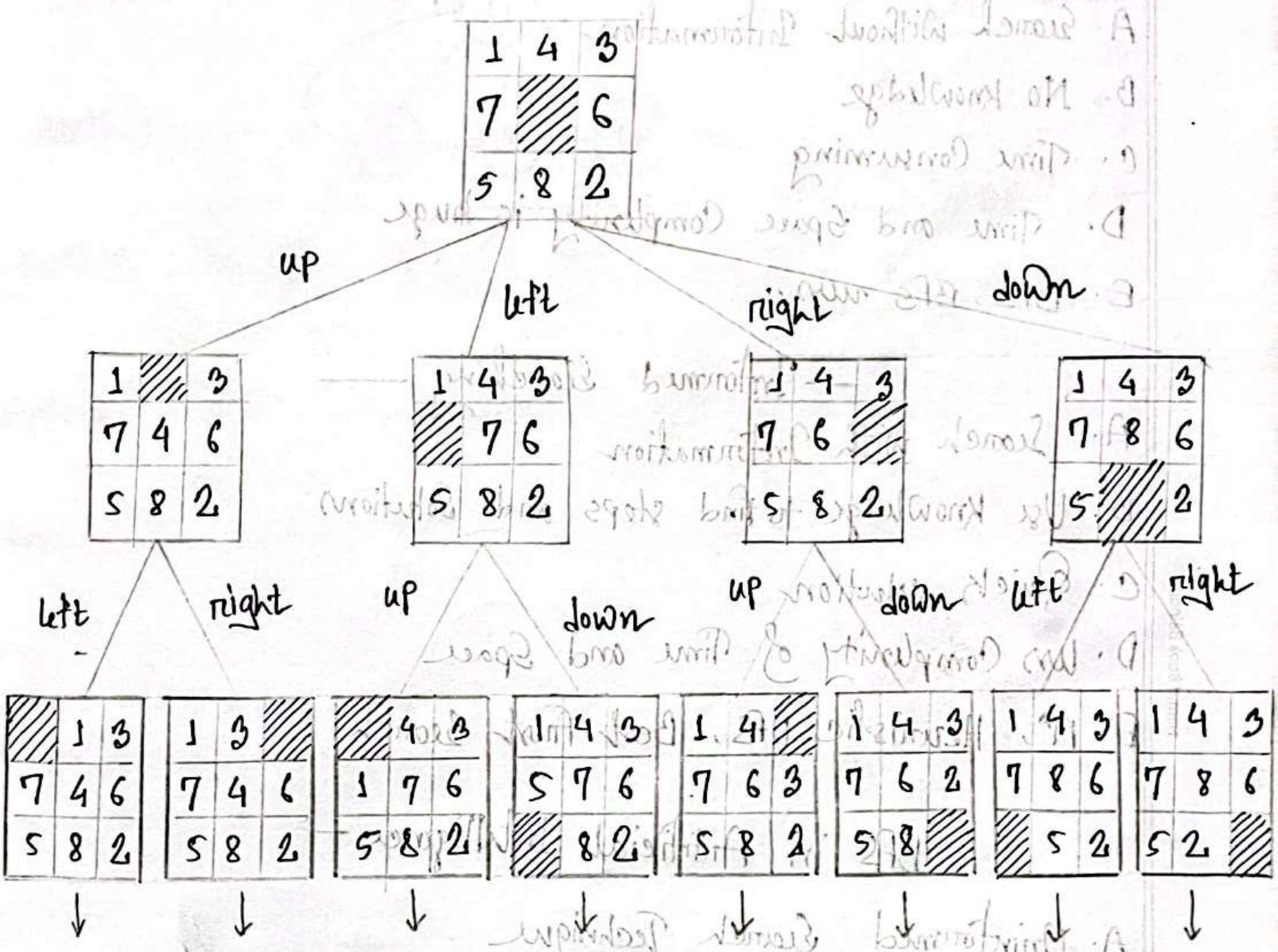
Subject : Date :

State Space Search



A*

Graph search → optimal → consistency

complete → admisible, consistency

Tree search → optimal, complete } → Admissible

Subject: | Date:

— Uninformed Searching —

A. Search Without Information

B. No Knowledge

C. Time Consuming

D. Time and Space Complexity is huge

E. DFS, BFS uses.

— Informed Searching —

A. Search with Information

B. Use Knowledge to find steps and Solutions

C. Quick solution

D. Less Complexity of Time and Space

E. A*, Heuristic DFS, Best First Search

— BFS in Artificial Intelligence —

A. Uninformed Search Technique

B. FIFO (Queue)

C. Shallowest Node

D. Complete

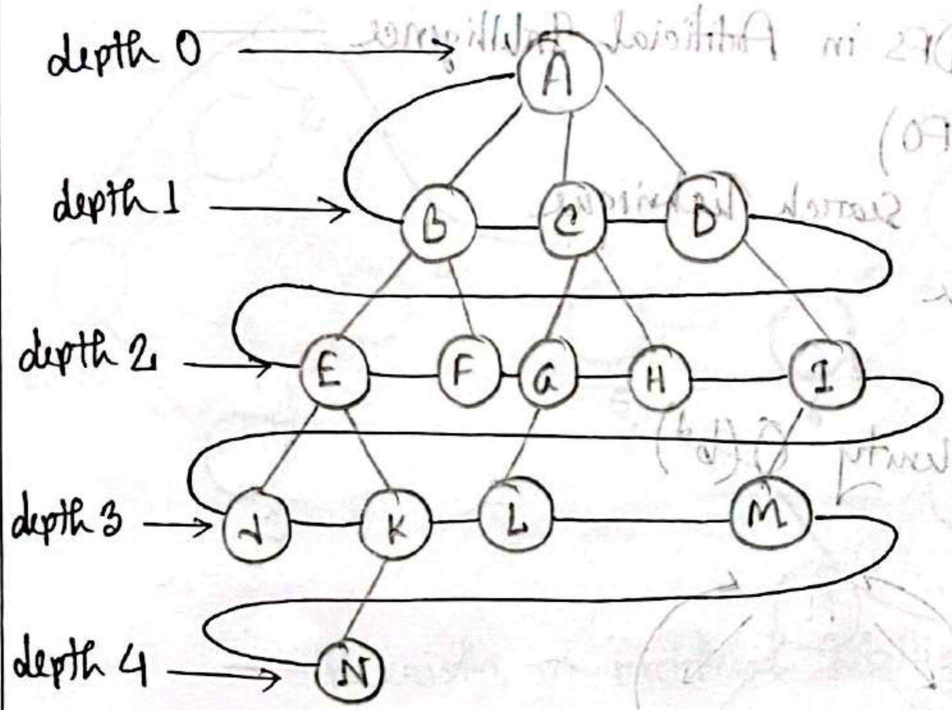
E. Time Complexity $O(b^d)$

b = Branch Factor (Child number)

d = depth

Subject:

Date:



for child node = 3

Complexity $= 3^4$

$= 81$

depth 0 → 3 A Nodes

depth 1 − (3×3) for BCD Nodes

depth 2 − (3×9) for ~~..~~ 9 Nodes

depth 3 − (3×15) for 15 Nodes

Subject: Date:

DFS in Artificial Intelligence

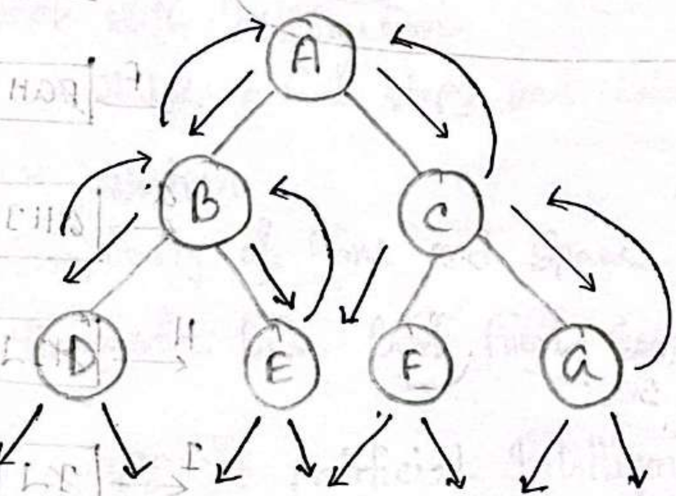
A. Stack (LIFO)

B. Uninformed Search Technique

C. Deepest Node

D. Incomplete

E. Time Complexity $O(b^d)$



[ACGFBED] Traversal Nodes Sequences

Bidirectional Search Algorithm

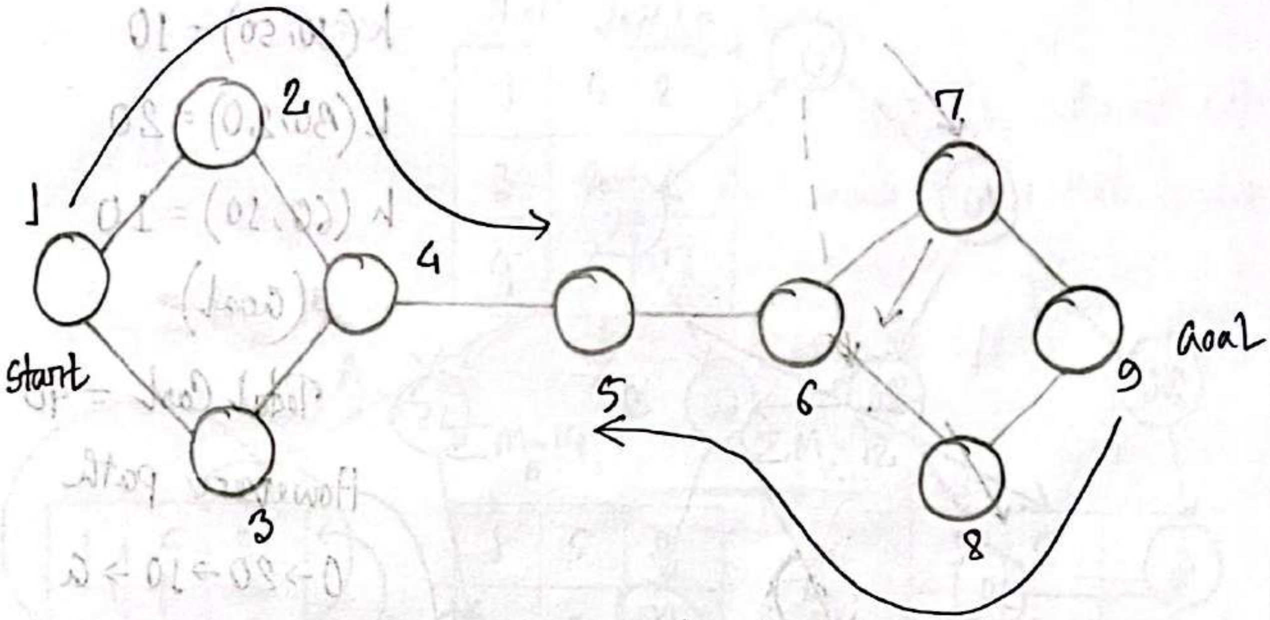
Two simultaneous search from on initial node to goal and backward from goal to initial, stopping when two meet.

Time Complexity: $2(b^{d/2})$

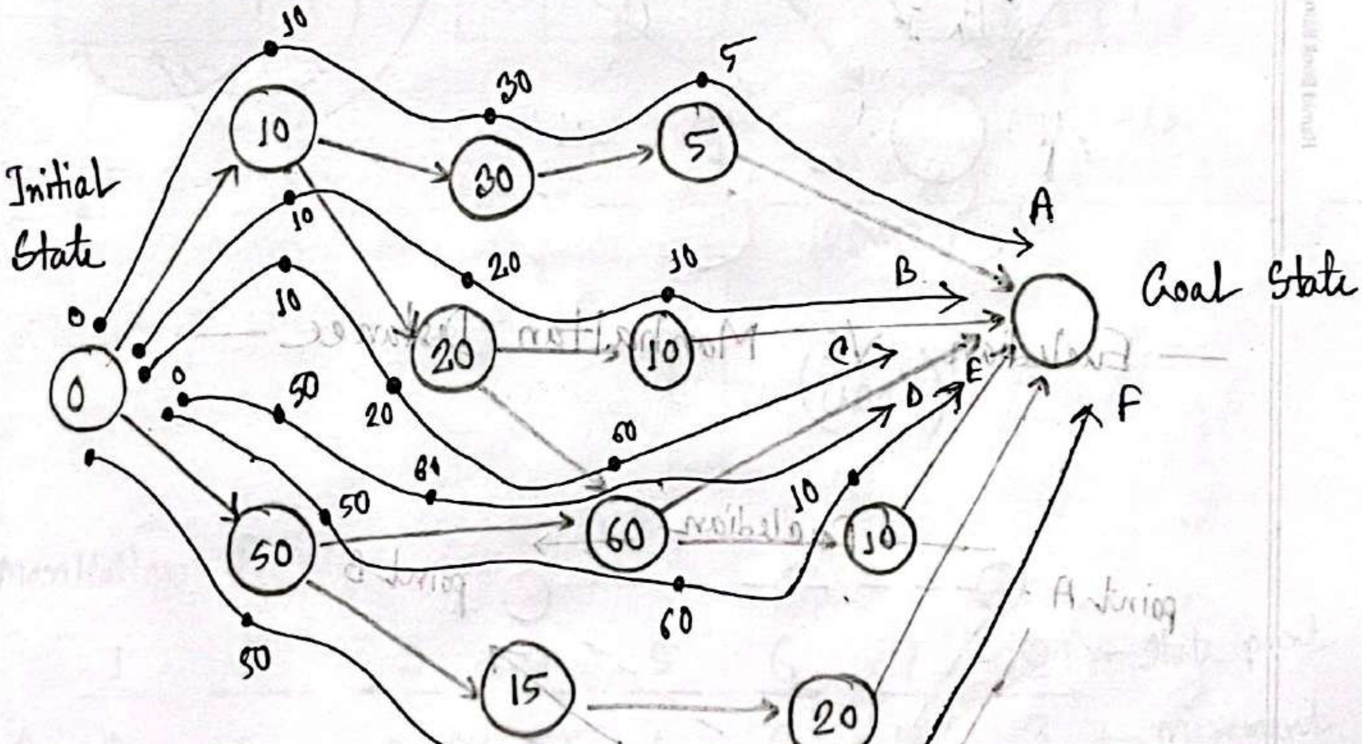
Complete in breadth first search

Not in depth first search

Subject: Date:

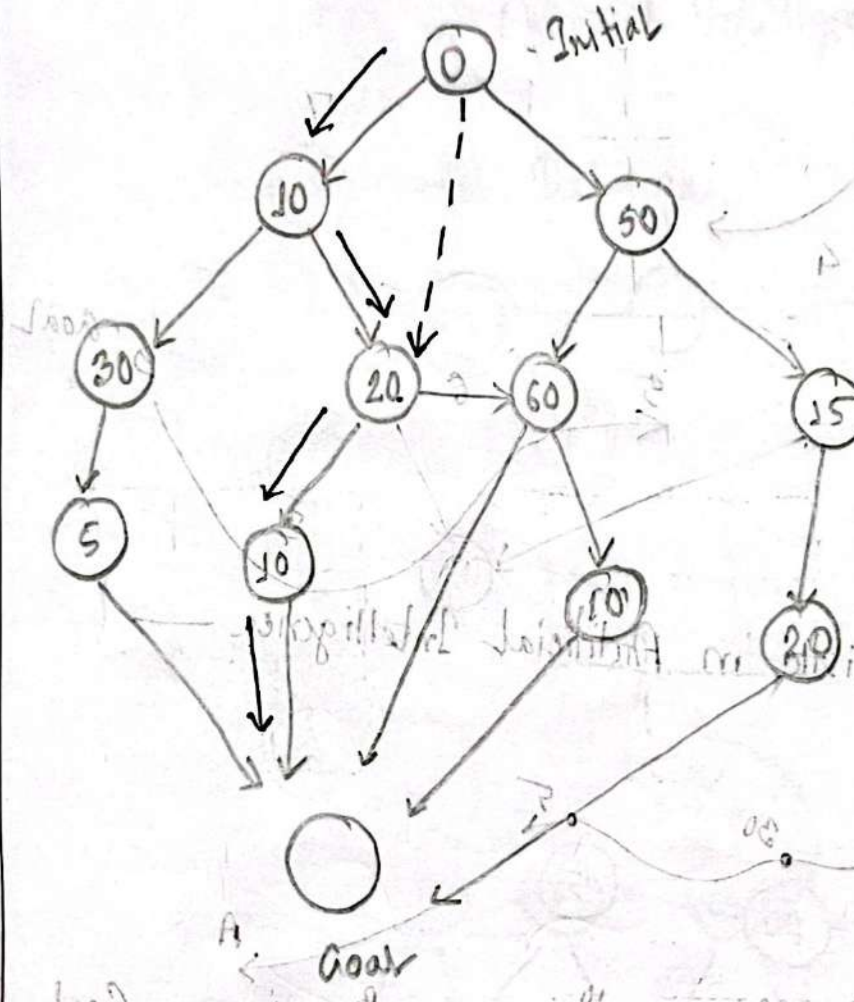


Heuristic in Artificial Intelligence



Subject :

Date :



h(10, 50) = 10

h(30, 20) = 20

h(60, 10) = 10

(Goal)

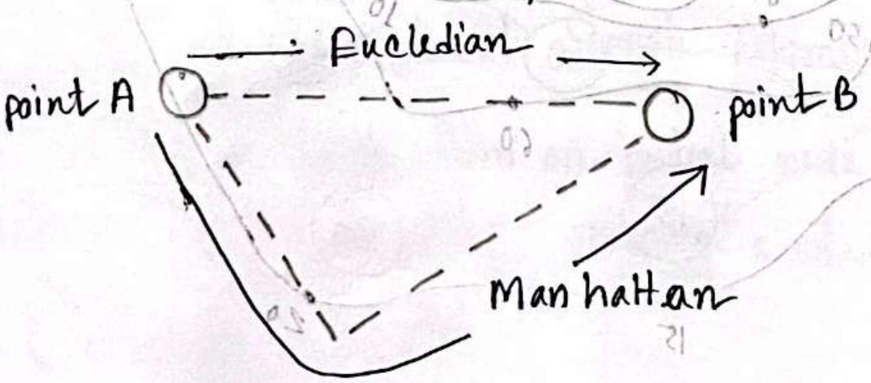
Total Cost = 40

However path

0 → 20 → 10 → G

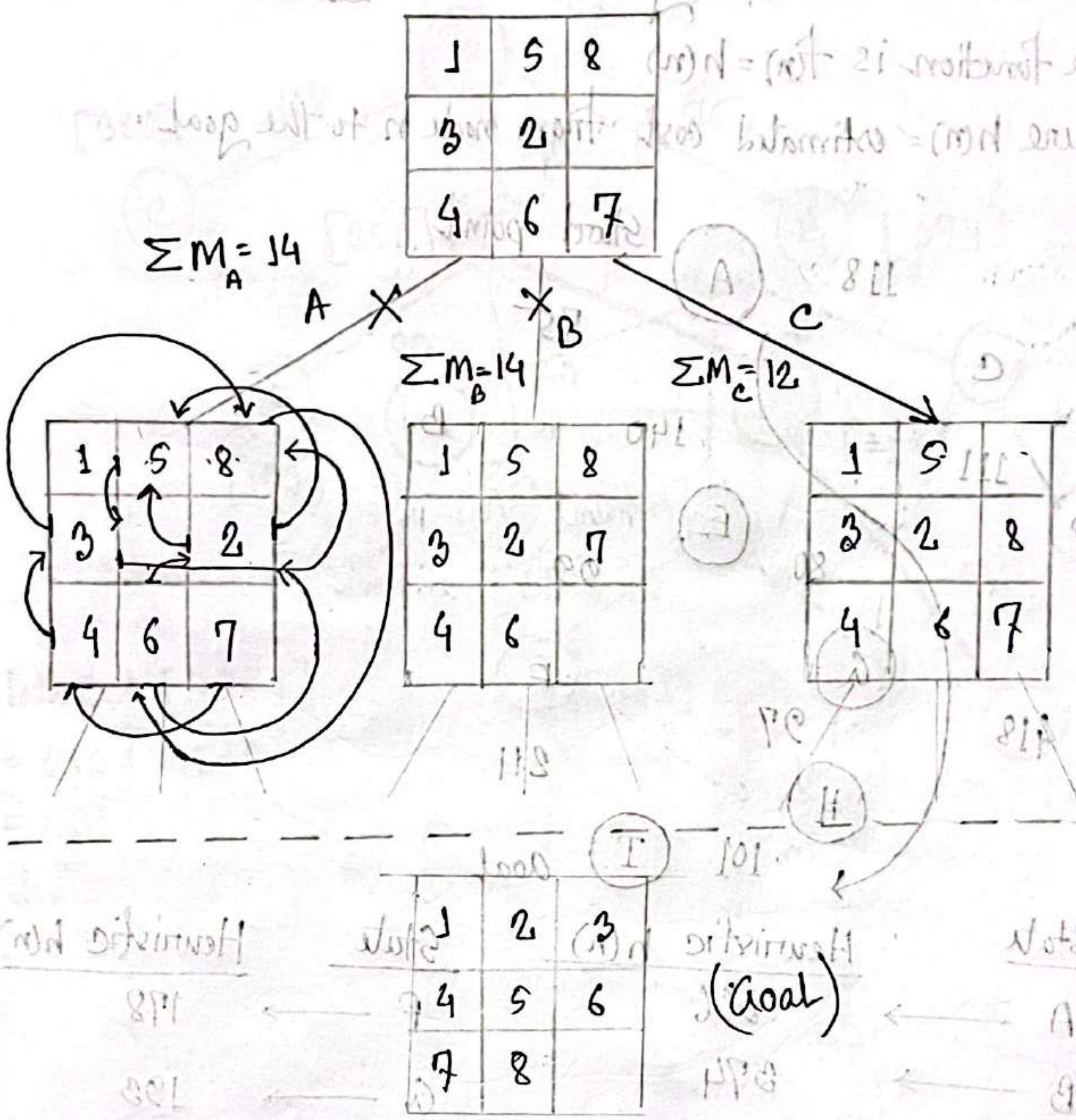
Cost 30 < 40

— Eucledian Vs Manhattan Distance —



Subject:

