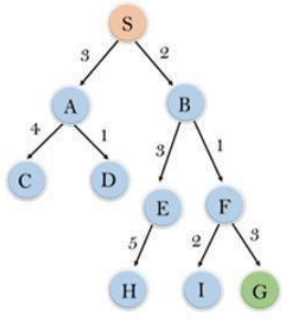
Heuristic (or informed) search algorithms:

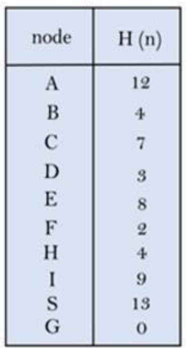
- A solution cost estimation is used to guide the search.
- The optimal solution, or even a solution, are not guaranteed.
- Some information about problem space (heuristic) is used to compute preference among the children for exploration and expansion.
- To solve large problems with large number of possible states, problem-specific knowledge needs to be added to increase efficiency of algorithm.





It expands nodes in the order of their heuristic values. It creates two lists, a closed list for the already expanded nodes and an open list for the created but unexpanded nodes.
In each iteration, a node with a minimum heuristic value is expanded, all its child nodes are created and placed in the closed list. Then, the heuristic function is applied to the child nodes and they are placed in the open list according to their heuristic value. The shorter paths are saved and the longer ones are disposed.





node	H (n)
A	12
B	4
C	7
D	3
E	8
F	2
H	4
I	9
S	13
G	0

Best-First Search

If we consider searching as a form of traversal in a graph, an uninformed search algorithm would blindly traverse to the next node in a given manner without considering the cost associated with that step. An informed search, like Best first search, on the other hand would use an evaluation function to decide which among the various available nodes is the most promising (or 'BEST') before traversing to that node.
The Best first search uses the concept of a Priority queue and heuristic search. To search the graph space, the BFS method uses two lists for tracking the traversal. An 'Open' list which keeps track of the current 'immediate' nodes available for traversal and 'CLOSED' list that keeps track of the nodes already traversed.

Variants of Best First Search

- Greedy best-first search
- A* best-first search

Best first search algorithm:

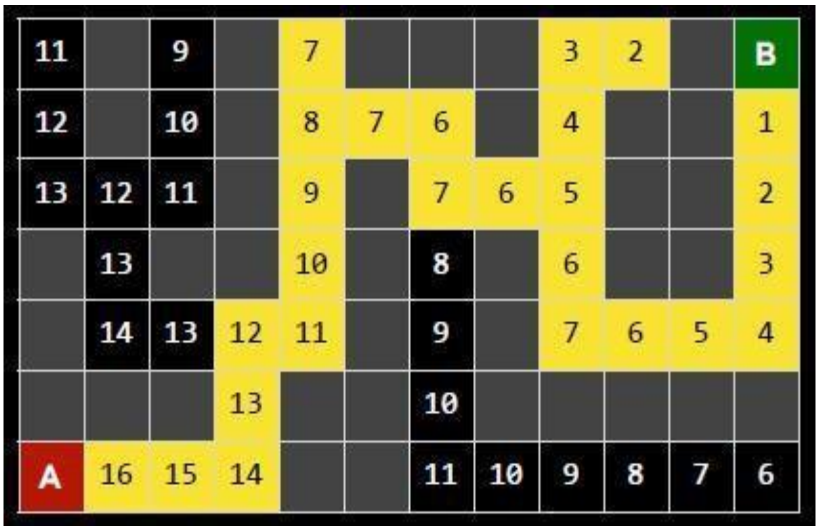
1. Create 2 empty lists: OPEN and CLOSED
2. Start from the initial node (say N) and put it in the 'ordered' OPEN list
3. Repeat the next steps until GOAL node is reached

1. If OPEN list is empty, then EXIT the loop returning 'False'
2. Select the first/top node (say N) in the OPEN list and move it to the CLOSED list. Also capture the information of the parent node
3. If N is a GOAL node, then move the node to the Closed list and exit the loop returning 'True'. The solution can be found by backtracking the path
4. If N is not the GOAL node, expand node N to generate the 'immediate' next nodes linked to node N and add all those to the OPEN list

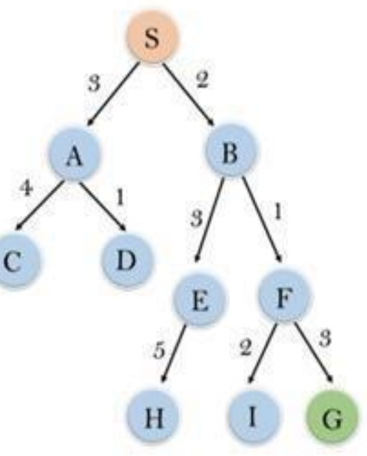
5. Reorder the nodes in the OPEN list in ascending order according to an evaluation function f(n)

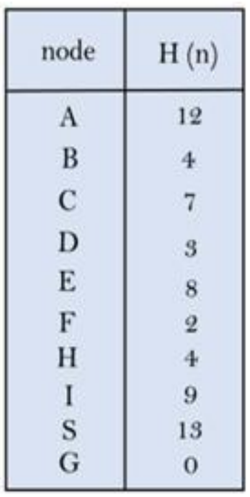
Greedy Best first search algorithm:

A search method of selecting the best local choice at each step in hopes of finding an optimal solution.
It is the combination of depth-first search and breadth-first search algorithms.
At each step, we choose the most promising node. In the greedy search algorithm, we expand the node which is closest to the goal node and the closest cost is estimated by heuristic function, i.e. f(n)= h(n).

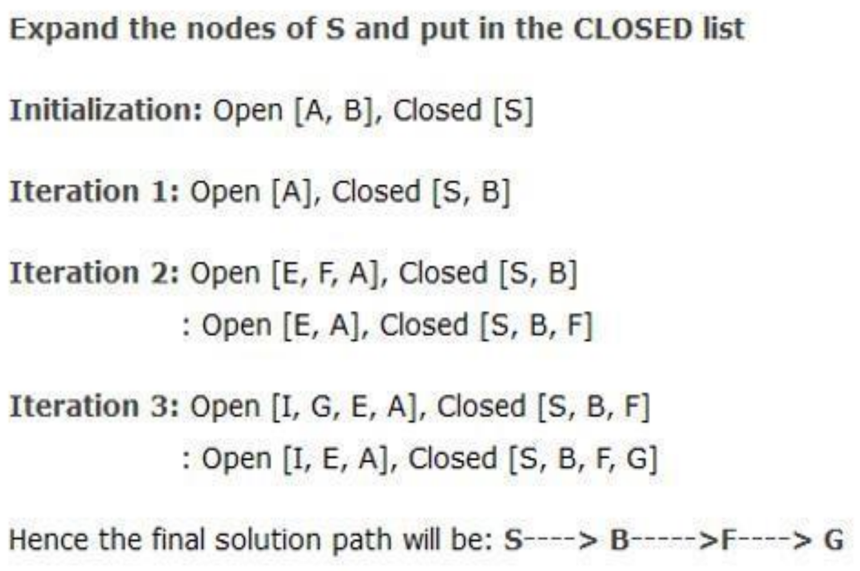


- Evaluation function *f(n) = h(n)* h(n) = estimated cost of the cheapest path from the state at node n to a goal state Greedy best-first search expands the node that appears to be closest to goal It is implemented using priority queue.





node	H (n)
A	12
B	4
C	7
D	3
E	8
F	2
H	4
I	9
S	13
G	0



Expand the nodes of S and put in the CLOSED list

Initialization: Open [A, B], Closed [S]

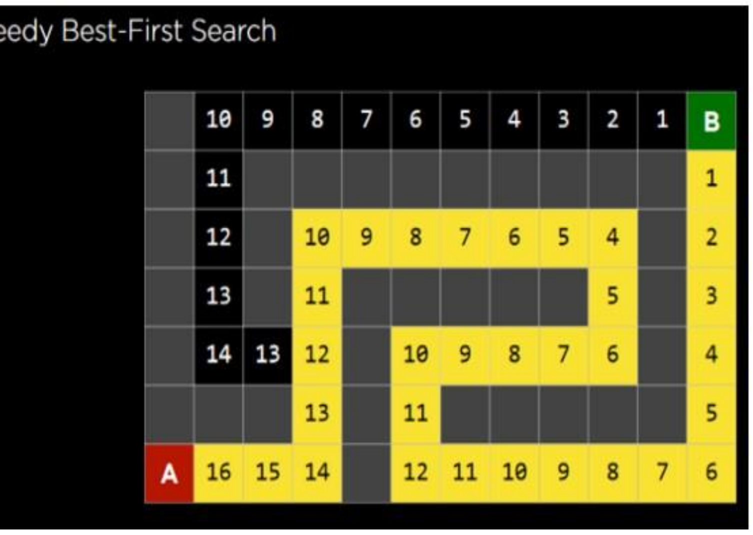
Iteration 1: Open [A], Closed [S, B]

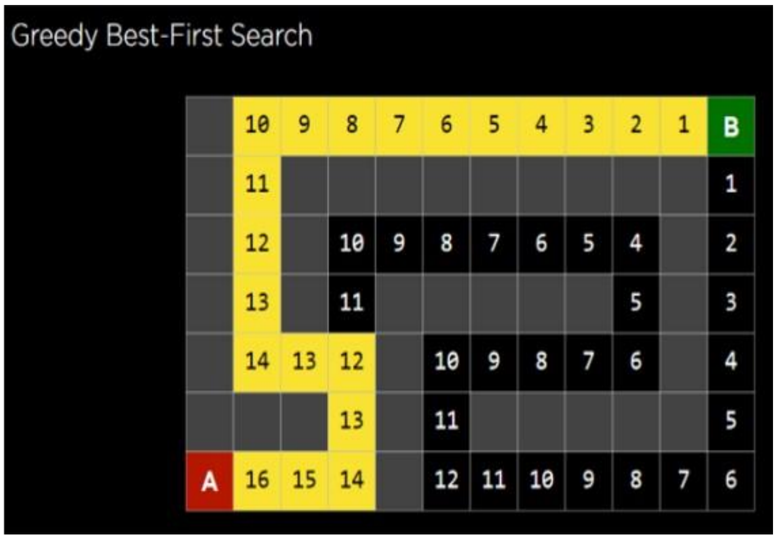
Iteration 2: Open [E, F, A], Closed [S, B]
: Open [E, A], Closed [S, B, F]

Iteration 3: Open [I, G, E, A], Closed [S, B, F]
: Open [I, E, A], Closed [S, B, F, G]