Date:



Manhattan Priority Function

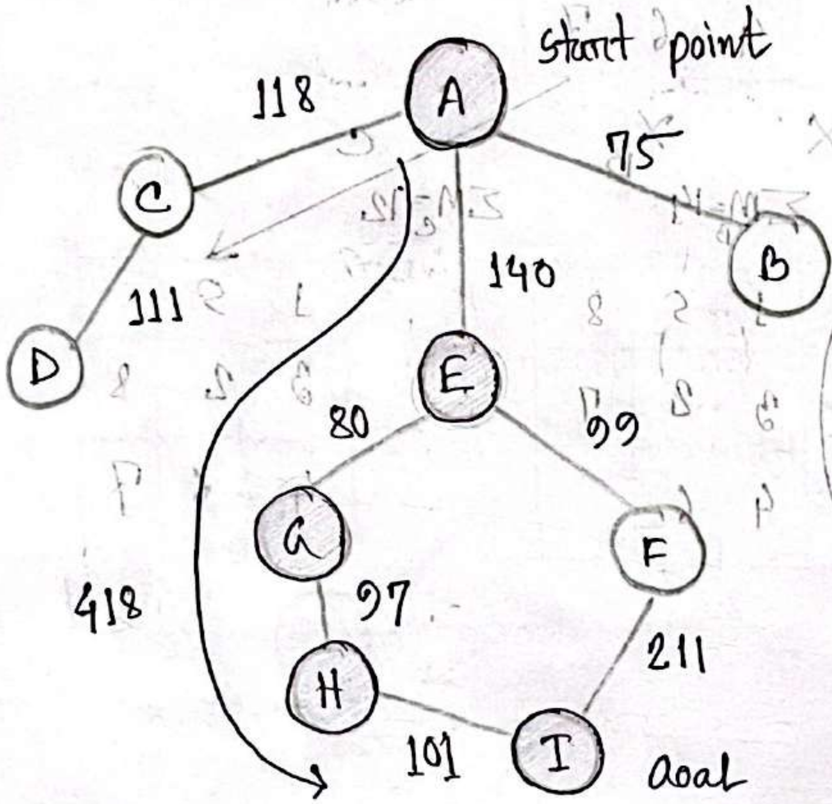
	1	2	3	4	5	6	7	8	← state point
A	0	2	3	1	1	2	2	3	← Movement
B	0	1	3	1	1	2	3	3	← Movement
C	0	1	3	1	1	2	2	2	← Movement

Subject: Date:

— Greedy Best-First Search —

The function is $f(n) = h(n)$

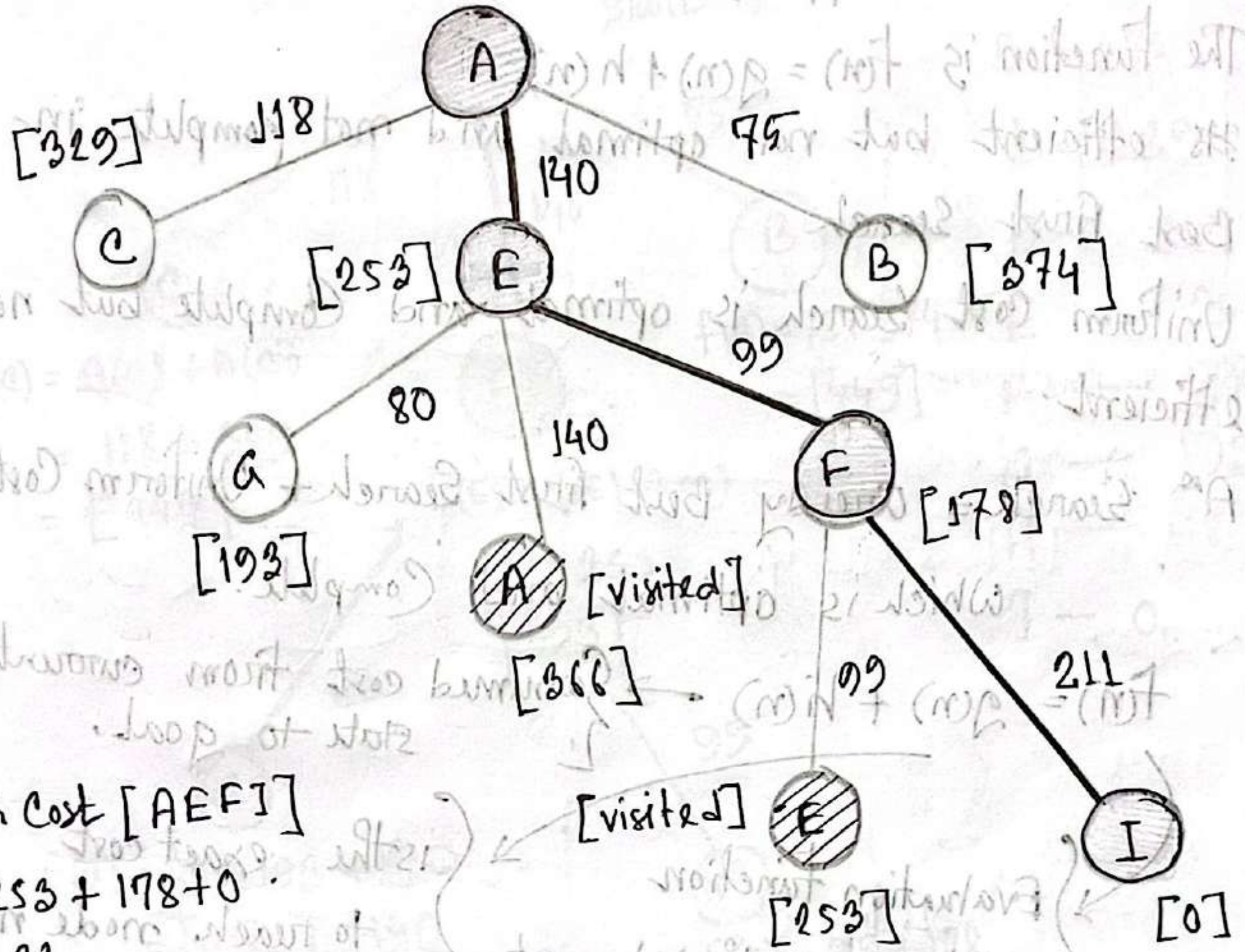
Where $h(n)$ = estimated cost from node n to the goal.



State	Heuristic $h(n)$
A →	366
B →	374
C →	329
D →	244
E →	253
F →	178
G →	193
H →	98
I →	0

Subject :

Date :



Path Cost [AEFI]

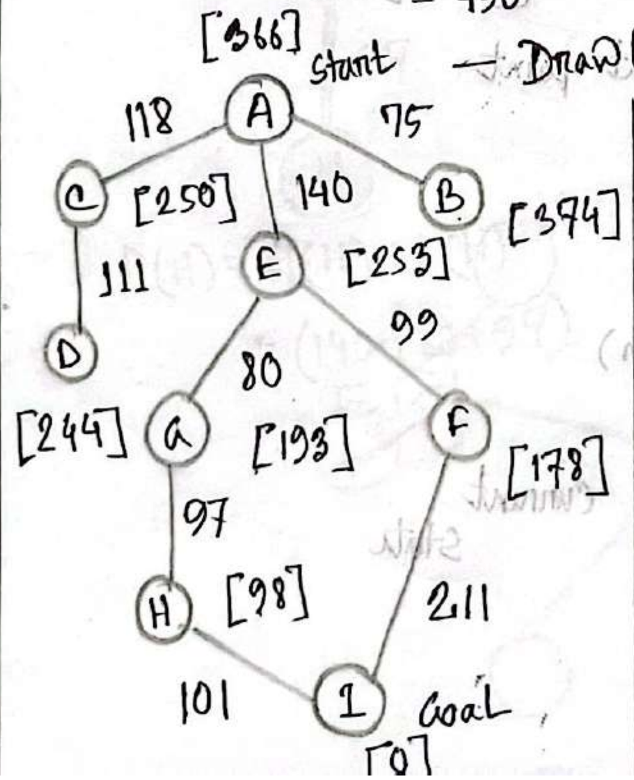
= 253 + 178 + 0

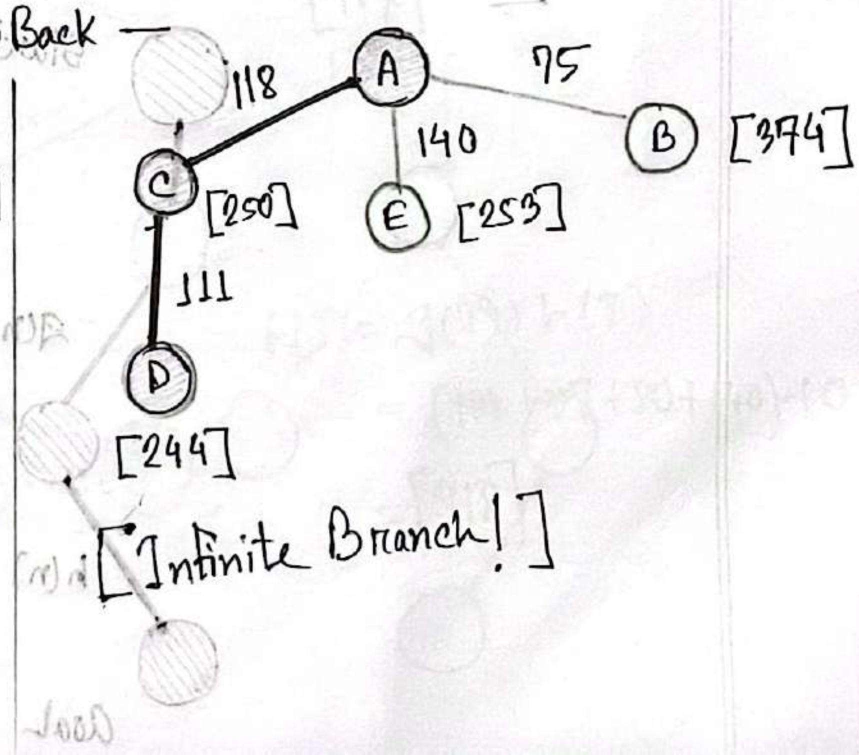
= 431

distance (AEFI) = 140 + 99 + 211

= 450

— Draw Back —





Subject : | Date :

— A* Search —

The function is $f(n) = g(n) + h(n)$

It's efficient but not optimal and not complete in Best First Search

Uniform Cost Search is optimal and Complete but not efficient.

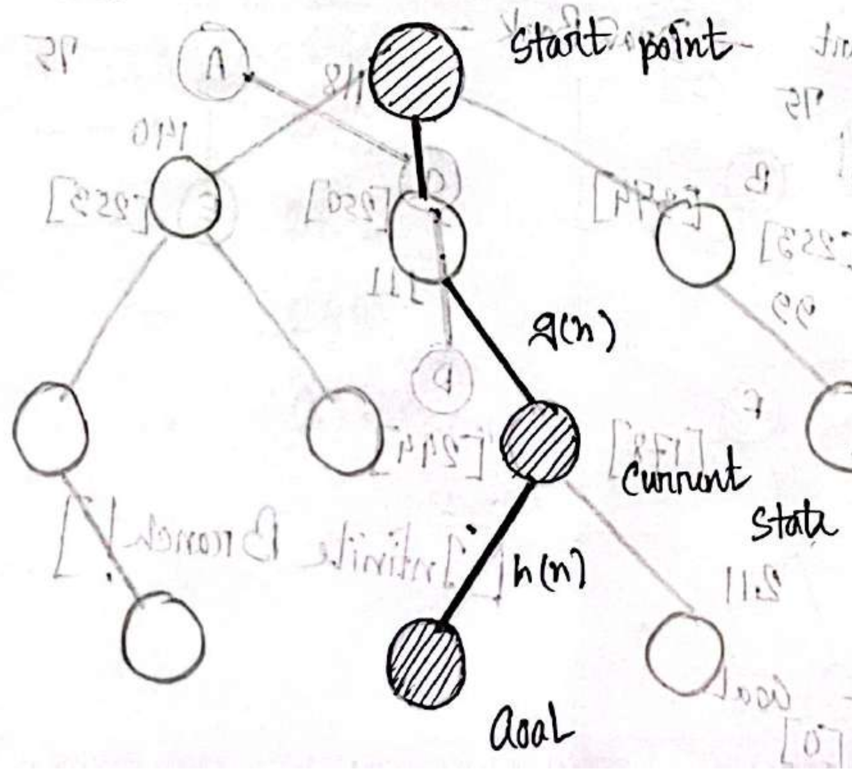
A* Search = Greedy Best first Search + Uniform Cost

which is optimal and Complete.

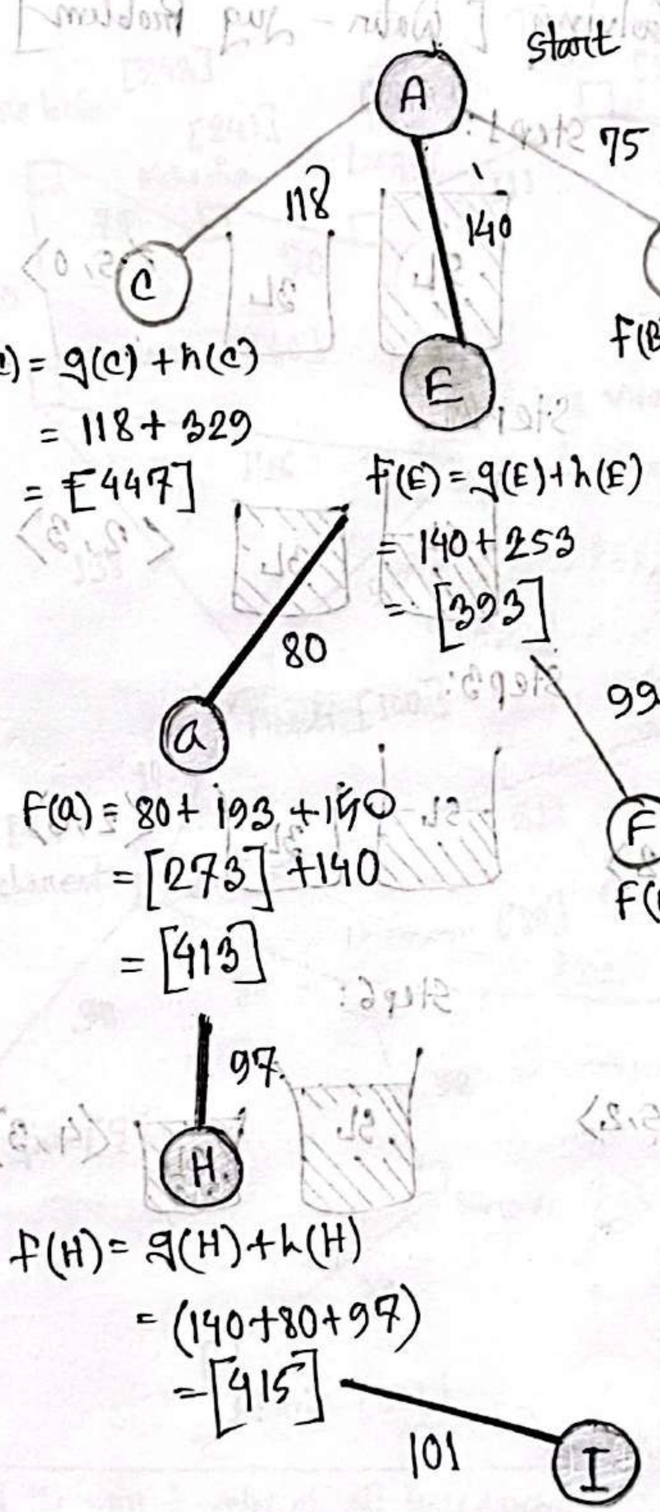
$f(n) = g(n) + h(n)$ → assumed cost from current state to goal.

$f(n)$ → Evaluation function which gives the cheapest solution cost

$g(n)$ → is the exact cost to reach node n from the initial state.



Subject : Date :



A — 366
B — 374
C — 329
D — 244
E — 253
F — 178
G — 193
H — 98
I — 0

$f(C) = g(C) + h(C)$
$= 118 + 329$
$= [447]$

$f(B) = 75 + 374$
$= [449]$

$f(E) = g(E) + h(E)$
$= 140 + 253$
$= [393]$

$f(G) = 80 + 193 + 140$
$= [273] + 140$
$= [413]$

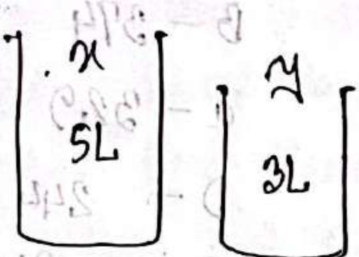
$f(F) = 99 + 178$
$= [277] + 140$
$= [417]$

$f(H) = g(H) + h(H)$
$= (140 + 80 + 97)$
$= [415]$

$f(I) = g(I) + h(I)$
$= (101 + 97 + 80 + 140) + 0$
$= [418]$

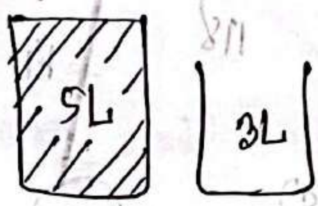
Subject: Date:

Problem Solving [Water - Jug Problem]



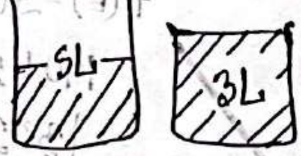
Jug x contains 5L
Jug y contains 3L
Goal is to have 4L in Jug x.

Step 1:



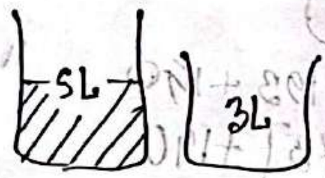
<5, 0>

Step 2:



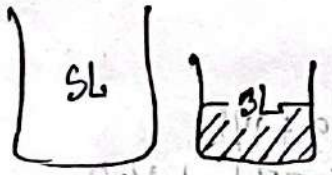
<2, 3>

Step 3:



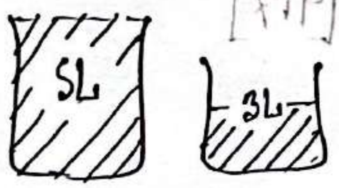
<2, 0>

Step 4:



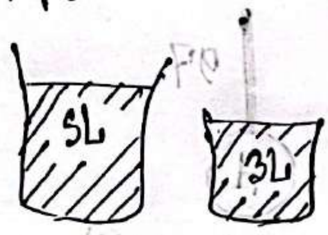
<0, 2>

Step 5:



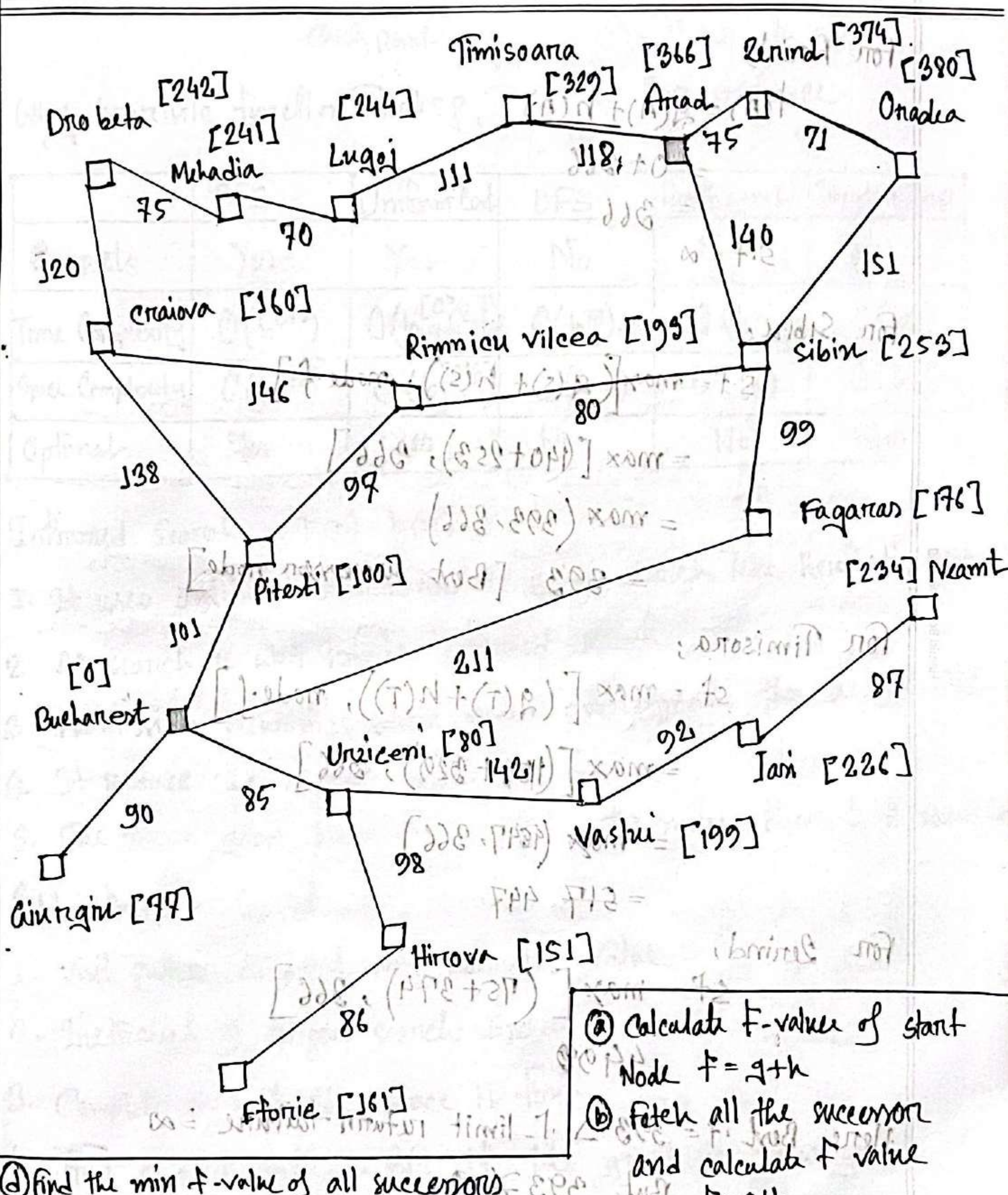
<5, 2>

Step 6:



<4, 3>

Subject: | Romania Map AI | Date:



ⓐ Calculate f-value of start Node $f = g + h$

ⓑ Fetch all the successor and calculate f value $f = g + h$

ⓒ find max(f_s, f_{parent}) This is the new f value of succ.

ⓓ Find the min f-value of all successors

ⓔ

Subject: Date:

For Arad;

$F(A) = g(A) + h(A)$

$= 0 + 366$

$= 366$

$s.f = \infty$

For Sibiu;

$s.f = \max[(g(s) + h(s)), node.f]$

$= \max[(140+253), 366]$

$= \max(393, 366)$

$= 393$ [Best successor node]

For Timisora;

$sf = \max[(g(T)+h(T)), node.f]$

$= \max[(118+329), 366]$

$= \max(447, 366)$

$= \cancel{517}\ 447$

For Zerind,

$sf = \max[(75+374), 366]$

$= 449$

Here Best $f = 393 > f\_limit$ return failure $= \infty$

But $393 \ngtr \infty$

Alt node $= \cancel{449}\ 447$

$f\_limit = \min(f\_limit, alt.f) = \min(\infty, 449) = \cancel{449}\ 447$

Hamid Book Bainding

Subject: Date:

Check point

$C^* =$ Total path cost
$\varepsilon =$ Avg step size.

Why heuristic function works?

	BFS	Uniform Cost	DFS	Depth Limit	Iterative Deep
Complete	Yes	Yes	No	No	Yes
Time Complexity	$O(b^{d+1})$	$O(b^{[C^*/\varepsilon]})$	$O(b^m)$	$O(b^l)$	$O(b^d)$
Space Complexity	$O(b^{d+1})$	$O(b^{[C^*/\varepsilon]})$	$O(bm)$	$O(bl)$	$O(bd)$
Optimal	Yes	Yes	No	No	Yes

Informed Search : $F(n) = h(n)$

1. It uses additional information to guide search like heuristics value
2. A* search is well known informed search
3. Admissible heuristics value never overestimate the actual cost
4. It reduce the search state
5. The more good heuristics value, the more efficient it would be

Uninformed Search:

1. visit paths without any heuristics value
2. Inefficient of larger search spaces
3. Complete only if the space is finite
4. Find shortest path in BFS if the graph is unweighted
5. Explore the entire search space with worst case
6. Both Tree and graph search

Subject: Date:

1. BFS is good for shortest path
2. DFS is good for deep search
3. A* for efficiency.
4. Due to heuristics value informed search uses less state than that of uninformed.
5. IDS (Iterative deepening search) uses DFS for low memory uses and BFS for completeness simultaneously.
6. A poor heuristics value can mislead the search and a good heuristics can ~~be~~ lead the A* search.
7. Overestimating the cost, may cause A* to skip the optimal path. So, heuristics never overestimate the actual path cost.
8. Bidirectional search ~~d~~ reduce the search space as it can search forward from the start and backward from the goal.

Subject : Date :

Sudoku Puzzle

Consistency Check $h(n) \leq c(n,m) + h(m)$

$h(A) = 3,\ h(B) = 2,\ c(A,B) = 1$ (Actual path cost)

$3 \leq 1 + 2$ True

$h(B) = 2,\ h(C) = 2;\ c(B,C) = 3$

$2 \leq 3 + 2$ True

$h(B) = 2,\ h(D) = 1,\ c(B,D) = 2$

$2 \leq 2 + 1$ True

$h(C) = 2,\ h(D) = 1,\ c(C,D) = 2$

$2 \leq 2 + 1$ True

$h(D) = 1,\ h(G) = 0,\ c(D,G) = 3$

$1 \leq 3 + 0$ True

Consistency satisfied.

Subject: | Date:

Admissible Graph

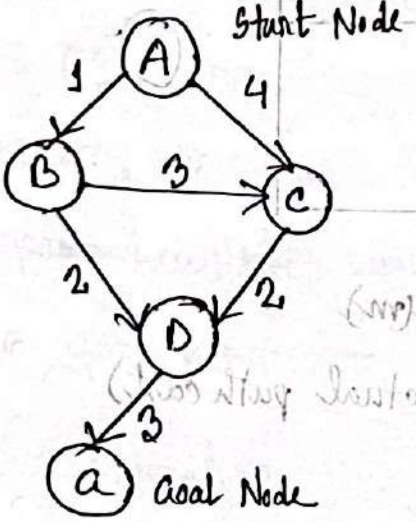
Conditions:

ⓐ Non-negative: $h(n)$ is admissible if $h(n) \geq 0$ for all nodes

ⓑ Never overestimate the actual cost. $h(n) \leq h^*(n)$

Underestimate $h(n) \leq h^*(n)$ | Here $h(n)$ = estimated value

Overestimate $h(n) \geq h^*(n)$ | $h^*(n)$ = actual value



Heuristic Table ($h(n)$)

equation: Minimum edge × steps to G

A → 3 steps and 1 edge = 3

B → 2 steps and 1 edge = 2

C → 2 steps and 1 edge = 2

D → 1 step and 1 edge = 1

G → 0 step and 0 edge = 0

Actual cost ($h^*(n)$)

A → C → D → G = 9

B → D → G = 5

C → D → G = 5

D → G = 3

G = 0

Admissible check

$h(n) = 3,\ h^*(n) = 9$

$h(n) \leq h^*(n)$ ✓

$h(n) = 2,\ h^*(n) = 5$

$h(n) \leq h^*(n)$ ✓

$h(n) = 2,\ h^*(n) = 5$

$h(n) \leq h^*(n)$ ✓

$h(n) = 1,\ h^*(n) = 3$

$h(n) \leq h^*(n)$ ✓

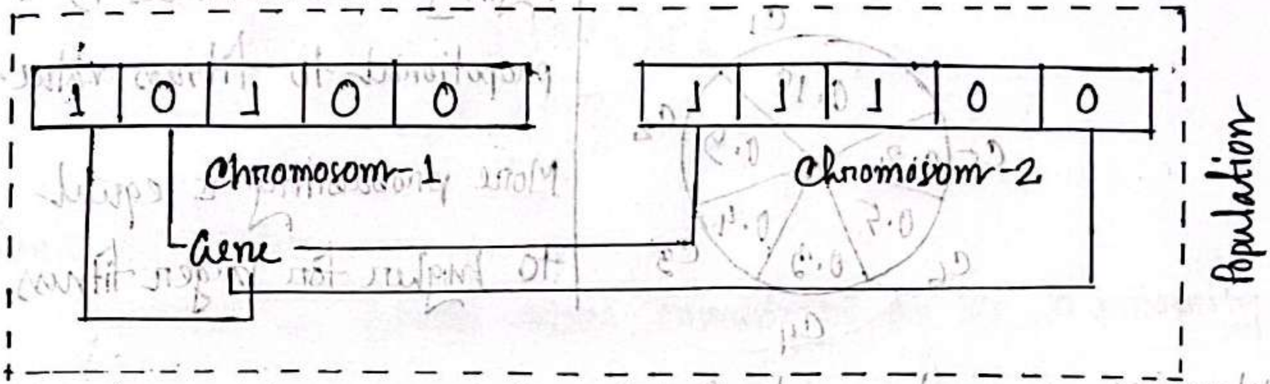
$h(n) = 0,\ h^*(n) = 0$

$h(n) \leq h^*(n)$ ✓

So, the graph is admissible

Subject :   Date :

Genetic Algorithm



A. Number of genes per chromosome

B. The codes value

C. The size of population per generation

D. Crossing over probabilities

E. Mutation Probabilities

F. Termination Criteria

Selection of Chromosome

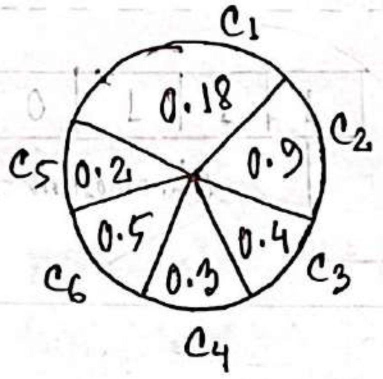
A. Roulette Wheel Selection

B. Rank Selection

C. Tournament Selection

Subject: Date:

Roulette Wheel Selection:



length of circumference is propotional to fitness value.

More probability is equal to higher for larger fitness.

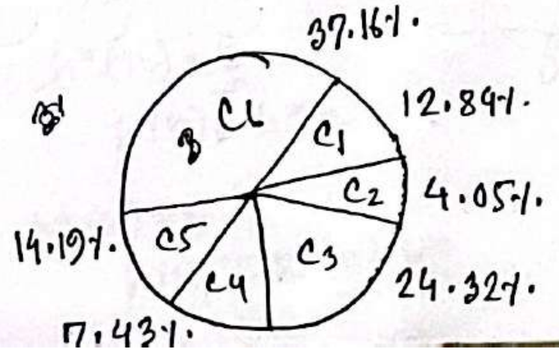
Determine parcentage of Roulette Wheel and Actual Count for six chromosomes With fitness value of 19, 6, 36, 11, 21 and 55.

Chromosome	Fitness	% of wheel	Probability	Expect Count	Actual Count
C_1	19	0.1284	12.84%	0.77	1
C_2	6	0.0405	4.05%	0.243	0
C_3	36	0.2432	24.32%	1.46	~~2~~ 1
C_4	11	0.0743	7.43%	0.44	1
C_5	21	0.1419	14.19%	0.85	1
C_6	55	0.3716	37.16%	2.23	2
	$\Sigma F = 148$				

Probality $\text{Count}_i = \frac{F_i}{\Sigma F}$ | For $C_1 = \frac{19}{148} = 0.1284$

$= 12.84\%$

Pie chart:



Subject: Date:

Expected Count = Probabilities $\times$ Total Chromosome

For, c_1 = 0.1284×6

= 0.77.

Rank Selection:

The less fitness values considered as the priopity rank.

From the previous table; Selective Pressure [1.0, 2.0]

Chromosome	Fitness	Rank	New Fitness	Previous %	New %
C_1	19	3	~~0.5~~ 0.9	12.84%	15%
C_2	6	1	~~0.75~~ 0.5	4.05%	8.33%
C_3	36	5	1.3	24.32%	21.667%
C_4	11	2	0.7	7.43%	11.667%
C_5	21	4	1.1	14.19%	18.334%
C_6	55	6	1.5	37.16%	25%
	$\Sigma F = 148$		$\Sigma nF = 6$		

$$Fit(X_n) = 2 - SP + 2(SP-1)\frac{(Pos-1)}{(Nind-1)}$$

SP → Selective pressure

(Pos − 1) → position

(Nind − 1) → Total number of chromosome

2, 2 → Constant value

$Fit(X_n)$ → Rank Number (position)

Here,

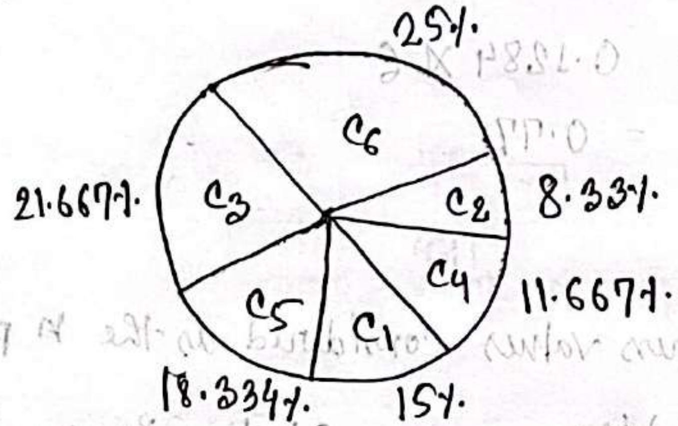
$$SP = \frac{2.0 + 1.0}{2}$$

$$= 1.5$$

Selective pressure = [1.0, Nind − 1]

Subject: | Date:

Pie chart:

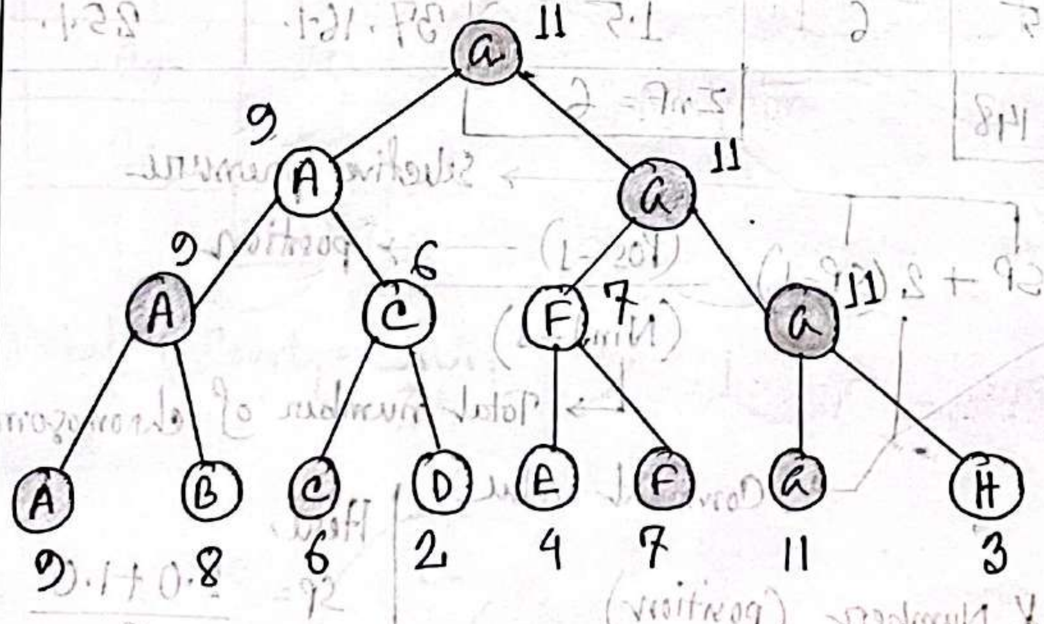


Tournament selection:

(a) select k individual from the population and perform a tournament among them.

(b) Select the best individual from the k individuals.

(c) Repeat the process 1 and 2 until we have the desired amount of population.