Hence the final solution path will be: **S----> B----->F----> G**

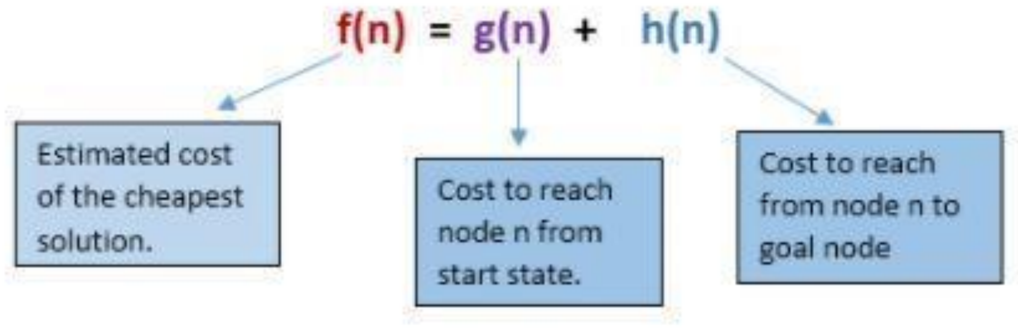
Disadvantage − It can get stuck in loops. It is not optimal.





A* search

It is best-known form of Best First search. It avoids expanding paths that are already expensive, but expands most promising paths first.

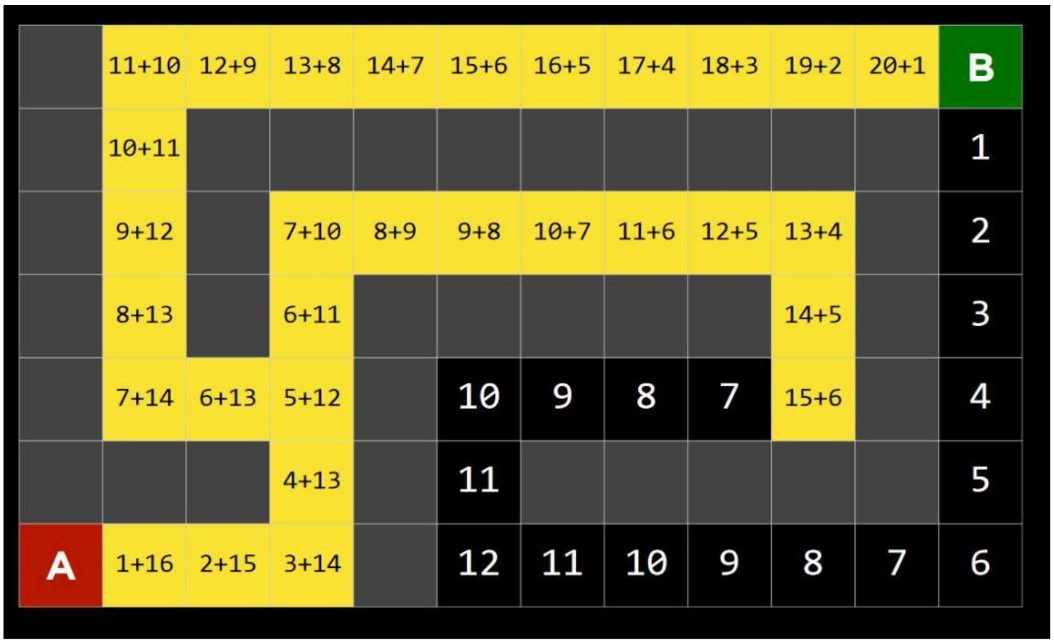

f(n) = g(n) + h(n)
Estimated cost of the cheapest solution.
Cost to reach node n from start state.
Cost to reach from node n to goal node

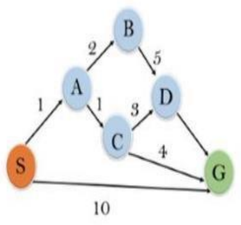
•Evaluation function f(n) = g(n) + h(n)

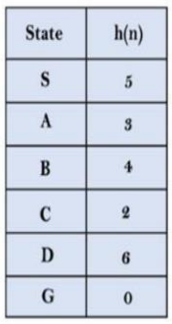
•g(n) = cost so far to reach n

•h(n)= estimated cost from n to goal

•f(n) = estimated total cost of path through n to goal







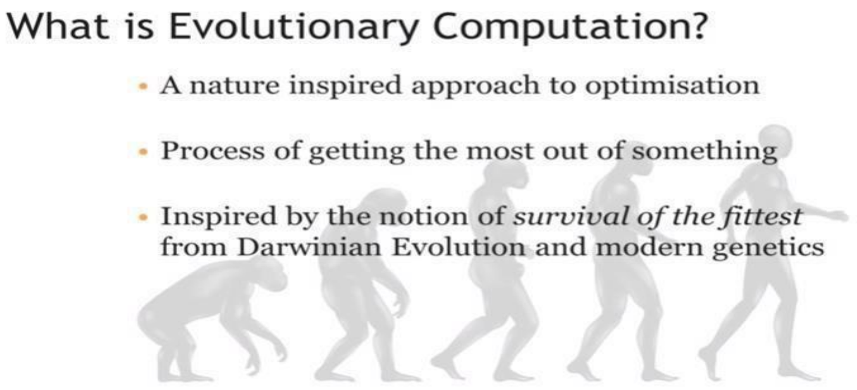
State	h(n)
S	5
A	3
B	4
C	2
D	6
G	0

Iteration1: {(S --> A, 4), (S-->G, 10)}

Iteration2: {(S - > A- >C, 4), (S--> A-- >B, 7), (S-- >G, 10)}

Iteration3: {(S --> A-- >C--- >G, 6), (S-- > A-->C--- >D, 11), (S--> A-->B, 7), (S-- >G, 10)} Iteration 4 will give the final result, as S--->A--->C--->G, it provides the optimal path with cost 6

Evolutionary algorithm:



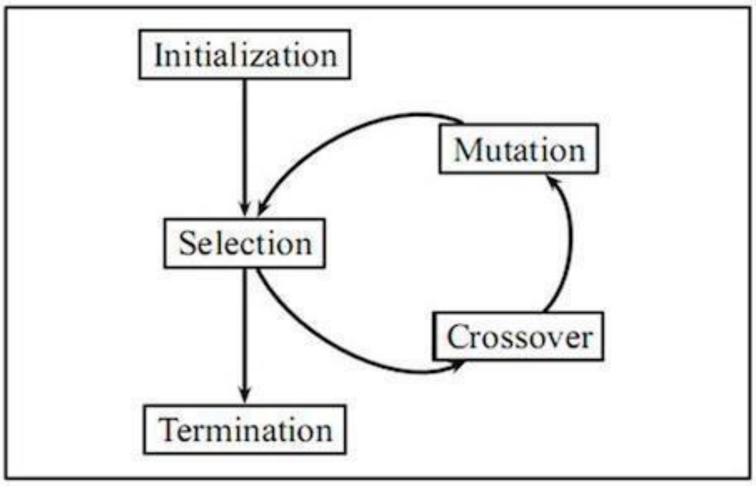
What is Evolutionary Computation?

- A nature inspired approach to optimisation
- Process of getting the most out of something
- Inspired by the notion of *survival of the fittest* from Darwinian Evolution and modern genetics

Evolutionary algorithms are a heuristic-based approach to solving problems that cannot be easily solved in polynomial time, such as classically NP-Hard problems, and anything else that would take far too long to exhaustively process.

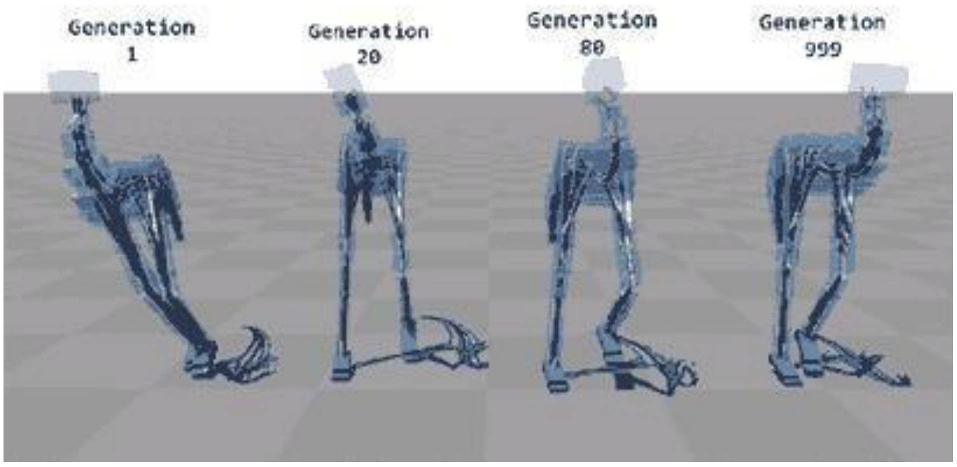
An EA contains four overall steps: initialization, selection, genetic operators, and termination. These steps each correspond, roughly, to a particular facet of natural selection, and provide easy ways to modularize implementations of this algorithm category.

Simply put, in an EA, fitter members will survive and proliferate, while unfit members will die off and not contribute to the gene pool of further generations, much like in natural selection.



Example:

Now, just to illustrate the result of this process I will show an example of an EA in action. The following figure shows several generations of dinosaurs learning to walk by optimizing their body structure and applied muscular forces. From left to right the generation increases, so the further right, the more optimized the walking process is. Despite the fact that the early generation dinosaurs were unable to walk, the EA was able to evolve the dinosaurs over time through mutation and crossover into a form that was able to walk.



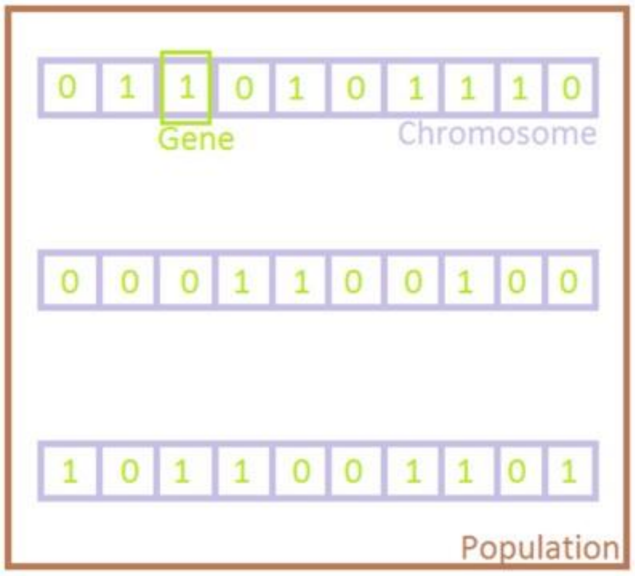
Genetic Algorithm:

A Genetic Algorithm (GA) is a meta-heuristic inspired by natural selection and is a part of the class of Evolutionary Algorithms (EA). We use these to generate high-quality solutions for optimization and search problems, for which, these use bio-inspired operators like mutation, crossover, and selection. In other words, using these, we hope to achieve optimal or near-optimal solutions to difficult problems.