Subject:

Date:

Crossover And Mutation

Crossover

0	1	0	0	1
1	0	1	1	0

0	1	0	1	0
1	0	1	0	1

(A) Single-site Crossover (B) Two-site Crossover (C) Crossover Mask

Single-site Crossover: (At point 5)

P_1	1	0	0	1	1	0	0	1	0
P_2	0	1	0	1	1	1	0	0	1

Strings Before mating

C_1	1	0	0	1	1	1	0	0	1
C_2	0	1	0	1	1	0	0	1	0

Strings After mating

Two-site Crossover: (At point 2 and 5)

P_1	1	0	1	0	1	0	0	1	0
P_2	0	1	0	1	0	1	0	0	1

C_1	1	0	0	1	0	0	0	1	0
C_2	0	1	1	0	1	1	0	0	1

Subject: Date:

Crossover Mask

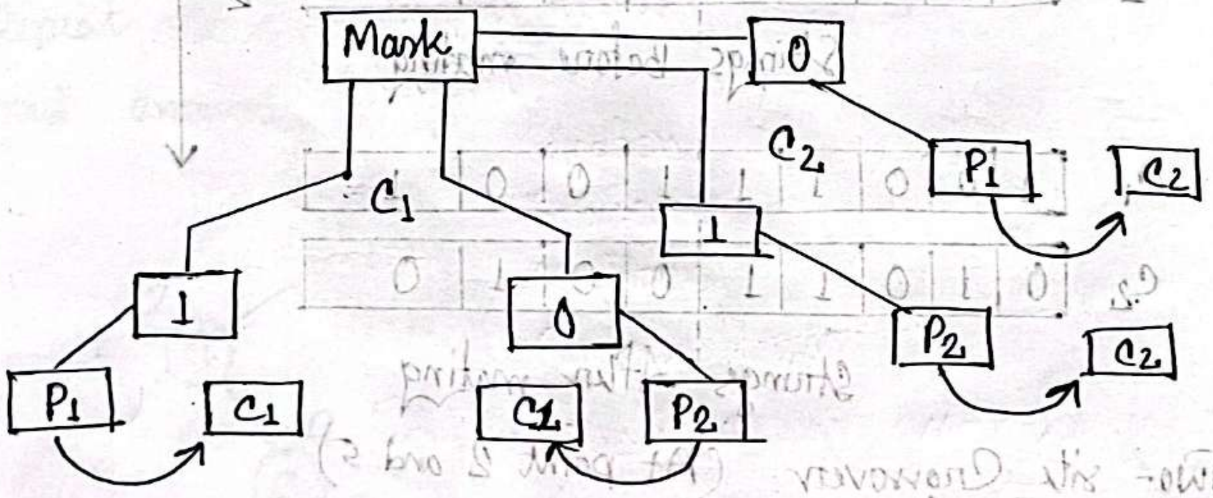
P_1	1	0	0	0	1	0	0	1	0
P_2	0	1	0	1	1	1	0	0	1

Strings Before mating

C_1	1	1	0	0	1	1	0	0	0
C_2	0	0	0	1	1	0	0	1	1

Strings After Mating

Mask	1	0	0	1	1	0	0	0	1



Subject: | Date:

Maximizing Function

Maximize the function $f(x) = x^2$, Where the range of x is [0,31] and maximum number of generation is 3

Step 1: Initialization

00000, 00001, 00010, 00011, ... 11111

max number of generation = 3

Step 2: Selection (1st generation)

Chromosome	Encoding	Value of X	Fitness x^2
C_1	00101	5	25
C_2	10000	16	256
C_3	01001	9	81
C_4	11100	28	784
			$\Sigma f(x) = 1146$

Step 3: Crossover (2nd generation)

Chromosome	Encoding	Value of X	x^2	After Crossover
C_1	0010|1	4	16	00100
C_2	1000|0	17	289	10001
C_3	01|001	12	144	01100
C_4	11|100	25	625	11001
			$\Sigma f(x) = 1074$	

Subject : Date :

Mutation (2nd generation)

Chromosome	Encoding	Values of x^2	$f(x)=x^2$	After Mutation
C_1	00100	16	4	00100
C_2	10001	441	21	10101
C_3	01100	144	12	01100
C_4	11001	729	27	11011
			$\Sigma f(x) = 1330$	

Pseudo-Code of Genetic Algorithm :

```
START
    Generate the initial population
    Compute fitness
    REPEAT
        Selection
        Crossover
        Mutation
        Compute fitness
    UNTIL population has over converged
STOP
```

Subject : Date :

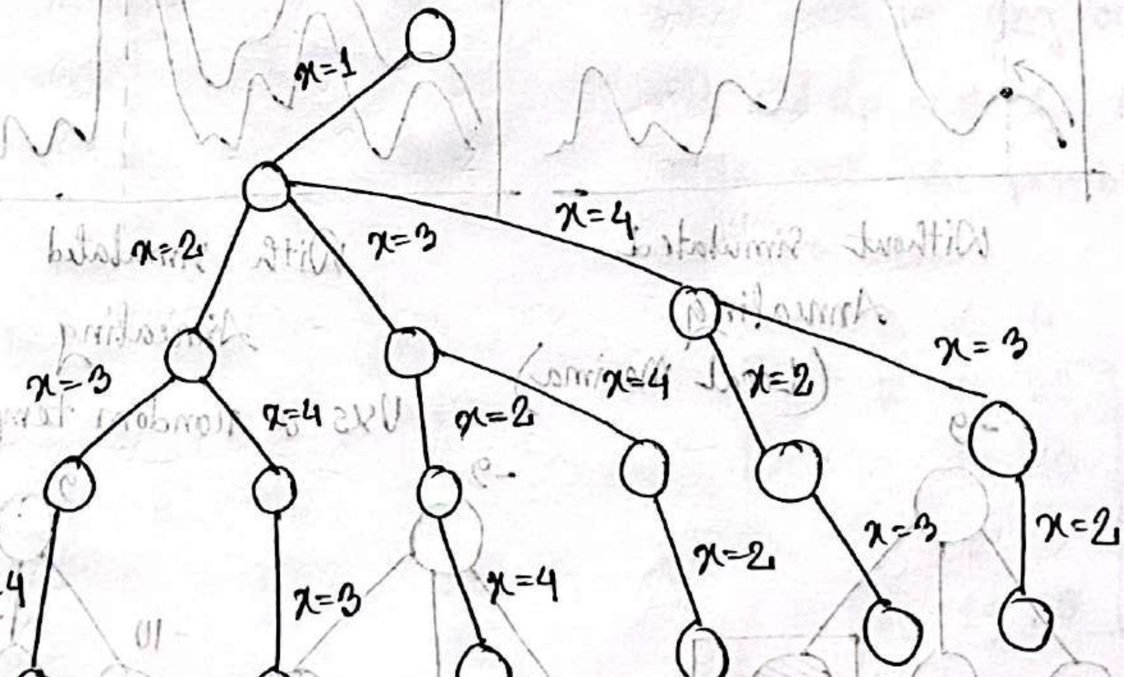
Queen Problem using Backtracking

	1	2	3	4
1	Q_1			
2		Q_2		
3			Q_3	
4				Q_4

4 Queens

$Q_1 \mid Q_2 \mid Q_3 \mid Q_4$

Queens can not be in same row and same columns.



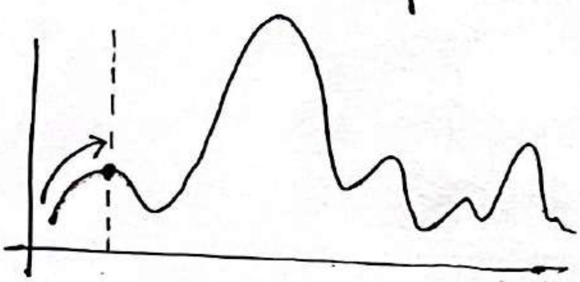
Subject: Date:

Local Search

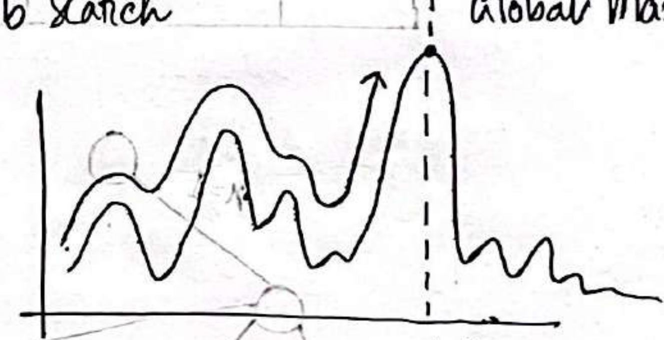
ⓐ Not concerned with path
ⓑ Solution itself matters
ⓒ Not concerned with perfect solution
ⓓ less memory required as its not using Backtracking
ⓔ less time required
ⓕ Applicable for large size of graph

Hill Climb search

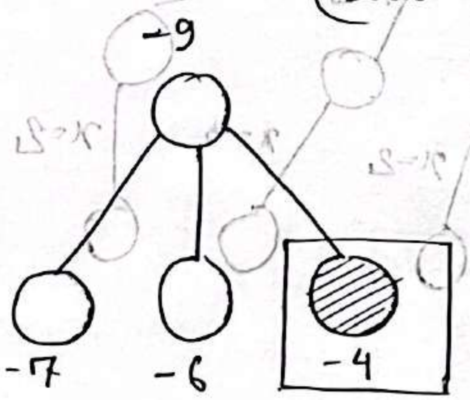
Global Maxima



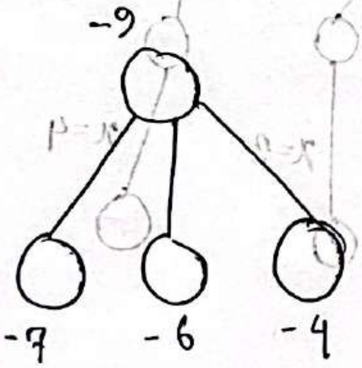
Without simulated Annealing (Local Maxima)



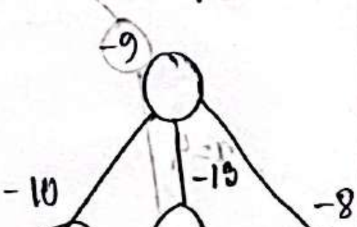
With simulated Annealing
Uses a random temperature



Hill climb Steepest
select the maximum



Hill climb Stochastic
select a random node



Hill Climb First choice
select the first occurrence of minimum.

Subject:

Date:

Simulated Annekling

It allowes downWard step also.

Advantages:

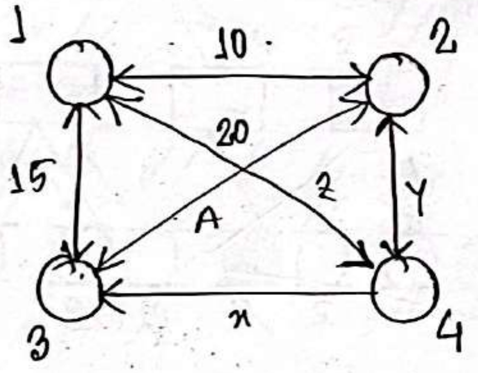
Always gives optimal Solutions

Disadvantages:

Slow process

Cann't tell if an optimal solution is found.

TSP (Travelling Salesman Problem)



Start from node 1, travel the entire path in any order and end up with the same node with the possible minimum cost.

ΔE = New Cost - Old Cost

if $\Delta E < 0$, accept

if $\Delta E > 0$, find the probability

$P = e^{-\Delta E/T}$.

$$\begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 0 & 10 & 15 & 20 \\ 5 & 0 & 25 & 10 \\ 15 & 30 & 0 & 5 \\ 15 & 10 & 20 & 0 \end{bmatrix} \end{array}$$

Higher temperature allows Worse Solutions to be accepted. helping escap local optima. When the temperature is high, P is close to Worse solutions.

Subject: Date:

Question Pattern

Pobular Problem

Travelling Salesman problem
- Local Search Algorithm
- Hill Climbing Algorithm

N- Queens Problem
- Hill Climbing
- Local Search

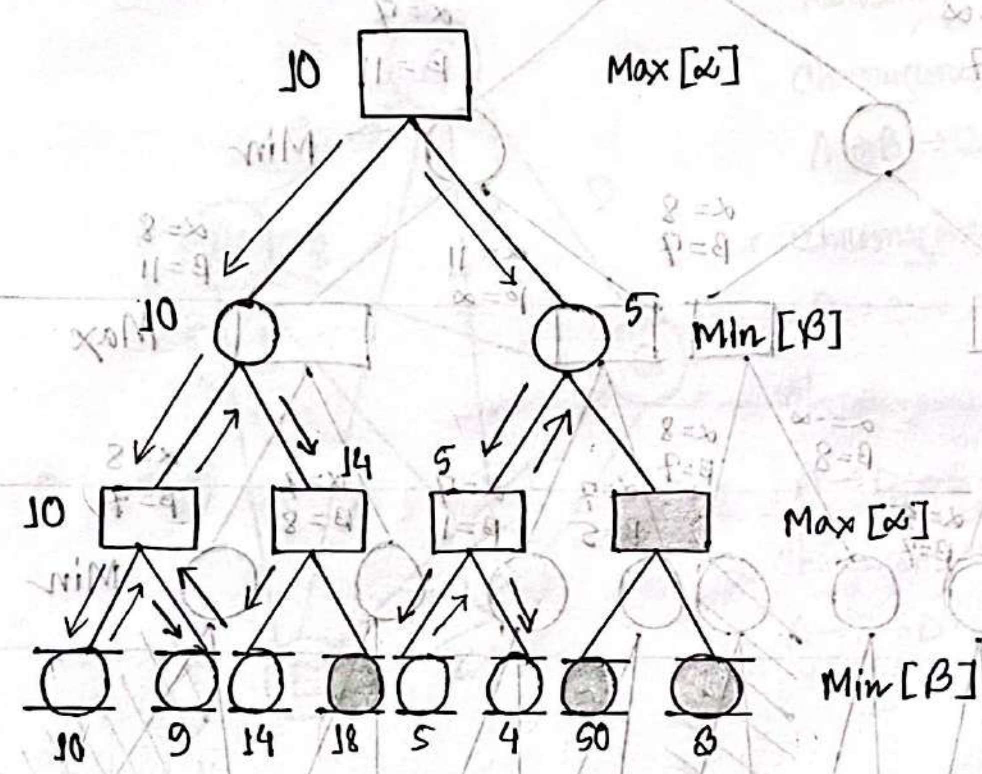
8- Putt Puzzle / 15- Puzzle Problem

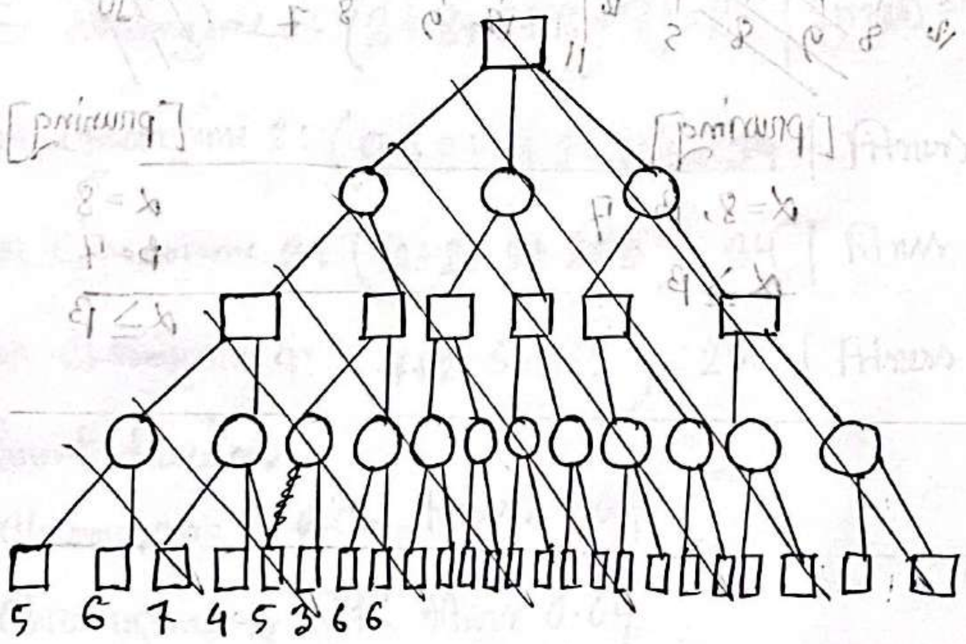
Subject: 						Date:

MiniMax using Alpha-Beta Pruning

Max contains Alpha

Min contains Beta

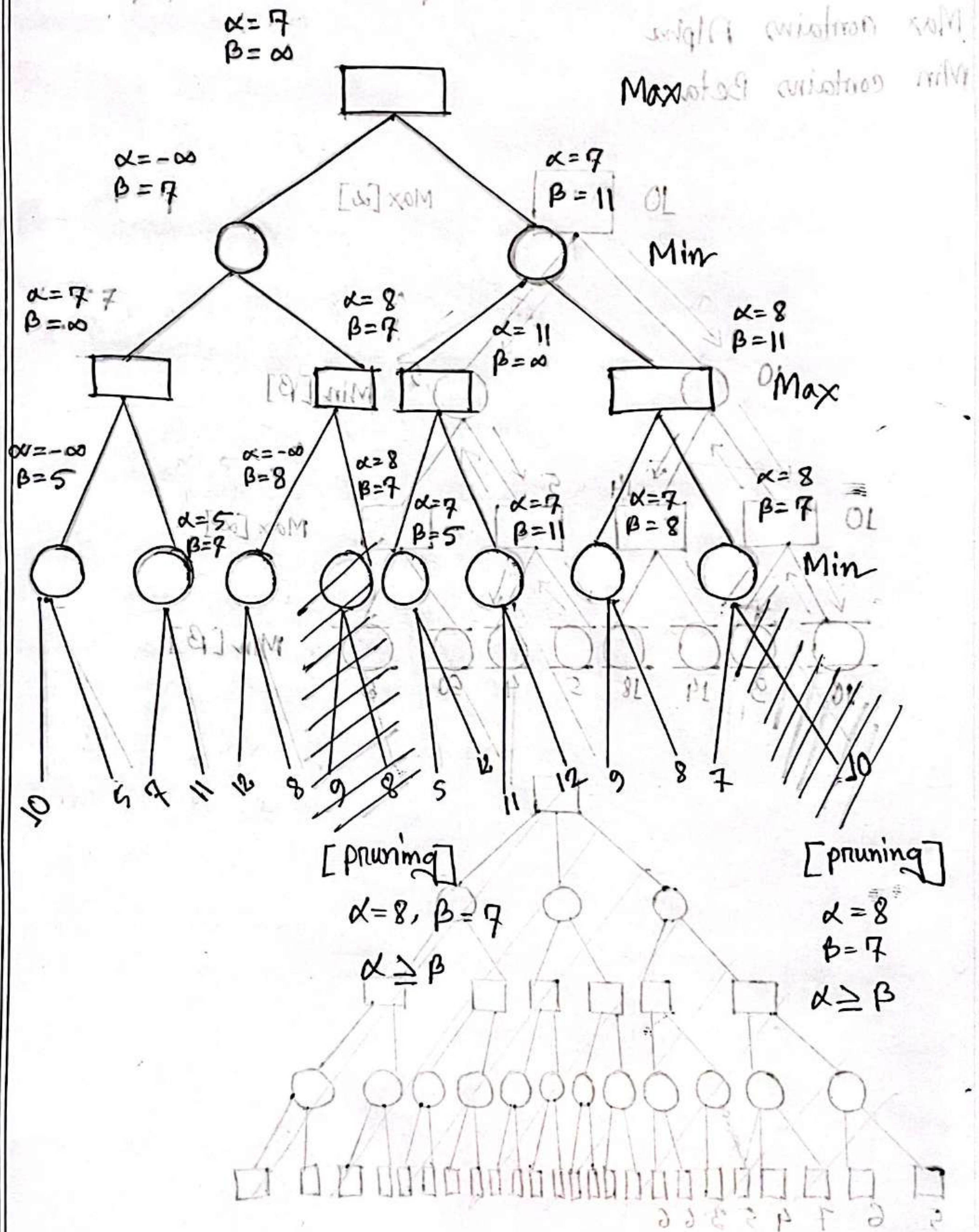




Subject :

Date :

Minimax Alpha Beta pruning Simulation

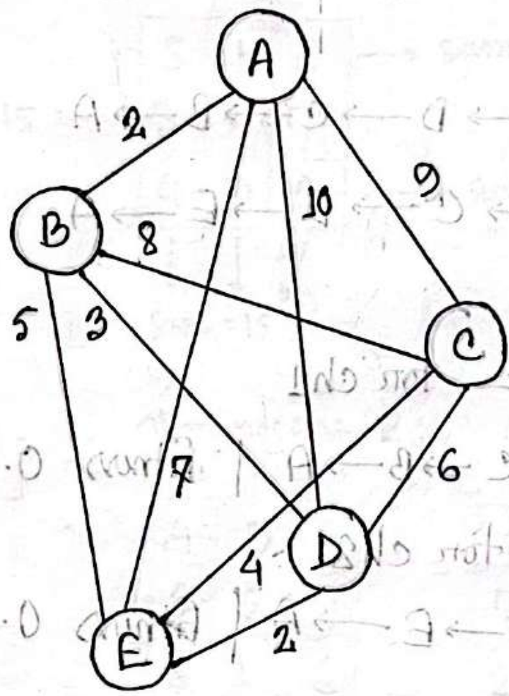


Subject: Date:

Genetic Algorithm Problem Solving AI

Problem 1:

Travelling Salesman Problem



Randomly selection:

Chromosome 1:
$A \to B \to C \to D \to E \to A$

Chromosome 2:
$A \to C \to B \to E \to D \to A$

Chromosome 3:
$A \to D \to E \to B \to C \to A$

Chromosome 4:
$A \to E \to D \to C \to B \to A$

Fitness Calculation:

For chromosome 1: $(2+8+6+2+7) = 25$ | Fitness $= \frac{1}{25} = 0.04$

For chromosome 2: $(9+8+5+2+10) = 34$ | Fitness $= 1/34 = 0.0294$

For chromosome 3: $(10+2+5+8+9) = 34$ | Fitness $= 1/34 = 0.0294$

For chromosome 4: $(7+2+6+8+2) = 25$ | Fitness $= 1/25 = 0.04$

Parent Selection:

Chromosome 1 with fitness 0.04

Chromosome 4 with fitness 0.04

Subject: Date:

Crossover:

Chromosome 1: A → B → C → D → E → A

Chromosome 2: A → E → D → C → B → A

Child 1: A → B → D → C → B → A

Child 2: A → E → C → D → E → A

Mutation:

Swap B with E for ch1

A → E → D → C → B → A | Fitness 0.040

Swap C with B for ch2

A → E → B → D → E → A | Fitness 0.0416

The answer A → E → B → D → E → A

Distance: 24

Fitness: 0.0416

Subject: Date:

Problem 2:

Rubik's Cube Problem

given that. The final state is

8	1	6	→ sum = 15
3	5	7	sum 15 ←
4	9	2	→ sum = 15

↓ ↓ ↙

sum = 15

Chromosome 1:

2	7	6
9	5	1
4	3	8

Chromosome 2:

4	9	2
3	5	7
8	1	6

Chromosome 3:

6	1	8
7	5	3
2	9	4

Chromosome 4:

9	3	5
4	1	8
7	2	6

Fitness calculation:

$$\text{fitness} = \frac{1}{1+|(15-\text{row sum})|+|(15-\text{column sum})|+|(15-\text{diagonal sum})|}$$

Chromosome 1:

Rows: $|15-15|+|15-15|+|15-15| = 0$

Column: $|15-15|+|15-15|+|15-15| = 0$

Diagonal: $|15-15|+|15-15| = 0$

Fitness $= \frac{1}{(0+1)} = 1$

Subject : | Date :

Chromosome 2:
Fitness: 1

Chromosome 3:
Fitness: 1.0

Chromosome 4:

Rows: $|15-17| + |15-13| + |15-15| = 4$

Columns: $|15-16| + |15-13| + |15-16| = 6$

Diagonals: $|15-16| + |15-17| = 3$

Fitness: 0.0833

ch2				ch1		
4	9	2		2	7	6
3	5	7	$\longleftrightarrow$	9	5	1
8	1	6		4	3	8

Parent 1: Chromosome 2

Parent 2: Chromosome 1

Child 1:

$$\begin{matrix} 4 & 9 & 2 \\ 9 & 5 & 1 \\ 8 & 1 & 6 \end{matrix}$$

swap 9 and 3

Child 2:

$$\begin{matrix} 2 & 7 & 6 \\ 3 & 5 & 7 \\ 4 & 3 & 8 \end{matrix}$$

swap 7 and 9

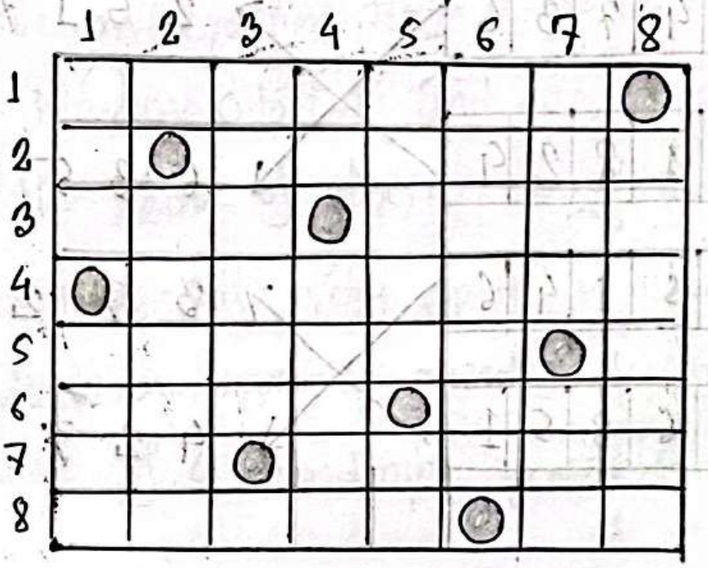
Mutation

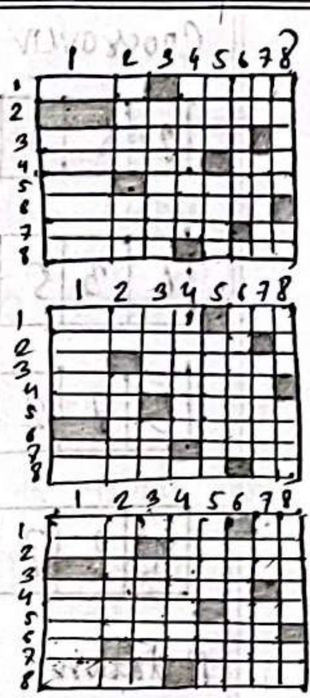
$$\begin{matrix} 4 & 3 & 2 \\ 9 & 5 & 1 \\ 8 & 1 & 6 \end{matrix} \qquad \begin{matrix} 2 & 9 & 6 \\ 3 & 5 & 7 \\ 4 & 3 & 8 \end{matrix}$$

Subject : | Date :

Problem 3:

8 Queen Problem





Representation:

[4, 2, 7, 3, 6, 8, 5, 1]

Here, index = column

Value = row

Chromosome 1 : [4, 2, 7, 3, 6, 8, 5, 1]

Chromosome 2 : [2, 5, 1, 8, 4, 7, 3, 6]

Chromosome 3 : [6, 3, 5, 7, 1, 8, 2, 4]

Chromosome 4 : [3, 7, 2, 8, 5, 1, 4, 6]

fitness:

$q_1 = 6$, $q_2 = 8$, $q_3 = 6$, $q_4 = 8$, $q_5 = 8$, $q_6 = 6$, $q_7 = 8$, $q_8 = 6$

total 56

$\Sigma_T = 240$

$\therefore$ fitness 23.33%

$q_1 = 8$, $q_2 = 8$, $q_3 = 6$, $q_4 = 8$, $q_5 = 6$, $q_6 = 6$, $q_7 = 8$, $q_8 = 6$

total 56

$\Sigma_T = 240$

$\therefore$ fitness 23.33%

$q_1 = 8$, $q_2 = 8$, $q_3 = 8$, $q_4 = 8$, $q_5 = 8$, $q_6 = 8$, $q_7 = 8$, $q_8 = 8$

total 64

$\Sigma T = 240$

$\therefore$ fitness: 35%

$q_1 = 8$, $q_2 = 8$, $q_3 = 8$, $q_4 = 8$, $q_5 = 8$, $q_6 = 8$, $q_7 = 8$, $q_8 = 8$

total 64

$\Sigma_T = 240$

Fitness: 35%

Subject : | Date :

Crossover :

2	5	1	8	4	7	3	6
6	3	5	7	1	8	2	4
3	7	2	8	5	1	4	6
4	2	7	3	6	8	5	1

2 5 1 7 1 8 2 4

6 3 5 8 4 7 3 6

3 7 2 3 6 8 5 1

4 2 7 8 5 1 4 6

Mutation :

2 5 1 8 1 8 2 4

6 3 5 8 4 4 3 6

3 7 2 1 6 8 5 1

4 2 1 8 8 1 4 6

Subject: Date:

Mid Question (Genetic Algorithm)

1. Suppose you have an equation $f(x) = x^2 - 5x + 6$. Assume x can be any number between 0 to 15. Find an appropiate value of x such that the value of $f(x) = 0$ using Genetic Algorithm

(a) Consider the fact that every population chromosome will have 4 genes. Illustrate an appropiate encoding system to create an initial population of 4 randomly generated chromosome.

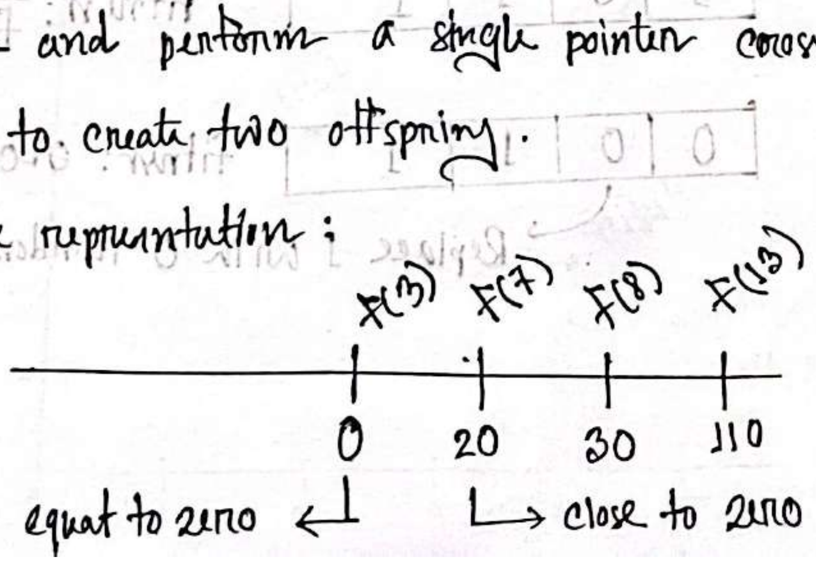
$$f(x) = x^2 - 5x + 6$$

BCD representation: (Encoding system)

8	4	2	1	
0	1	1	1	chromosome 1
0	0	1	1	chromosome 2
1	0	0	0	chromosome 3
1	1	0	1	chromosome 4

(b) Using an appropiate fitness function deduce the 2 fittest chromosome and perform a single pointer crossover from the middle to create two offspring.

Number Line representation:



Subject: Date:

parent 1 : fitness ∞

0	0	1	1

parent 2 : fitness 0.05

0	1	1	1

Fitness calculation

$$\text{fitness} = \frac{1}{f(x)}$$

for $f(3)$, fitness $= \infty$

for $f(7)$, fitness $= 0.05$

for $f(8)$, fitness $= 0.0333$

for $f(13)$, fitness $= 9.09 \times 10^{-3}$

Single point crossover:

Assume, point = 2

0	0	1	1
0	1	1	1

→

0	0	1	1
0	1	1	1

(c) Explain how you can mutate the offspring.

Mutation → Replace 0 with 1 randomly

1	0	1	1

fitness: $\frac{1}{f(11)} = 0.01388$

0	0	1	1

fitness: ~~0.05~~ ∞

→ Replace 1 with 0 randomly

Subject:	Date:

(d) Explain your opinion on whether Genetic Algorithm can be treated as a class of Local Search Algorithms or not.

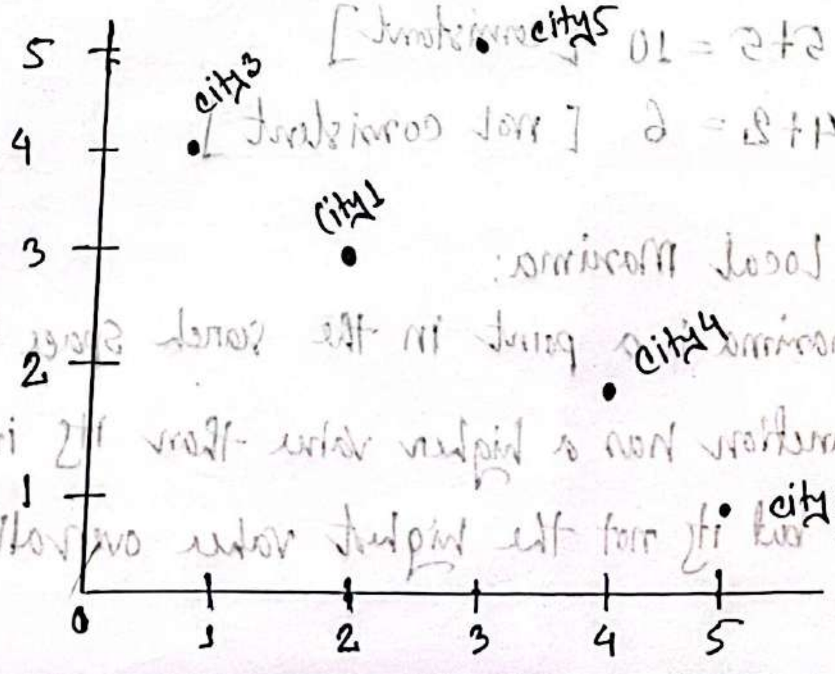
GA is better to be considered as global search algorithm for its diversity and mutation nature.

2. Ruhan is a sale person. His point is to find the optimal path that visit all cities at once and return to the starting point.

City 1	City 2	City 3	City 4	City 5
(2,3)	(5,1)	(1,4)	(4,2)	(3,5)

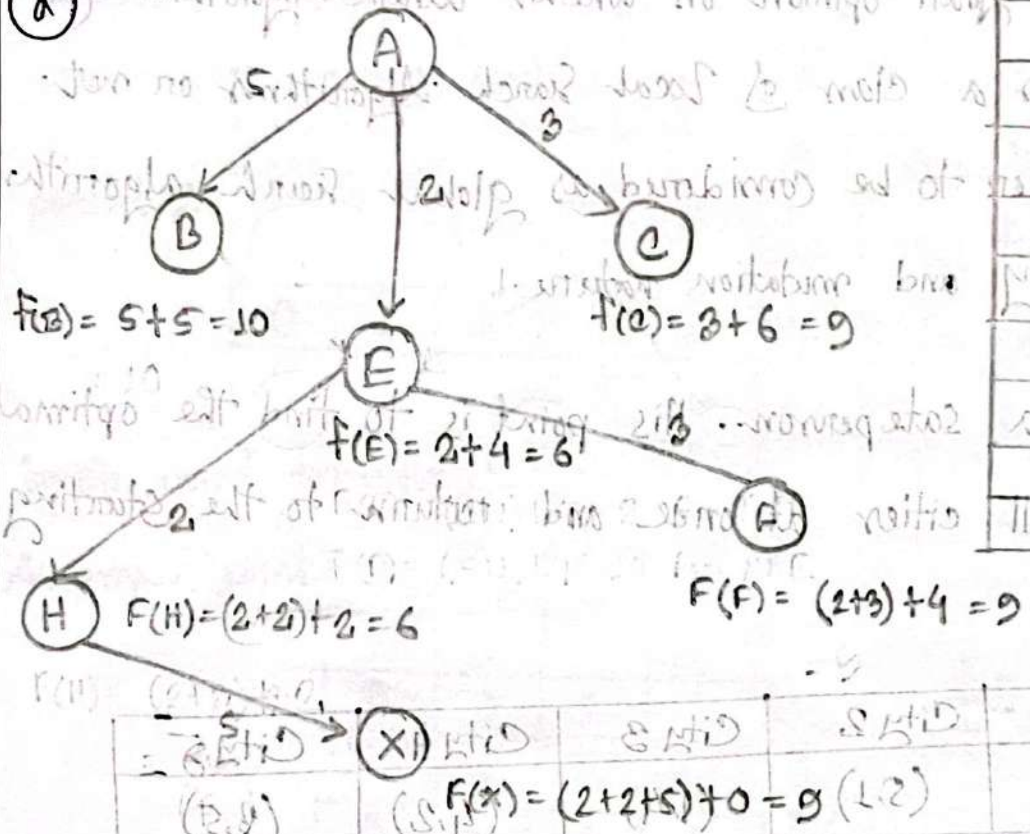
Use genetic algorithm to find solutions.

(a) Encode the problem and create four parent chromosomes. Then determine an appropiate fitness function and choose parents.



Subject: Date:

3. (a)



Node	h-value
A	7
B	5
C	6
D	3
E	4
F	4
G	2
H	2
X	0

(b) Check the Consistancy.

Hueristic value of Node 1 $\leq$ Hueristic value of node 2 + path cost

Node A:

$7 \leq 5+5 = 10$ [consistant]

$7 \leq 4+2 = 6$ [not consistent].

(c) Define Local Maxima:

A local maxima is a point in the search space where the objective function has a higher value than its immediate neighbours but its not the highest value overall.

Subject: | Date:

Example: Hill Climbing. it will stuck at local maxima but there will be a better solution elsewhere.

(d) Define local minima:

A local minima is a point where the objective function has a lower value than its imediate neighbour. but not the lowest at all.

Hill climbing: often stuck in local maxima

Simulated Annealing: Escape local maxima/minima

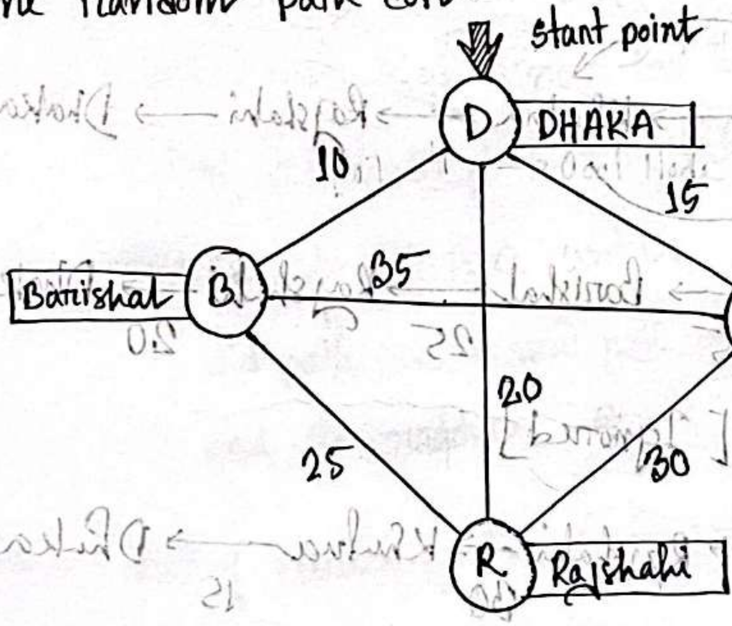
GA: Avoid local traps and find global optima

Subject: Date:

Question 2:

Travelling Salesman Problem is a path travelling problem where the algorithm finds that path which is the shortest among all other paths possible from start point to start point after visiting every other nodes exactly one time.

To solve the problem lets assume that we have 4 cities and they are interconnected as mesh topology structure with some random path cost.



Approch:

ⓐ Select any initial Route:

$$D \xrightarrow[10]{} B \xrightarrow[35]{} K \xrightarrow[30]{} R \xrightarrow[20]{} D$$

Total distance = 95 unit

Subject: Date:

(b) Generate Random path:

We will generate random path by swapping any two cities.

Then we will calculate the path cost.

If the total cost $\geq$ previous cost:

Ignore!

If the total cost $<$ previous cost:

Accepted!

previous path:

Dhaka ⟶ Barishal ⟶ Khulna ⟶ Rajshahi ⟶ Dhaka

Dhaka $\xrightarrow[15]{}$ Khulna $\xrightarrow[35]{}$ Barishal $\xrightarrow[25]{}$ Rajshahi $\xrightarrow[20]{}$ Dhaka

Total path cost: 95 [Ignored]

Dhaka $\xrightarrow[10]{}$ Barishal $\xrightarrow[25]{}$ Rajshahi $\xrightarrow[30]{}$ Khulna $\xrightarrow[15]{}$ Dhaka

Total path cost: 80 [Accepted]

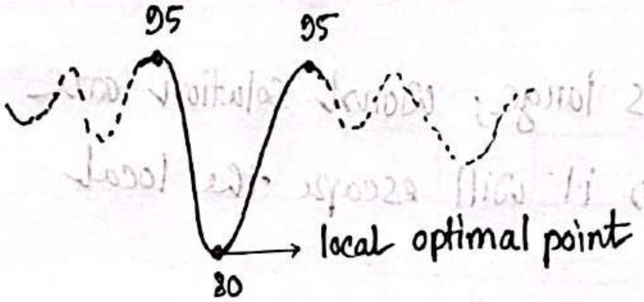
Dhaka $\xrightarrow[20]{}$ Rajshahi $\xrightarrow[25]{}$ Barishal $\xrightarrow[35]{}$ Khulna $\xrightarrow[15]{}$ Dhaka

Total path cost: 95 [Ignored and Stop searching]

Return: 80, Path: D ⟶ B ⟶ R ⟶ K ⟶ D

Subject: Date:

Visual Representation:



Question 3:

step A: Initialize a random path and find the path cost.

step B: Initialize a random temperature. The selected temprature neither too high nor too low.

step C: Define a cooling factor between 0 and 1.

step D: Repeat step A, B, C until:

old path cost − new pathcost < 0

and temperature value is very small.