This algorithm work in four steps:

- Individuals in population compete for resources, mate.
- Fittest individuals mate to create more off-springs than others.
- Fittest parent propagates genes through generation; parents may produce off-springs better than either parent.
- Each successive generation evolves to suit its ambience.

In optimization, we try to find within this search space the point or set of points that gives us the optimal solution. Each individual is like a string of characters/integers/floats and the strings are like chromosomes.



Phases in Genetic Algorithms:

Five phases are considered in a genetic algorithm.

- Initial population
- Fitness function
- Selection
- Crossover
- Mutation

Example Of GA:

1. **Initialization & Fitness**

We toss a fair coin 10 times and get the following initial population:

s1=1111010101 (s1) = 7   s2=0111000101

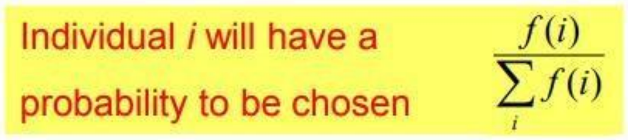
(s2) = 5   s3=1110110101 (s3) = 7

s4=0100010011 (s4) = 4   s5=1110111101

(s5) = 8   s6=0100110000 (s6) = 3

2. **Selection:**

We randomly (using a biased coin) select a subset of the individuals based on their fitness.



Individual i will have a probability to be chosen $\frac{f(i)}{\sum_i f(i)}$

Suppose that, after performing selection, we get the following population:

s1 ` = 1111010101 (s1)

s2 ` = 1110110101 (s3)

s3 ` = 1110111101 (s5)

s4 ` = 0111000101 (s2)

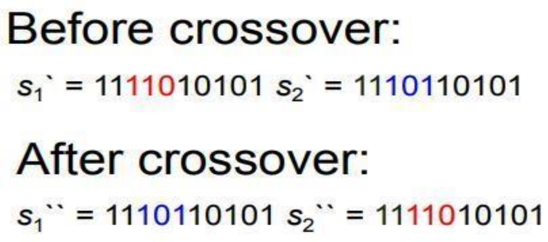
s5 ` = 0100010011 (s4)

s6 ` = 0100110000 (s6)

You can analyze here the fi values 7,7,8 for respective populations have the highest and nearest fitness function calculations we select the starting two pairs S1' and S2' for crossover process according to their fitness calculation.

3. **Crossover:**

Next we mate strings for crossover. For each couple we first decide whether to actually perform the crossover or not. If we decide to actually perform crossover, we randomly extract the crossover points, for instance from point 2 and 5.



Before crossover:

s_1` = 1111010101 $s_2$` = 1110110101

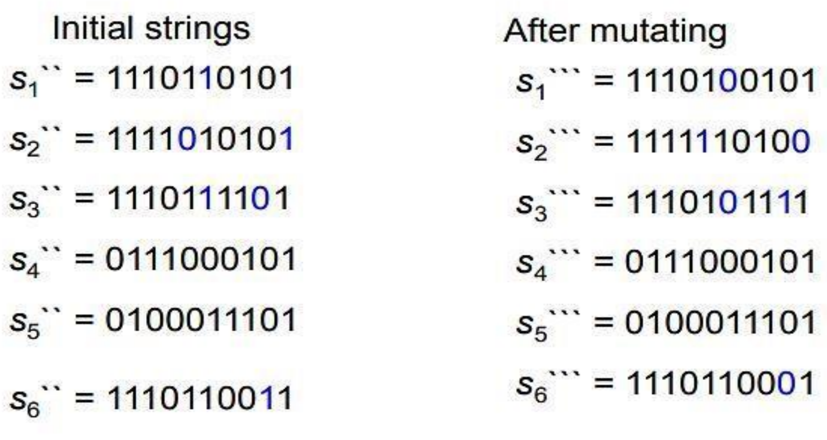
After crossover:

s_1`` = 1110110101 $s_2$`` = 1111010101

4. **Mutation:**

The final step is to apply random mutations: for each bit that we are to copy to the new population we allow a small probability of error (for instance 0.1).

Here, we also perform the crossover between the remaining pairs at a certain instance points.