Step E: Every iteration / Function call

compute $\Delta f = f(S_{new}) - f(S_{old})$

if $\Delta f < 0$; ACCEPTED

else:

ACCEPTED with propability $= e^{\frac{-\Delta E}{T}}$

$T = \alpha$ (cooling factor) $* T$

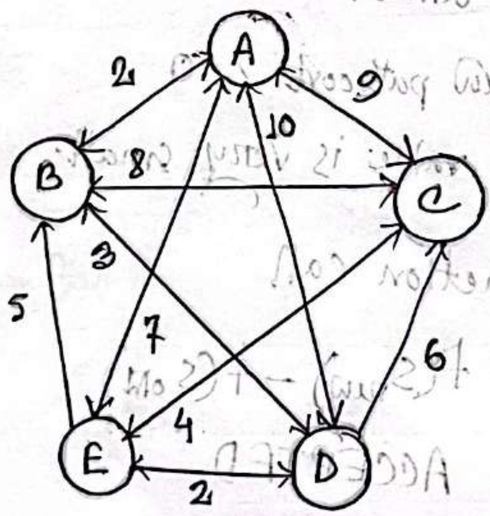
Subject: Date:

Step F: Every iteration the temperature (T) will decrease.
If probabity is large; worst solution are accepted means it will escape the local minima.

Step G: When The T is very low to make the next search, The program will stop finding.

Step H: Return the path with path cost.

Question 4:



Subject: | Date:

Iteration 1:

Initialization of chromosome:

Chromosome 1:

A	B	C	D	E	A

Chromosome 2:

A	C	B	E	D	A

Chromosome 3:

A	D	E	B	C	A

Chromosome 4:

A	E	D	C	B	A

Fitness Calculations:

For chromosome 1: $(2+8+6+2+7) = 25$

Fitness: $\frac{1}{25} = 0.04$ [Highest fitness]

For chromosome 2: $(9+8+5+2+10) = 34$

Fitness: $\frac{1}{34} = 0.0294$

For chromosome 3: $(10+2+5+8+9) = 34$

Fitness: $\frac{1}{34} = 0.0294$

For chromosome 4: $(7+2+6+8+2) = 25$

Fitness: $\frac{1}{25} = 0.04$ [Highest fitness]

Subject: Date:

Crossover:

chromosome 1: A → B → C → D → E → A

chromosome 2: A → E → D → C → B → A

Child 1: A → B → D → C → B → A

child 2: A → E → C → D → E → A

Mutation:

swap B with E for child 1

swap E with B for child 2

Child 1: A → B → D → C → E → A

Child 2:

A → E → C → D → B → A

Iteration 2:

Initialize chromosome:

chromosome 1:

A	B	D	C	E	A

chromosome 2:

A	E	C	D	B	A

Child 1 and child 2 in previous iteration

chromosome 3:

A	C	B	E	D	A

Chromosome 4:

A	D	E	B	C	A

Subject: | Date:

Fitness Calculations:

For chromosome 1: $(2+3+6+4+7) = 22$

fitness $= 1/22 = 0.04545$ [Highest fitness]

For chromosome 2: $(7+4+6+3+2) = 22$

Fitness $= 1/22 = 0.04545$ [Highest fitness]

For chromosome 3: $(9+8+5+2+10) = 34$

Fitness $= 0.029411$

For chromosome 4: $(10+2+5+8+9) = 34$

Fitness $= 1/34 = 0.029411$

Crossover:

chromosome 1: A ⟶ B ⟶ D ⟶ C ⟶ E ⟶ A

Chromosome 2: A ⟶ E ⟶ C ⟶ D ⟶ B ⟶ A

Child 1: A ⟶ B ⟶ D ⟶ D ⟶ B ⟶ A

Child 2: A ⟶ E ⟶ C ⟶ C ⟶ E ⟶ A

Mutation:

swap D with C for child 1

Swap C with B for child 2

Subject: Date:

Child 1: A ⟶ B ⟶ C ⟶ D ⟶ B ⟶ A

Child 2: A ⟶ E ⟶ B ⟶ C ⟶ E ⟶ A

Iteration 3:

Initialize chromosomes

Chromosome 1:

A	B	C	D	B	A

Chromosome 2:

A	E	B	C	E	A

Chromosome 3:

A	C	B	E	D	A

Chromosome 4:

A	D	E	B	C	A

Fitness Calculations:

For chromosome 1: $(2+8+6+3+2) = 21$

Fitness $= 1/21 = 0.04761$ [Highest fitness]

For chromosome 2: $(7+5+8+4+7) = 31$

Fitness $= 1/31 = 0.03225$ [Highest fitness]

For chromosome 3: $(9+8+5+2+10) = 34$

Fitness $= 1/34 = 0.029411$

For chromosome 4: $(10+2+5+8+9) = 34$

Fitness $= 1/34 = 0.029411$

Subject: | Date:

Crossover:

Chromosome 1: A ⟶ B ⟶ C ⟶ D | ⟶ B ⟶ A

Chromosome 2: A ⟶ E ⟶ B ⟶ C | ⟶ E ⟶ A

Child 1: A ⟶ B ⟶ C ⟶ D ⟶ E ⟶ A

Child 2: A ⟶ E ⟶ B ⟶ C ⟶ B ⟶ A

Mutations:

Swap B with D in child 1

swap B with D in child 2

Child 1: A ⟶ D ⟶ C ⟶ D ⟶ E ⟶ A

fitness $= 1/(10+6+6+2+7) = 0.032258$

Child 2: ~~A ⟶ E ⟶ B ⟶ C ⟶ D ⟶ A~~

~~fitness $= 1/(7+5+8+6+10)$~~

A ⟶ E ⟶ D ⟶ C ⟶ B ⟶ A

fitness: $\dfrac{1}{(7+2+6+8+2)} = 1/25 = 0.04$

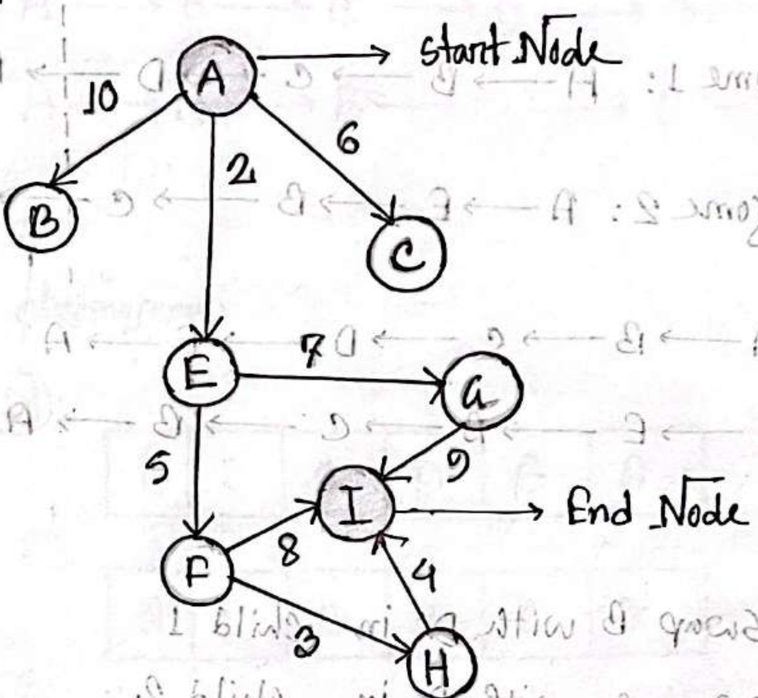
Final Path: A ⟶ E ⟶ D ⟶ C ⟶ B ⟶ A

Cost: 25, fitness: 0.04.

Subject :   Date :

Question 1 :

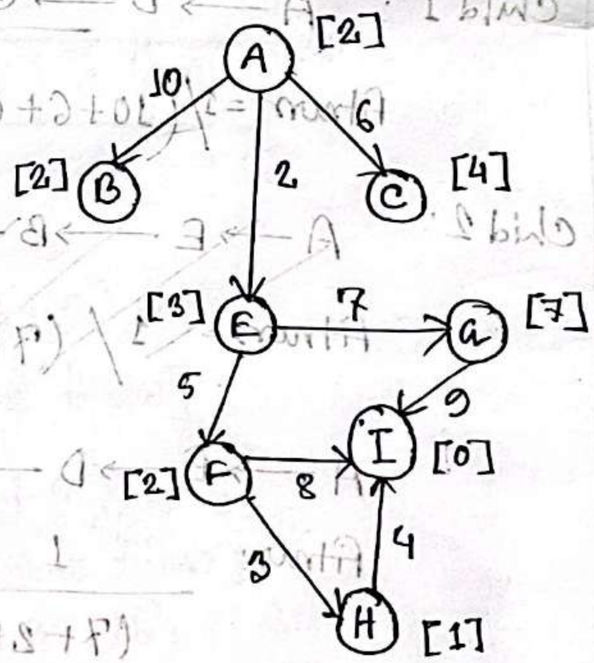
[A, B]



B.

C.

Node	Heuristic Value
A	2
B	2
C	4
~~B~~ I	0
E	3
F	2
G	7
H	1



Subject: Date:

D.

Start Node = A

End Node = I

Find the lowest / shortest path from A to I

path 1: $A \xrightarrow{2} E \xrightarrow{5} F \xrightarrow{8} I$, cost = 15

path 2: $A \xrightarrow{2} E \xrightarrow{7} G \xrightarrow{9} I$, cost = 18

path 3: $A \xrightarrow{2} E \xrightarrow{5} F \xrightarrow{3} H \xrightarrow{4} I$, cost = 14

∴ The shortest path: $A \longrightarrow E \longrightarrow F \longrightarrow H \longrightarrow I$

The graph is admissible if Node A, E, F, H and I are admissible simultaneously.

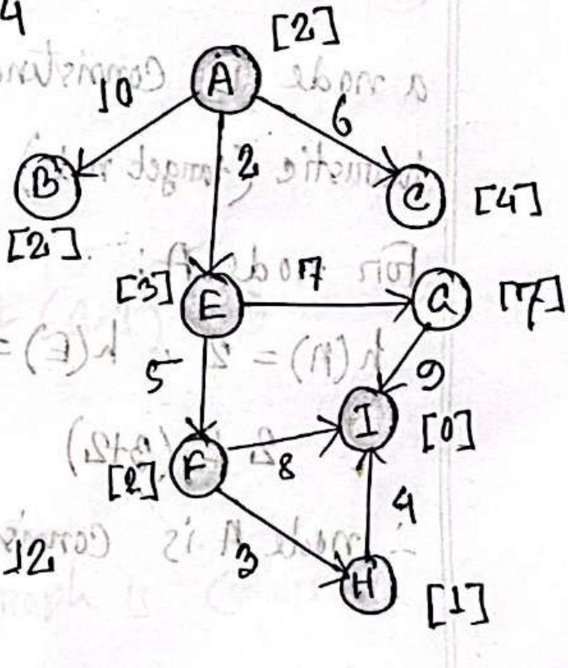
For checking admissibility,

Admissible if, heuristic value of $(n) \leq$ actual path cost from (n)

For Node A:

$h(A) = 2$ and $h^*(A) = (2+5+3+4) = 14$

∴ $h(A) \leq h^*(A)$



For Node B:

$h(B) = 2$ [edge node]

For Node C:

$h(C) = 4$ [edge node]

For Node E:

$h(E) = 3$ and $h^*(E) = (5+3+4) = 12$

∴ $h(E) \leq h^*(E)$

Subject: Date:

For Node F:

$h(F) = 2$ and $h^*(F) = (3+4) = 7$

$\therefore h(F) \leq h^*(F)$

For Node H:

$h(H) = 1$ and $h^*(H) = 4$

$\therefore h(H) \leq h^*(H)$

For Node I:

$h(I) = 0$ and $h^*(I) = 0$

$\therefore h(I) \leq h^*(I)$

For Node G:

$h(G) = 7$ and $h^*(G) = 9$

$\therefore h(G) \leq h^*(G)$

$\therefore$ The graph is admissible. [proved]

E. The shortest path from Node A to Node I is

$A \longrightarrow E \longrightarrow F \longrightarrow H \longrightarrow I$

a node is consistence if

heuristic (target node) $\leq$ heuristic (next node) + path cost (m, n)

For node A:

$h(A) = 2$, $h(E) = 3$, path cost $(A, E) = 2$

$2 \leq (3+2)$

$\therefore$ node A is consistance.

Subject: Date:

For node E:

$h(E) = 3$, $h(F) = 2$, path cost $(E, F) = 5$

$h(G) = 7$, path cost $(E, G) = 7$

$3 \leq (2+5)$ and $3 \leq (7+7)$

$\therefore$ node E is consistance.

For node F:

$h(F) = 2$, $h(I) = 0$, path cost $(F, I) = 8$

$h(H) = 1$, path cost $(F, H) = 3$

$2 \leq (0+8)$ and $2 \leq (1+3)$

$\therefore$ Node F is consistance.

For node H:

$h(H) = 1$, $h(I) = 0$, path cost $(I, H) = 4$

$1 \leq (0+4)$

$\therefore$ node H is consistance.

For node G:

$h(I) = 0$, $h(G) = 7$, path cost $(G, I) = 9$

$0 \leq (7+9)$

$\therefore$ node G is consistance.

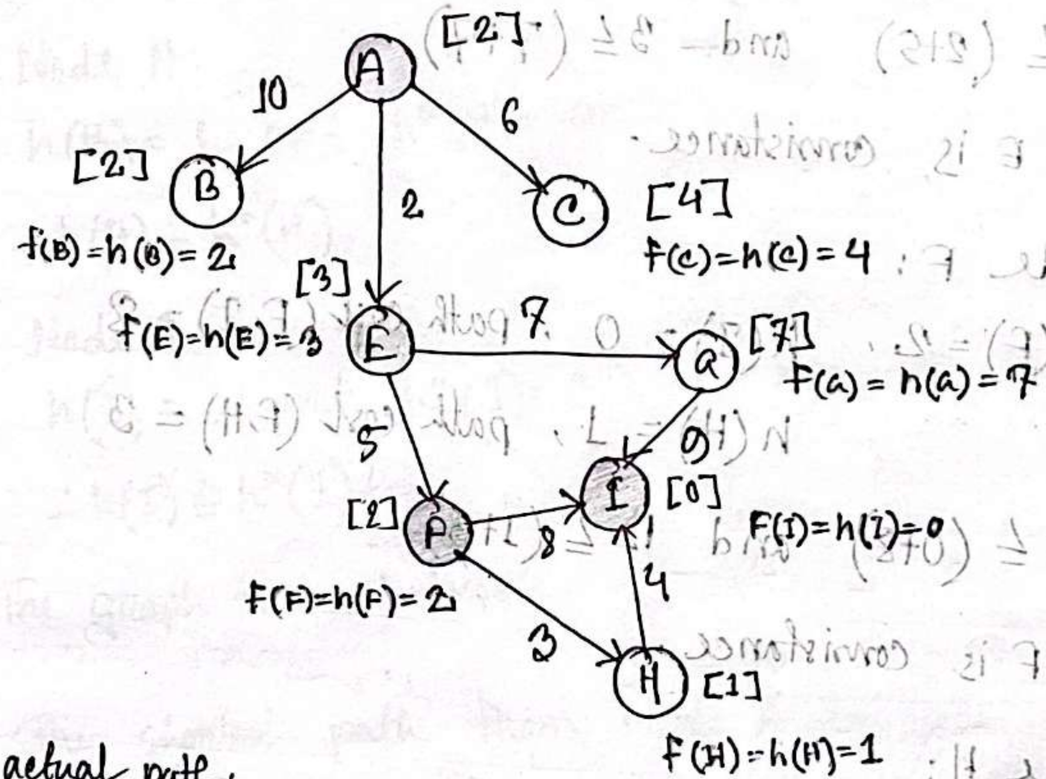
As all nodes are consistance. The graph is Consistant.

Subject:　　　　　　　　　　　　　　　　　　　　Date:

F.

Technique 1:

Greedy Best First Search



The actual path:

$A \xrightarrow{2} E \xrightarrow{5} F \xrightarrow{3} H \xrightarrow{4} I$

Total path cost = 14

GBFS path:

$A \xrightarrow{2} E \xrightarrow{5} F \xrightarrow{8} I$

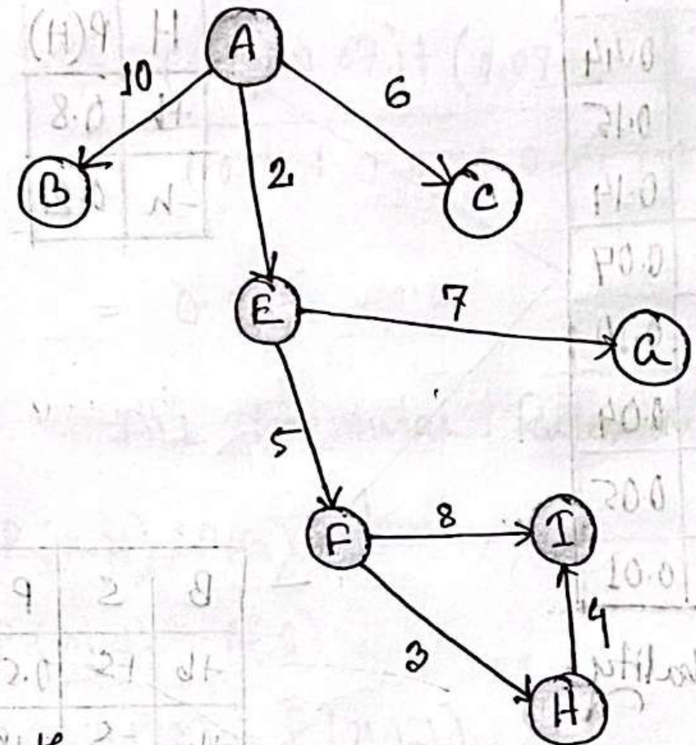
Total path cost = 15

Conclusion: Completeness but not optimal.

Subject: Date:

Technique 2:

Uniform Cost Search Algorithm



The actual path:

$A \xrightarrow{2} E \xrightarrow{5} F \xrightarrow{3} H \xrightarrow{4} I$

Total path cost: 14

UCS path:

$A \xrightarrow{2} E \xrightarrow{5} F \xrightarrow{3} H \xrightarrow{4} I$

Total path cost = 14

Conclusion: Complitness and optimal simultaneously.

Not always complete and optimal.