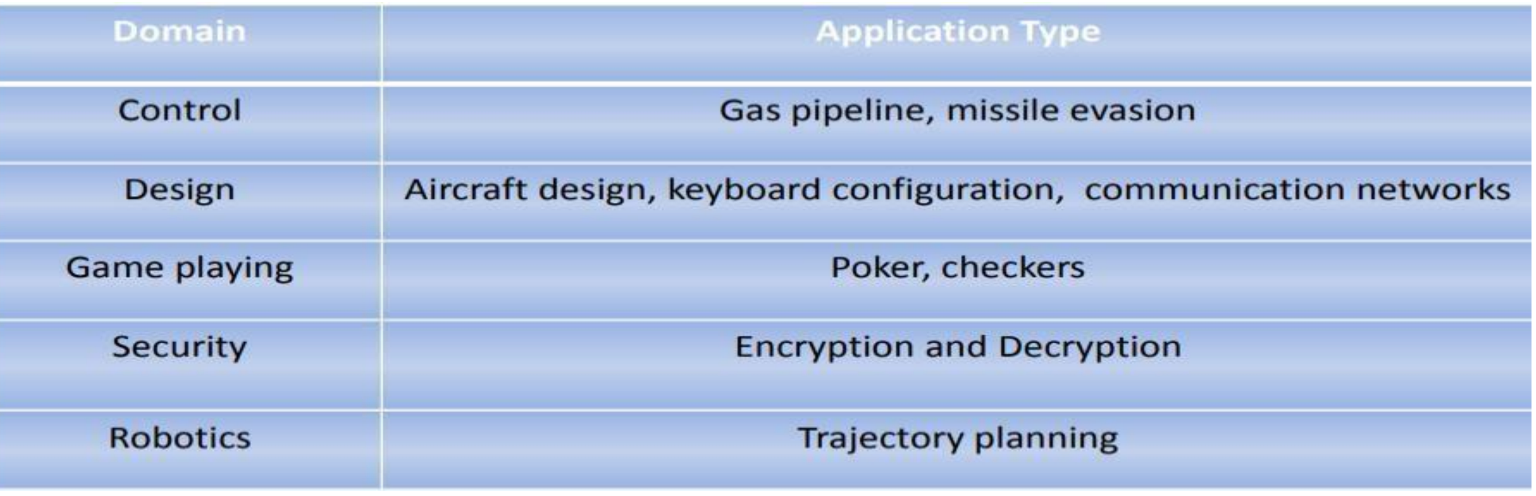
Initial strings	After mutating
s_1`` = 1110110101	s_1``` = 1110100101
s_2`` = 1111010101	s_2``` = 1111110100
s_3`` = 1110111101	s_3``` = 1110101111
s_4`` = 0111000101	s_4``` = 0111000101
s_5`` = 0100011101	s_5``` = 0100011101
s_6`` = 1110110011	s_6``` = 1110110001

In one generation, the total population fitness changed from 34 to 37, thus improved by ~9%. At this point, we go through the same process all over again, until a stopping criterion is met.

Benefits of Genetic Algorithms:

- Concept is easy to understand.
- Modular, separate from application.
- Supports multi-objective optimization.
- Always an answer; answer gets better with time.
- Easy to exploit previous or alternate solutions.
- Flexible building blocks for hybrid applications.

GA Applications:

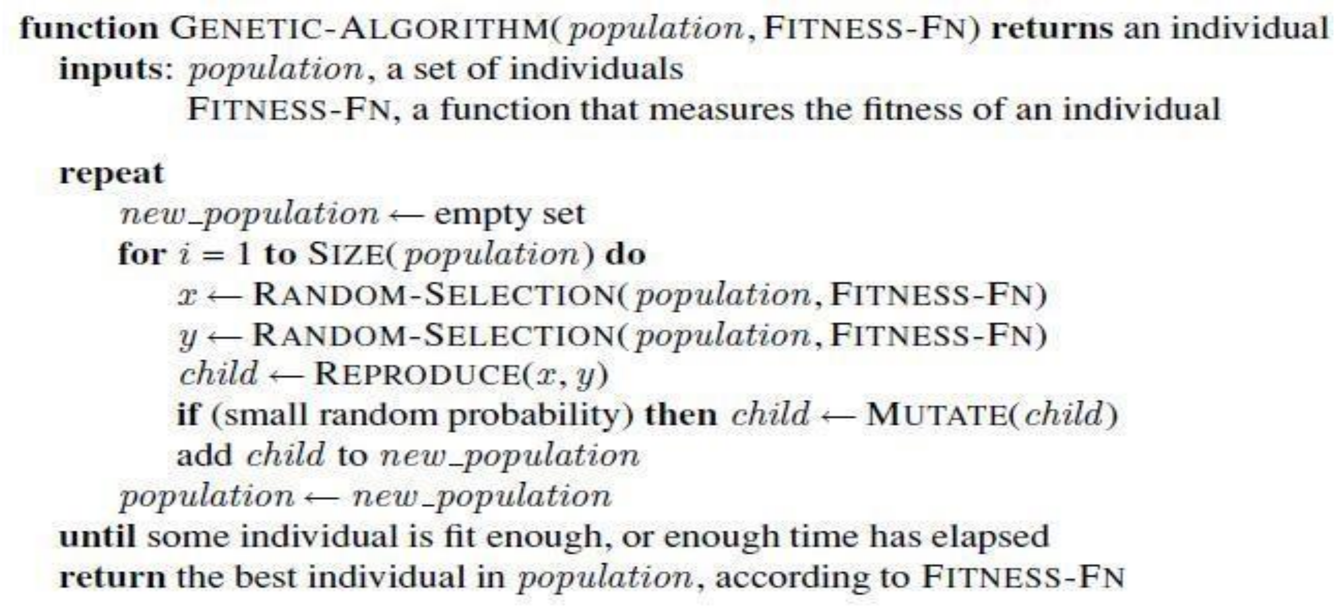


Domain	Application Type
Control	Gas pipeline, missile evasion
Design	Aircraft design, keyboard configuration, communication networks
Game playing	Poker, checkers
Security	Encryption and Decryption
Robotics	Trajectory planning

Limitations of Genetic Algorithms

- Not suitable for simple problems with available derivative information.
- Stochastic; no guarantee of the result solution being optimal.
- Frequent calculation of fitness value is computationally expensive for some problems. No guarantee of convergence to the optimal solution if not implemented properly.

Pseudo Code of Genetic Algorithm:



```
function GENETIC-ALGORITHM(population, FITNESS-FN) returns an individual
  inputs: population, a set of individuals
          FITNESS-FN, a function that measures the fitness of an individual

  repeat
     new_population ← empty set
     for i = 1 to SIZE(population) do
        x ← RANDOM-SELECTION(population, FITNESS-FN)
        y ← RANDOM-SELECTION(population, FITNESS-FN)
        child ← REPRODUCE(x, y)
        if (small random probability) then child ← MUTATE(child)
        add child to new_population
     population ← new_population
  until some individual is fit enough, or enough time has elapsed
  return the best individual in population, according to FITNESS-FN
```

<u>Lab Tasks</u>

Task#1:

Given a target string, the goal is to produce target string starting from a random string of the same length. In the following implementation, following analogies are made:

- Characters A-Z, a-z, 0-9 and other special symbols are considered as genes
- A string generated by these characters is considered as chromosome/solution/Individual
- Population size= 70
- Target string to be generated: TARGET = "Artificial Intelligence Lab"
- Fitness score is the number of characters which differ from characters in target string at a particular index. So, individual having lower fitness value is given more preference.

Task#2:

Implement genetic algorithm to the Traveling Salesman Problem.

We have a set of four cities A, B, C, and D. The distances between the cities are also given to us. Here (4-1)! That is 3! Route can be generated. The tour with A □B □C □D □A will be the optimal route for given problem.

Initial Population (set of solutions) = 6

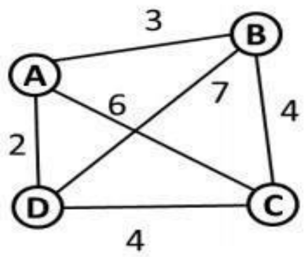


Fig. 1: Traveling salesman problem

Fitness (Quality of solution) = each solution is generally represented as a string of binary numbers, known as a chromosome. The most common fitness function for TSP is the length of the route. However, the 'shorter' the route is - the better.

Crossover point, exchange data after ‘1’ instances of the list.

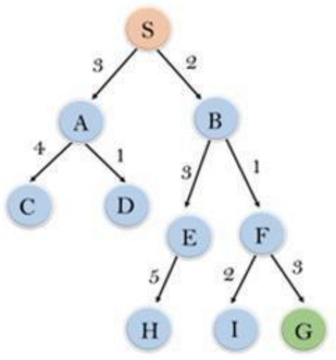
Mutation point perform between 2nd and 4th item.

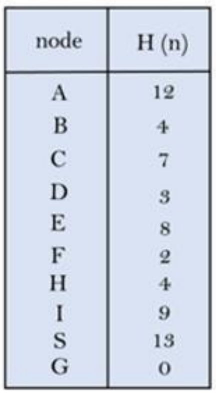
The pseudo code for genetic algorithm for implementing the Traveling salesman problem:

```
GeneticAlgorithm ( )
{
Initialize population of routes of cities randomly with a function Random ( )
Evaluate the fitness of each individual route using function Fitness ( )
While the fitness criteria is not satisfied do
{
        Selection of two routes for reproduction using select function
            (Select (parent_route1, parent_route2))
           Perform crossover on the selected parent routes with crossover function
            (child_route = Crossover (parent_route1, parent_route2))
        Perform mutation on the newly generated child routes with mutation function
            (Mutation (child_route) )
        Evaluate the fitness of child_route and replace the parent population with child_route
} }
```

Task#3:

Implement the following tree using greedy algorithm having a destination node H.

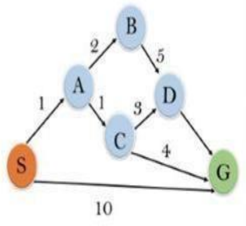


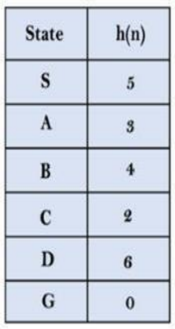


node	H (n)
A	12
B	4
C	7
D	3
E	8
F	2
H	4
I	9
S	13
G	0

Task#4:

Implement the following graph using A* search algorithm starting from Node S to Node D.