Subject : | Date :

Probability

H	B	S	P
+h	+b	+s	0.44
+h	+b	-s	0.15
+h	-b	+s	0.14
+h	-b	-s	0.07
-h	+b	+s	0.10
-h	+b	-s	0.04
-h	-b	+s	0.05
-h	-b	-s	0.01

H	P(H)
+h	0.8
-h	0.2

B	S	P
+b	+s	0.54
+b	-s	0.19
-b	+s	0.19
-b	-s	0.08

Conditional Probability

	+b		-b	
	+s	-s	+s	-s
+h	0.44	0.15	0.14	0.07
-h	0.10	0.04	0.05	0.01

ⓐ $P(+h|+b) = \frac{P(+h,+b)}{P(+b)} = \frac{0.44+0.15}{(0.44+0.15+0.1+0.04)} = 0.81$

ⓑ $P(-h,-s|+b) = \frac{P(-h,-s,+b)}{P(+b)} = \frac{0.04}{(0.44+0.15+0.1+0.04)} = 0.05$

ⓒ $P(-h|+s,-b) = \frac{P(-h,+s,-b)}{P(+s,-b)} = \frac{0.05}{(0.14+0.05)} = 0.26$

Subject : Date :

(d) $$P(+h \vee -s \mid -b) = \frac{\left(P(+h,-b) + P(-s,-b) - P(+h,-s,-b)\right)}{P(-b)}$$

$$= \frac{(0.14+0.07)+(0.07+0.01)-0.07}{(0.14+0.05+0.07+0.01)}$$

$$= 0.81$$

ALL Formulae, Rules and Principles

$P(x|y) = P(x,y) / P(y)$

$P(x,y) = P(x|y) * P(y)$

$P(x \vee y) = P(x) + P(y) - P(x \wedge y)$

$P(x,y) = P(x) * P(y)$

$P(x,y|z) = P(x|z) * P(y|z)$

$P(A') = 1 - P(A)$

$P(A|B) = P(A \cap B) / P(B)$

$P(A \cap B) = P(A) * P(B)$ if A and B are independent

$$P(A|B) = \frac{\left(P(B|A) * P(A)\right)}{P(B)}$$ (Bayes' Theorem)

Subject : | Date :

Marginal Probability:

	Male	Female	Total
GOT	0.16	0.24	0.40
TBBT	0.2	0.05	0.25
Others	0.1	0.25	0.35
Total	0.46	0.54	1

(a) What is the marginal probability of people watching TBBT?

$P(TBBT) = 0.25$

(b) What is the joint probability of a person being female and liking TBBT?

$P(F \wedge T) = 0.05$

(c) What is the probability of a person liking GOT given that person is male?

$$P(GOT \mid Male) = \frac{P(GOT \wedge Male)}{P(Male)} = \frac{0.16}{0.46} = 0.347$$

(2)

Subject: | Date:

Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Big| \quad P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$P(A \cap B) \rightarrow$ Join probability

$P(A) \rightarrow$ Marginal Probability

Hence. $P(A|B) \times P(B) = P(B|A) \times P(A)$

Then, $P(A|B) = \frac{\left(P(B|A) \times P(A)\right)}{P(B)}$ [Bayes' Theorem]

* Are Male viewers and GOT independent?

$P(M \cap GOT) = 0.16$

$P(A \cap B) = P(A) \times P(B)$

$P(M) = 0.46$

$P(GOT) = 0.40$

$P(M) \times P(GOT) = 0.46 \times 0.40 = 0.184$

Since $P(M \cap GOT) \neq P(M) \times P(GOT)$, so not independent.

	Smart		Not Smart	
	Study	Not Study	Study	Not Study
Prepared	0.432	0.16	0.084	0.008
Not Prepared	0.048	0.16	0.036	0.072

Is smart conditionally independent of prepared given study?

$$P(SM \wedge PRE \mid STU) = \frac{P(SM \wedge PRE \wedge STU)}{P(STU)} = \frac{0.432}{0.6} = 0.72$$

Subject : | Date :

Again,

$P(SM \wedge PrE \mid STU) = P(SM \mid STU) * P(PrE \mid STU)$

$P(SM \mid STU) = \dfrac{P(SM \wedge STU)}{P(STU)} = \dfrac{0.432 + 0.048}{0.6} = 0.8$

$P(PrE \mid STU) = \dfrac{P(PrE \wedge STU)}{P(STU)} = \dfrac{0.516}{0.6} = 0.85$

$\therefore P(SM \mid STU) * P(PrE \mid STU) = 0.8 \times 0.85 = 0.68 \neq 0.72$

$\therefore$ Not conditionally independent

A person is brought in front of a jury. The jury finds the defendant guilty in 98% of the cases in which he committed a crime and it finds the defendant not guilty 97% of the cases when the defendant has not committed a crime. Only 0.8% of the population has committed a crime.

If a random person is found guilty by the jury what is more likely, criminal or not?

Subject : | Date :

Bayes' Theorem

Naive Bayes

1

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Problem 1:

Bag A contains 3 red balls
4 black balls

Bag B contains 5 red balls
6 black balls

One ball is drawn at random from one of the bags and $\boxed{\text{its red.}}$ Given

Find the probability that it was from bag B.

Solution:

Bag A: 3 Red and 4 Black

Bag B: 5 Red and 6 Black

P (Bag B given that ball is red) = $P(B_B|Red)$

$P(B_A) = \frac{1}{2}$, $P(B_B) = \frac{1}{2}$

$P(Red|B_A) = \frac{3}{7}$, $P(Red|B_B) = \frac{5}{11}$

$$= \frac{P(B_B) \cdot P(Red|B_B)}{P(B_A) \cdot P(Red|B_A) + P(B_B) \cdot P(Red|B_B)}$$

$$= \frac{\frac{1}{2} \times \frac{5}{11}}{\left(\frac{1}{2} \times \frac{3}{7}\right) + \left(\frac{1}{2} \times \frac{5}{11}\right)}$$

$$= \frac{35}{68}$$

Subject: Date:

Problem 2:

Given three identical boxes A, B, C.

A has $\longrightarrow$ 2 gold coins

B has $\longrightarrow$ 2 silver coins

C has $\longrightarrow$ 1 gold and 1 silver coin.

What is the probability that the other coin in the box is also of gold if a person choose a box at random and take out a coin which is $\boxed{\text{gold?}}$ given

$P(B) = 1/3$, $P(A) = 1/3$, $P(C) = 1/3$

$P(G|B) = 0/2$, $P(G|A) = 2/2$, $P(G|C) = 1/2$

$$P(A|G) = \frac{P(A) \times P(G|A)}{P(A) \times P(G|A) + P(B) \times P(G|B) + P(C) \times P(G|C)}$$

$$= \frac{1/3 \times 2/2}{(1/3 \times 2/2) + (1/3 \times 0) + (1/3 \times 1/2)}$$

$$= \frac{2/6}{2/6 + 1/6}$$

$$= 1/3 \times 2$$

$$= 2/3$$

Subject : | Date : -

Problem 3:

Bag X → contains 1 White, 2 Red, 3 Green balls

Bag Y → contains 2 White, 3 Red, 1 Green balls

Bag Z → contains 3 White, 1 Red, 2 Green balls

Two balls were choosen from the bag and it was 1W and 1R. Find the probability that the balls so drawn came from the Bag Y.

Let E = Getting 1W and 1R

$$P(BY|E) = \frac{P(Y) \times P(E|Y)}{P(X) \times P(E|X) + P(Y) \times P(E|Y) + P(Z) \times P(E|Z)}$$

$$= \frac{\frac{1}{3} \times ({}^2C_1 \times {}^3C_1) / {}^6C_2}{\frac{1}{3} \times \frac{({}^1C_1 \times {}^2C_1)}{{}^6C_2} + \frac{1}{3} \frac{({}^2C_1 \times {}^3C_1)}{{}^6C_2} + \frac{1}{3} \frac{({}^3C_1 \times {}^1C_1)}{{}^6C_2}}$$

$$= 6/11$$

Subject : Date :

Problem 4:

Suppose that 5% of men and 0.25% of women have grey hair. A grey haired persons selected at random. What is the probability of this person being male? There are equal number of males and females.

$$P(a|M) = \frac{5\%}{100\%} \quad P(a|W) = \frac{0.25\%}{100\%}$$

$$P(M|a) = \frac{P(a|m) \cdot P(M)}{P(a|m) \cdot P(m) + P(a|W) \cdot P(W)}$$

$$= \frac{5/100 \times 1/2}{\left(5/100 \times 1/2\right) + \left(\frac{0.25}{100} \times 1/2\right)}$$

$$= \frac{1/40}{1/40 + \left(\frac{1}{800}\right)}$$

$$= 20/21$$

Subject: Date:

Problem 5:

In a certain college, 4% of the boys and 1% of girls are taller than 1.8m. 60% of the students are girls. If a student is selected at random and is found to be taller than 1.8m. What is the probability that the student is a girl?

E = taller than 1.8m

$$P(E/B) = \frac{4}{100}, \quad P(E/G) = \frac{1}{100}, \quad P(B) = \frac{40}{100}, \quad P(G) = \frac{60}{100}$$

$$P(G/E) = \frac{P(E|G) \cdot P(G)}{P(E|G) \cdot P(G) + P(E|B) \cdot P(B)}$$

$$= \frac{1/100 \times \frac{60}{100}}{\left(\frac{1}{100} \times \frac{60}{100}\right) + \left(\frac{4}{100} \times \frac{40}{100}\right)}$$

$$= 60/220$$

Subject: Date:

Problem 6:

If a machine is correctly setup, it produces 90% acceptable items. If it is incorrectly set up, it produces only 40% aceptable items. Past experience shows that 80% of the set ups are correctly done. After a setup the machine produces 2 acceptable items.

Find the probability that the machine is correctly setup.

A = getting 2 acceptable items

$$P(C|A) = \frac{P(A|C) \times P(C)}{P(A|C) \times P(C) + P(A|C') \times P(C')}$$

$P(C) = \frac{80}{100}$ $\quad P(C') = \left(1 - \frac{80}{100}\right)$

$= 0.8$ $\quad = 0.2$

$P(A|C) = \left(\frac{90}{100} \times \frac{90}{100}\right)$ $\quad P(A|C') = \left(\frac{40}{100} \times \frac{40}{100}\right)$

$= 0.81$ $\quad = 0.16$

$$P(C|A) = \frac{0.81 \times 0.8}{(0.81 \times 0.8) + (0.16 \times 0.2)}$$

$$= \frac{81}{85} = 0.95$$

Subject: Date:

Problem 7:

In a factory which manufactures bolts, machine A, B, C manufactures respectively 25%, 35%, 40% of the bolts output 5%, 4%, 2% are defective bolts. A bolt is drawn at random and found to be defective.

What is the probability that it is manufactured by the machine B?

$$P(B|D) = \frac{P(D|B) \times P(B)}{P(D|B) \times P(B) + P(D|A) \times P(A) + P(D|C) \times P(C)}$$

$P(A) = 1/3$, $P(B) = 1/3$, $P(C) = 1/3$

$P \quad = \frac{25}{100} \quad = \frac{35}{100} \quad = \frac{40}{100}$

$P(D|A) = 5/100$, $P(D|B) = 4/100$, $P(D|C) = 2/100$

$$P(B|D) = \frac{\left(4/100 \times \frac{35}{100}\right)}{\left(4/100 \times \frac{35}{100}\right) + \left(5/100 \times 25/100\right) + \left(2/100 \times \frac{40}{100}\right)}$$

$$= \frac{0.04 \times 0.35}{(0.04 \times 0.35) + (0.05 \times 0.25) + (0.02 \times 0.4)}$$

$= 0.405797$

Subject: Date:

Problem 8:

Companies B1, B2, B3 produces 30%, 45%, 25% of the cars respectively. It is known that 2%, 3% and 2% of these cars produced from are defected.

(a) What is the probability that a car purchased is defective?

$P(B_1) = \frac{30}{100}$, $P(B_2) = 45/100$, $P(B_3) = 25/100$

$P(D|B_1) = 2/100$, $P(D|B_2) = 3/100$, $P(D|B_3) = 2/100$

$$P(D) = P(D|B_1)P(B_1) + P(D|B_2)P(B_2) + P(D|B_3)P(B_3)$$

$$= \left(\frac{2}{100} \times \frac{30}{100}\right) + \left(\frac{3}{100} \times \frac{45}{100}\right) + \left(\frac{2}{100} \times \frac{25}{100}\right)$$

$$= (0.02 \times 0.3) + (0.03 \times 0.45) + (0.02 \times 0.25)$$

$$= 0.0245$$

(b) If a car purchased is found out to be defective, What is the probability that this car is produced by Company B1?

$$P(B1|D) = \frac{P(D|B1) \times P(B1)}{P(D|B1) \times P(B1) + P(D|B_2) \times P(B_2) + P(D|B_3) \times P(B_3)}$$

Subject: Date:

$P(B_1) = \frac{30}{100}$, $P(B_2) = \frac{45}{100}$, $P(B_3) = \frac{25}{100}$

$P(D|B_1) = \frac{2}{100}$, $P(D|B_2) = 3/100$ $P(D|B_3) = 2/100$

$$P(B_1|D) = \frac{2/100 \times \frac{30}{100}}{\left(2/100 \times \frac{30}{100}\right) + \left(\frac{3}{100} \times \frac{45}{100}\right) + \left(2/100 \times \frac{25}{100}\right)}$$

$$= \frac{0.02 \times 0.3}{(0.02 \times 0.3) + (0.03 \times 0.45) + (0.02 \times 0.25)}$$

$$= \frac{6 \times 10^{-3}}{6 \times 10^{-3} + 0.0135 + 5 \times 10^{-3}}$$

$$= \frac{6 \times 10^{-3}}{0.0245}$$

$$= 0.244897$$

Problem: 9

A lot of IC chips contains 2% defective chips. Each is tested before delivery. The tester itself is not totally reliable. Probability of tester says the chip is good when it is really good is 0.95 and the probability of tester says the chip is defective when it is actually defective is 0.94. If a tested device is indicated to be defected, what is the probability that it is actually defective.

Subject : Date :

$P(T|D) = \cancel{\frac{P(D|T) \times P(T)}{P(D|T) \times P(T)}}$

T = tester say good

T' = tester say defective

$P(D) = 2/100$; $P(D') = 1 - 2/100$

$P(T|D') = 0.95$ $P(\neg T|D) = 0.94$

$P(\neg T|\neg D) = 1 - P(T|D')$
$= 1 - 0.95$
$= 0.05$

$$P(D|T') = \frac{P(\neg T|D) \times P(D)}{P(\neg T|D) \times P(D) + P(\neg T|\neg D) \times P(\neg D)}$$

$$= \frac{0.94 \times 0.02}{(0.94 \times 0.02) + (0.05 \times 0.98)}$$

$$= 0.27.$$

Subject: Date:

Naive Bayes

Dataset

Day	Outlook	Temperature	Humidity	Windy	Play Tennis
1	Sunny	Hot	H	W	No
2	Sunny	Hot	H	S	No
3	Overcast	Hot	H	W	Yes
4	Rain	Mild	H	W	Yes
5	Rain	Cool	N	W	Yes
6	Rain	Cool	N	S	No
7	Overcast	Cool	N	S	Yes
8	Sunny	Mild	H	W	No
9	Sunny	Cool	N	W	Yes
10	Rain	Mild	N	W	Yes
11	Sunny	Mild	N	S	Yes
12	Overcast	Mild	H	S	Yes
13	Overcast	Hot	N	W	Yes
14	Rain	Mild	H	S	No

$P(\text{Play Tennis} = \text{Yes}) = \frac{9}{14} = 0.64$

$P(\text{Play Tennis} = \text{No}) = \frac{5}{14} = 0.36$

Outlook	Y	N
Sunny	2/9	3/5
Overcast	4/9	0
Rain	3/9	2/5

Hum.	Y	N
H	3/9	4/5
N	6/9	1/5

Windy	Y	N
W	6/9	2/5
S	3/9	3/5

Temp	Y	N
Hot	2/9	2/5
Mild	4/9	2/5
Cool	3/9	1/5

Subject: Date:

(Outlook = sunny, Temp = cool, Hum = high, Wind = Strong)

$$V_{NB} = \underset{v_j \in \{yes, No\}}{\operatorname{argmax}} \; P(v_j) \prod_i P(a_i \mid v_j)$$

$$= \underset{v_j \in \{yes, No\}}{\operatorname{argmax}} \; P(v_j)$$

V_{NB} (yes):

$P(yes)\, P(sunny \mid Y)\, P(cool \mid Y)\, P(high \mid Y)\, P(strong \mid Y)$

$= 0.64 \times 0.222 \times 0.3333 \times 0.3333 \times 0.3333$

$= 0.00526$

V_{NB} (No):

$P(No)\, P(sunny \mid N)\, P(cool \mid N)\, P(high \mid N)\, P(strong \mid N)$

$= 0.36 \times 0.6 \times 0.2 \times 0.8 \times 0.6$

$= 0.020736$

$$V_{NB}(Yes) = \frac{V_{NB}(Yes)}{V_{NB}(Yes) + V_{NB}(No)} = \frac{0.00526}{0.00526 + 0.020736}$$

$= 0.2023$

$$V_{NB}(No) = \frac{V_{NB}(No)}{V_{NB}(No) + V_{NB}(Yes)} = \frac{0.020736}{0.00526 + 0.020736}$$

$= 0.79766$

Subject: Date:

Dataset

No	Color	Legs	Heights	Smelly	Species
1	White	3	S	Y	M
2	Green	2	T	N	M
3	Green	3	S	Y	M
4	White	3	S	Y	M
5	Green	2	S	N	H
6	White	2	T	N	H
7	White	2	T	N	H
8	White	2	S	Y	H

$P(M) = 4/8 = 0.5$

$P(H) = 4/8 = 0.5$

Color	M	H
W	2/4	3/4
G	2/4	1/4

Legs	M	H
2	1/4	4/4
3	3/4	0/4

Height	M	H
T	3/4	2/4
S	1/4	2/4

Smell	M	H
Y	3/4	1/4
S	1/4	3/4

(Color = Green, Legs = 2, Height = Tall, Smelly = No)

$(M) = P(M) * P(G|M) * P(2|M) * P(Tall|M) * P(No|M)$

$= 0.5 \times 2/4 \times 1/4 \times 3/4 \times 1/4 = 0.0117$

$(H) = P(H) * P(G|H) * P(2|H) * P(Tall|H) * P(No|H)$

$= 0.5 \times 1/4 \times 1 \times 2/4 \times 3/4 = 0.047$

Subject :

Date :

Decision Tree

Day	Outlook	Temp	Humidity	Wind	Play Tennis
01	Sunny	Hot	HA	Weak	No
02	Sunny	Hot	HA	Strong	No
03	Overcast	Hot	HA	Weak	Yes
04	Rain	Mild	HA	Weak	Yes
05	Rain	Cold	NR	Weak	Yes
06	Rain	Cold	NR	Strong	No
07	Overcast	Cold	NR	Strong	Yes
08	Sunny	Mild	HA	Weak	No
09	Sunny	Cold	NR	Weak	Yes
10	Rain	Mild	NR	Weak	Yes
11	Sunny	Mild	NR	Strong	Yes
12	Overcast	Mild	HA	Strong	Yes
13	Overcast	Hot	NR	Weak	Yes
14	Rain	Mild	HA	Strong	No

Entropy : $\sum_{i=1}^{n} -P \log_2 P$

$E(\text{Desision}) = -P(\text{Yes}) \log_2 P(\text{Yes}) + \left(-P(\text{No}) \log_2 P(\text{No})\right)$

$= -\frac{9}{14} \log_2\left(\frac{9}{14}\right) + \left(-\frac{5}{14} \log_2\left(\frac{5}{14}\right)\right)$

$= 0.9402859$

Subject : Date :

[Outlook]

$E(\cancel{Yes} \mid \text{Sunny}) \Rightarrow$

$-P(\text{Yes} \mid \text{Sunny}) \log_2 (\text{Yes} \mid \text{Sunny}) - P(\text{No} \mid \text{Sunny}) \log_2 (\text{No} \mid \text{Sunny})$

$= -\frac{2}{5} \log_2 \left(\frac{2}{5}\right) - P\ \frac{3}{5} \log_2 \left(\frac{3}{5}\right)$

$= 0.971$

$E(\text{Overcast}) \Rightarrow$

$-P(\text{Yes} \mid \text{overcast}) \log_2 (\text{Yes} \mid \text{overcast}) - P(\text{No} \mid \text{overcast}) \log_2 (\text{No} \mid \text{overcast})$

$= -4/4 \log_2 (4/4) - 0/4 \log_2 (0/4)$

$= 0$

$E(\text{Rain}) \Rightarrow$

$-P(\text{Yes} \mid \text{Rain}) \log_2 (\text{Yes} \mid \text{Rain}) - P(\text{No} \mid \text{Rain}) \log_2 (\text{No} \mid \text{overcast})$

$= -3/5 \log_2 (3/5) - 2/5 \log_2 (2/5)$

$= 0.971.$

Information Gain: [Outlook]

$E(\text{Decision}) - P(\text{Sunny}) \times E(\text{Sunny}) - P(\text{Rain}) \times E(\text{Rain}) - P(OC) \times E(OC)$