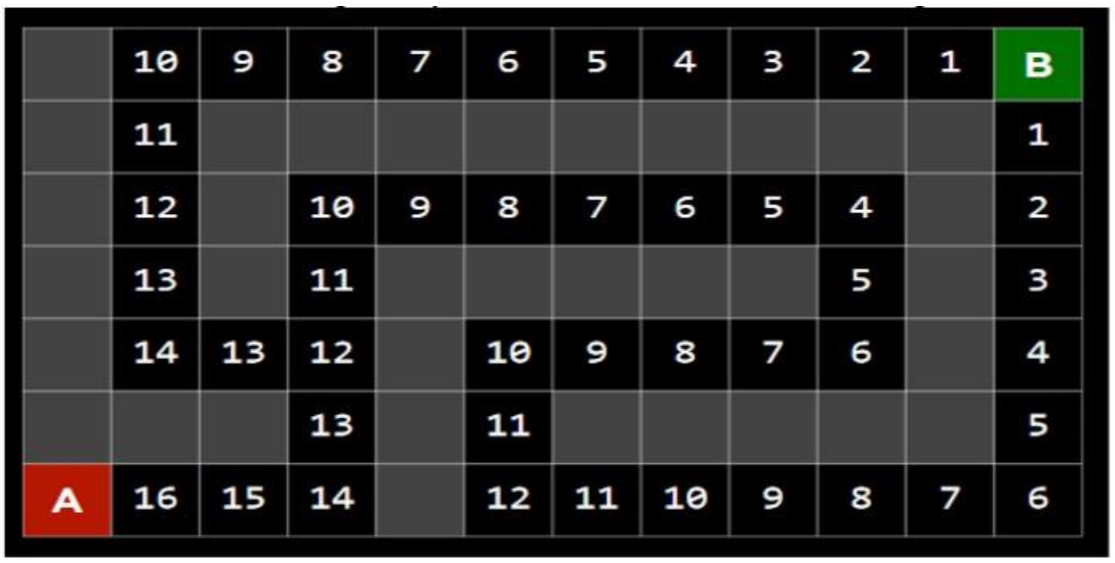




State	h(n)
S	5
A	3
B	4
C	2
D	6
G	0

Task#5:

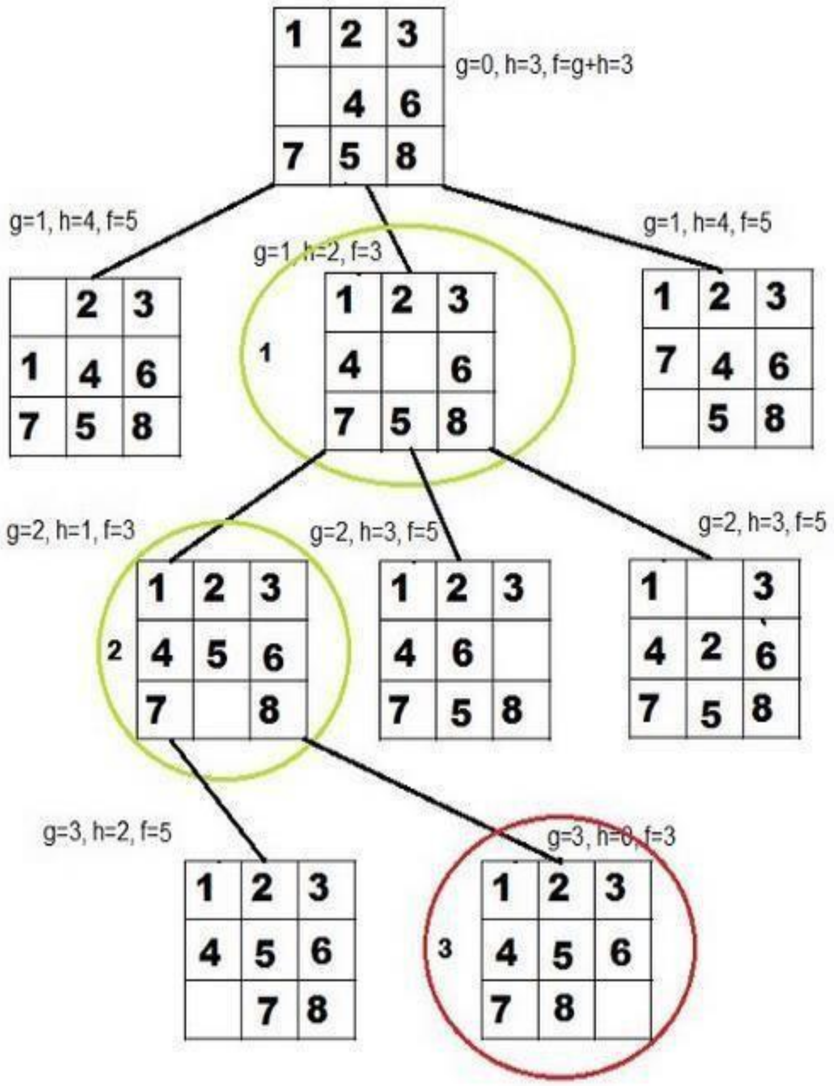
Solve the below maze using Greedy Best First and A* Best First search algorithm.



	10	9	8	7	6	5	4	3	2	1	B
	11										1
	12		10	9	8	7	6	5	4		2
	13		11						5		3
	14	13	12		10	9	8	7	6		4
			13		11						5
A	16	15	14		12	11	10	9	8	7	6

Task#6:

Write a program to solve the 8-puzzle problem using Heuristics (h(n)) for A*



g=0, h=3, f=g+h=3
g=1, h=4, f=5
g=1, h=2, f=3
g=1, h=4, f=5
1
g=2, h=1, f=3
g=2, h=3, f=5
g=2, h=3, f=5
2
g=3, h=2, f=5
g=3, h=0, f=3
3