$= 0.9402 - (5/14 \times 0.971) - (5/14 \times 0.971) - (4/14 \times 0)$

$= 0.246.$

Subject: Date:

[Temp]

E(Hot) ⇒

$-P(yes|Hot)\log_2(yes|Hot) - P(No|Hot) \cdot \log_2(No|Hot)$

$= -\frac{2}{4}\log_2\left(\frac{2}{4}\right) - \frac{2}{4}\log_2\left(\frac{2}{4}\right)$

$= 1.0$

E(Mild) ⇒

$-P(yes|Mild)\log_2(yes|Mild) - P(No|Mild)\log_2(No|Mild)$

$= -\frac{4}{6}\log_2\left(\frac{4}{6}\right) - \frac{2}{6}\log_2\left(\frac{2}{6}\right)$

$= -4/6\log_2(4/6) - 2/6\log_2(2/6)$

$= 0.918295$

E(Cold) ⇒

$-P(yes|Cold)\log_2(yes|Cold) - P(No|Cold)\log_2(No|Cold)$

$= -\frac{3}{4}\log_2\left(\frac{3}{4}\right) - \left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right)$

$= 0.811278$

Information gain:

$E(D) - P(Hot) \times E(Hot) - P(Mild) \times E(Mild) - P(Cold) \times E(Cold)$

$= 0.9402 - \left(\frac{4}{14} \times 1\right) - \left(6/14 \times 0.918\right) - \left(\frac{4}{14} \times 0.811278\right) = 0.0289$

Subject: Date:

[Humidity]

E(High) ⇒

$-P(\text{Yes}|\text{High})\log_2(\text{Yes}|\text{High}) - P(\text{No}|\text{High}).\log_2(\text{No}|\text{High})$

$= -\frac{3}{7}\log_2\left(\frac{3}{7}\right) - \left(\frac{4}{7}\right)\log_2\left(\frac{4}{7}\right)$

$= 0.985228136$

E(Normal) ⇒

$-P(\text{Yes}|\text{Normal})\log_2(\text{Yes}|\text{Normal}) - P(\text{No}|\text{Normal})\log_2(\text{No}|\text{Normal})$

$= -\frac{6}{7}\log_2\left(\frac{6}{7}\right) - \frac{1}{7}\log_2\left(\frac{1}{7}\right)$

$= 0.59167277$

Information Gain:

$E(D) - P(\text{High}) \times E(\text{High}) - P(\text{Normal}) \times E(\text{Normal})$

$= 0.94028 - \left(\frac{7}{14} \times 0.9852\right) - \left(\frac{7}{14} \times 0.59167\right)$

$= 0.1516$

Subject: Date:

[Wind]

E(Weak) ⇒

$-P(\text{Yes}|\text{Weak}) \log_2(\text{Yes}|\text{Weak}) - P(\text{No}|\text{Weak}) \log_2(\text{No}|\text{Weak})$

$= -\frac{6}{8} \log_2\left(\frac{6}{8}\right) - \frac{2}{8} \log_2\left(\frac{2}{8}\right)$

$= 0.8112781$

E(Strong) ⇒

$-P(\text{Yes}|\text{Strong}) \log_2(\text{Yes}|\text{Strong}) - P(\text{No}|\text{Strong}) \log_2(\text{No}|\text{Strong})$

$= -\frac{3}{6} \log_2\left(\frac{3}{6}\right) - \frac{3}{6} \log_2\left(\frac{3}{6}\right)$

$= 1.0$

Information Gain:

$E(D) - P(\text{Weak}) \times E(\text{Weak}) - P(\text{Strong}) \times E(\text{Strong})$

$= 0.9402859 - \left(\frac{8}{14} \times 0.81127\right) - \left(\frac{6}{14} \times 1\right)$

$= 0.0478.$

Subject: Date:

Summary:

$\boxed{\text{Gain(Outlook)} = 0.246}$ Gain (Humidity) = 0.1516

Gain (Temp) = 0.0289 Gain (Wind) = 0.0478

Outlook:

Sunny: 01, 02, 08, 09, 11 $\longrightarrow$ X

Overcast: 03, 07, 12, 13 $\longrightarrow$ Y

Rain: 04, 05, 06, 10, 14 $\longrightarrow$ Z

$\boxed{X}$ Outlook = Sunny

Day	Temp	Humidity	Wind	Play Tennis
D1	Hot	High	Weak	No
D2	Hot	High	Strong	No
D8	Mild	High	Weak	No
D9	Cool	Normal	Weak	Yes
D11	Mild	Normal	Strong	Yes

$\boxed{\text{Temp}}$

$$E(\text{Sunny}) = -P(\text{Yes})\log_2 P(\text{Yes}) - P(\text{No})\log_2 P(\text{No})$$

$$= -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right)$$

$$= 0.97$$

$$E(\text{Hot}) = -P(\text{Yes}|\text{Hot})\log_2 P(\text{Yes}|\text{Hot}) - P(\text{No}|\text{Hot})\log_2 P(\text{No}|\text{Hot})$$

$$= -\frac{0}{2}\log_2(0/2) - P(2/2)\log_2(2/2)$$

$$= 0$$

Subject : Date :

$E(Mild) \Rightarrow -P(yes|Mild)\log_2(yes|Mild) - P(No|Mild)\log_2(No|Mild)$

$= -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2})$

$= 1.0$

$E(Cold) \Rightarrow -P(yes|Cold)\log_2(yes|Cold) - P(No|Cold)\log_2(No|Cold)$

$= -\frac{1}{1}\log_2(\frac{1}{1}) - (\frac{0}{1})\log_2(\frac{0}{1})$

$= 0$

Information Gain:

$E(S) - P(Hot) \times E(Hot) - P(Mild) \times E(Mild) - P(Cold) \times E(Cold)$

$= 0.97 - 0 - 1 \times \frac{2}{5} - 0$

$= 0.570.$

Humidity

$E(High) \Rightarrow -P(yes|High)\log_2(yes|High) - P(No|High)\log_2(No|High)$

$= -\frac{0}{3}\log_2(\frac{0}{3}) - (\frac{3}{3})\log_2(\frac{3}{3})$

$= 0$

$E(Normal) = -P(yes|Normal)\log_2(yes|Normal) - P(No|Normal)\log_2(No|Norm$

$= -(\frac{2}{2})\log_2(\frac{2}{2}) - (\frac{0}{2})\log_2(\frac{0}{2})$

$= 0$

IG:

$E(S) - P(H) \times E(H) - P(N) \times E(N) = 0.97 - 0 - 0$

$= 0.97$

Subject: Date:

$E(Weak) \Rightarrow -P(Yes|Weak)\log_2(Yes|Weak) - P(No|Weak)\log_2(No|Weak)$

$= -\frac{1}{3}\log_2(\frac{1}{3}) - \frac{2}{3}\log_2(\frac{2}{3})$

$= 0.9182$

$E(Strong) \Rightarrow -P(Yes|Strong)\log_2(Yes|Strong) - P(No|Strong)\log_2(No|Strong)$

$= -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2})$

$= 1.0$

Information Gain:

$\boxed{\text{Wind}}$

$E(S) - P(W) \times E(W) - P(S) \times E(S)$

$0.971 - \frac{3}{5} \times 0.9182 - \frac{2}{5} \times 1$

$\Rightarrow 0.02008$

Summary:

Gain(Temp): 0.570

$\boxed{\text{Gain(Humidity): 0.97}}$

Gain(Wind) = 0.02008

$E(Hot) = 0$

$E(Mild) = -\frac{2}{3}\log_2(\frac{2}{3}) - \frac{1}{3}\log_2(\frac{1}{3})$

$= 0.91829$

$E(Cold) = -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2})$

$= 1.0$

$E(Hot) = 0.0$

$\boxed{\text{R}}$ outlook = Rain

Day	Temp	Humidity	Wind	PlayTennis
D4	Mild	High	Weak	Yes
D5	Cold	Normal	Weak	Yes
D6	Cold	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No

$E(Rain) = -\frac{3}{5}\log_2(\frac{3}{5}) - \frac{2}{5}\log_2(\frac{2}{5})$

$= 0.970950$

$E(High) = -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2}) = 1.0$

$E(Normal) = -\frac{2}{3}\log_2(\frac{2}{3}) - \frac{1}{3}\log_2(\frac{1}{3}) = 0.9182$

Subject: | Date:

Information Gain: [Temp]

E(Rain) − P(Hot) × E(Hot) − P(Mild) × E(Mild) − P(Cold) × E(Cold)

$= 0.970 - 0 - (3/5 \times 0.918) - (2/5 \times 1.0)$

$= 0.0192$

Information Gain: [Humidity]

E(Rain) − P(High) × E(High) − P(Normal) × E(Normal)

$= 0.970 - (2/5 \times 1.0) - (3/5 \times 0.9182)$

$= 0.0192$

$E(\text{Weak}) = -3/3 \log_2(3/3) - (0/3 \log_2(0/3)) = 0.0$

$E(\text{Strong}) = -0/2 \log_2(0/2) - 2/2 \log_2(2/2) = 0.0$

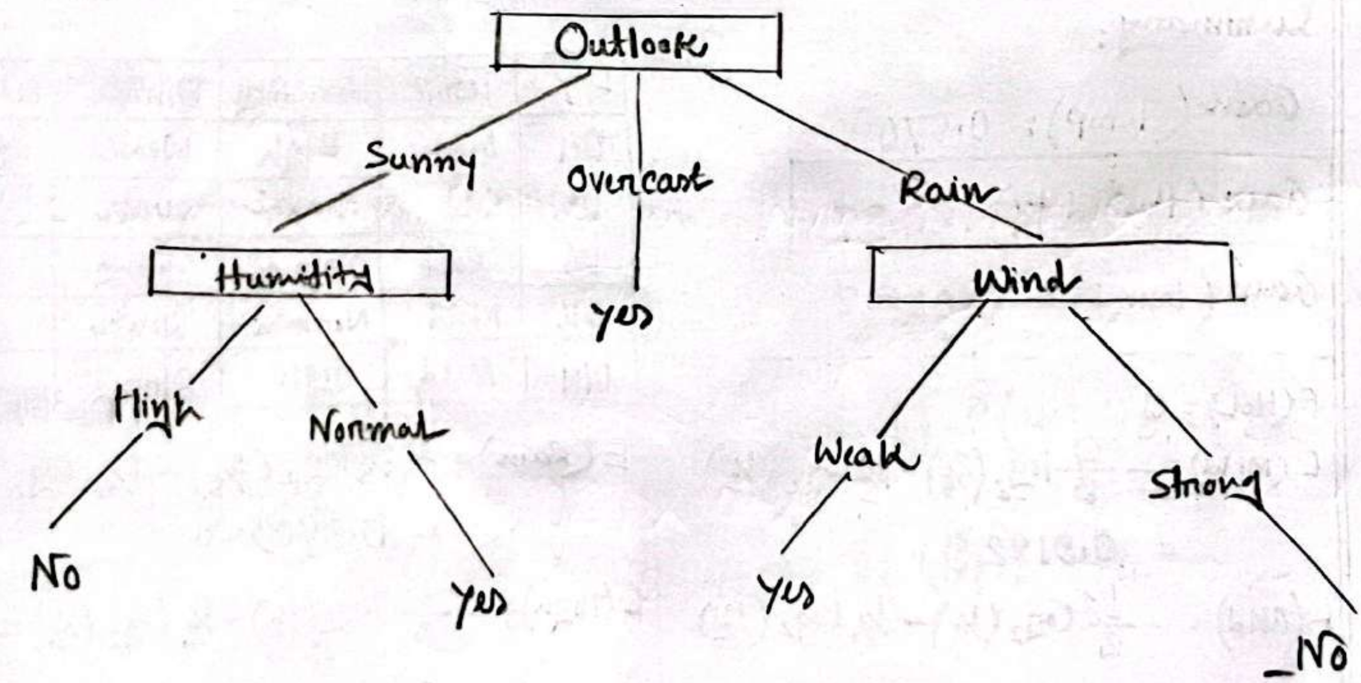
Temp =	0.0192
Humidity =	0.0192
Wind =	0.97

Information Gain: [Wind]

E(Rain) − P(Weak) × E(Weak) − P(Strong) × E(Strong)

$= 0.97 - 0 - 0$

$= 0.97$



Subject : Date :

Problems

(a) A patient went to the hospital for a malaria test. The doctor informed him that their test can successfully diagnose malaria positive given that patient is actually malaria positive 94% of the time.

Also, the probability of having no malaria and getting a malaria negative test result is 4%.

Meanwhile, 27% of people in general who come for tests are malaria positive. Now if the patient is already diagonosed malaria negative, then calculate the probability of the patient actually being malaria negative.

Given that,

$P(\text{Positive test} \mid \text{Actually positive}) = 0.94$

$P(\text{Negative test} \mid \text{Actually negative}) = 0.04$

$P(\text{Actually positive}) = 0.27$

$P(\text{Actually negative}) = 1 - 0.27 = 0.73$

$P(\text{Actually negative} \mid \text{Negative test})$

$$= \frac{P(\text{Negative test} \mid \text{Actually negative}) \times P(\text{Actually negative})}{P(\text{Negative test})}$$

Hamid Book Bainding

Subject: Date:

$$P(\text{Negative test}) = P(\text{Negative test} \mid \text{Actually negative}) \times P(\text{Actually neg})$$
$$+ P(\text{Negative test} \mid \text{Actually positive}) \times P(\text{Actually positive})$$

$$P(\text{Negative test} \mid \text{Actually positive}) = \cancel{1-0.04} \quad 1-0.94$$
$$= \cancel{0.96} \quad 0.06$$

$$\therefore P(\text{Actually negative} \mid \text{Negative test})$$

$$= \frac{(0.04 \times 0.73)}{(0.04 \times 0.73) + (0.06 \times 0.27)}$$

$$= \frac{0.0292}{0.0292 + 0.0162}$$

$$= 0.64317$$

(b) Is being malaria positive and having a positive test result independent of each other?

Assume,
A = Malaria positive
B = Malaria positive test result

We know, $P(A \cap B) = P(A) \times P(B)$
$$= 0.27 \times 0.94$$
$$= 0.2538$$

Given, $P(A) = 0.27$.

Subject: | Date:

$P(B) = P(\text{positive test})$.

$= P(B|A) \times P(A) + P(B|A') \times P(A')$

$= P(\text{positive test} \mid \text{Actually positive}) \times P(\text{Actually positive})$

$+ P(\text{positive test} \mid \text{Actually negative}) \times P(\text{Actually negative})$

$= (0.94 \times 0.27) + ([1-0.04] \times 0.73)$

$= 0.2538 + \cancel{0.0438}\ 0.7008$

$= \cancel{0.2976}\ 0.9546$

$P(A \cap B) = 0.2538$

$P(A) = 0.27$

$P(B) = 0.9546$

$P(A) \times P(B) = 0.27 \times 0.9546$

$= 0.257742$

$P(A \cap B) \neq P(A) \times P(B)$

$\therefore$ They are not independent.

Subject: Date:

Covid-19 test all over the World aren't 100% accurate. A patient is actually positive in 85% of the cases When the test comes out to be positive. A person is actually positive in 10% of the cases When the test comes out to be negative.

ⓐ Of all the people Who tested for Covid-19, 70% of them actually had the disease. If 1000 people participated in the tests, calculate the probability of a person's test results being positive.

$P(\text{Actually positive} \mid \text{positive test}) = 0.85$

$P(\text{Actually positive} \mid \text{negative test}) = 0.10$

$P(\text{Actually positive}) = 0.70$

Total people $= 1000$

$P(\text{Actually Positive} \mid \text{positive test})$

$$= \frac{P(\text{positive test} \mid \text{Actually Positive}) \times P(\text{Actually Positive})}{P(\text{positive test})}$$

$$= \frac{\overset{x}{\cancel{0.85}} \times 0.70}{P(\text{positive test})} = 0.85 = \frac{0.70x}{P(\text{positive test})}$$

Subject: Date:

$$P(\text{positive test}) = \frac{0.70x}{0.85} \quad \text{--- (b)}$$

$$P(\text{Actually positive} \mid \text{negative test}) = 0.10$$

$$0.10 = \frac{P(\text{negative test} \mid \text{Actually positive}) \times P(\text{Actually positive})}{P(\text{negative test})}$$

$$0.10 = \frac{(1-x) \times 0.70}{1 - P(\text{positive test})}$$

$$0.10 = \frac{(1-x) \times 0.70}{1 - \frac{0.70x}{0.85}} \quad \text{or,} \quad 0.10\left(1 - \frac{0.70x}{0.85}\right) = (1-x) \times 0.70$$

or, $0.085 - 0.70x = 0.595 - 0.595x$

or, $x = \frac{0.51}{0.525} \approx 0.9714$

$\therefore P(\text{positive test} \mid \text{Actually positive}) = 0.9714$

$$P(\text{positive test}) = \frac{0.70 \times (0.9714)}{0.85}$$

$$= 0.8$$

$\therefore P(\text{positive test}) = 0.80$

total population = 1000

$\therefore$ people tested positive = (1000×0.80)

$= 800$ people.

Subject: Date:

Consider two medical tests, A and B, for a virus. Test A is 95% effective at recognizing the virus when it is present, but has a 10% false positive rate (including that the virus is present, when it is not). Test B is 90% effective at recognizing the virus, but has a 5% false positive rate. The two test use independent methods of identifying the virus. The virus is carried by 1% of all the people. Say that a person is tested for the virus using only one of the tests. And the test comes back positive for carrying the virus.

Which test returning positive, is more indicative of someone really carrying the virus?

$$P(\text{Virus} \mid \text{Positive}) = \frac{P(\text{positive} \mid \text{virus})\, P(\text{virus})}{P(\text{positive})}$$

$$P(\text{positive}) = P(\text{positive} \mid \text{virus})\, P(\text{virus}) + P(\text{positive} \mid \text{No virus})\, P(\text{No virus})$$

For test A:

$$\frac{0.95 \times 0.01}{(0.95 \times 0.01) + (0.1 \times 0.99)} = 0.0876$$

For test B:

$$\frac{0.90 \times 0.01}{(0.90 \times 0.01) + (0.05 \times 0.99)} = 0.1538$$

Subject: | Date:

Linear Regression

$y = a_0 + a_1 x$
$y = b + mx$] same

Linear Regression equation is given by $y = a_0 + a_1 x + \text{error}$

where,

$$a_1 = \frac{(\overline{xy}) - (\bar{x})(\bar{y})}{\overline{x^2} - (\bar{x})^2}$$

$$a_0 = \bar{y} - a_1 \bar{x}$$

x_i (week)	y_i (Sales in Thousands)
1	1.2
2	1.8
3	2.6
4	3.2
5	3.8

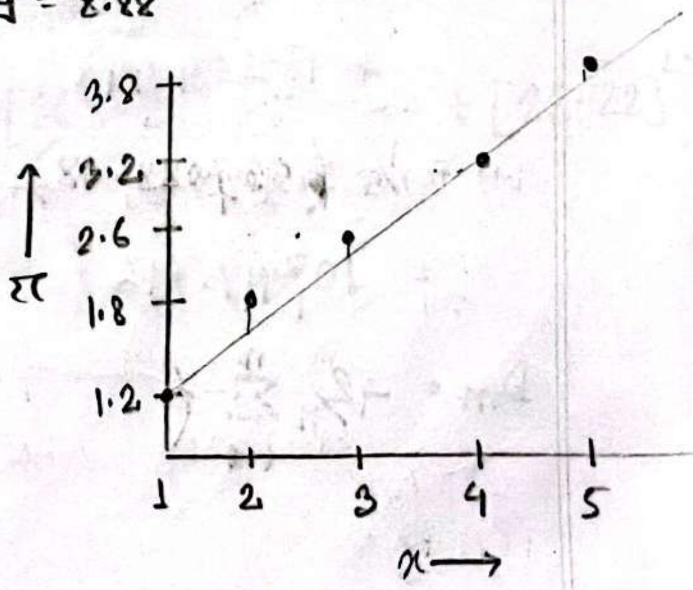
	x_i (week)	y_i (Sales in Thousands)	x_i^2	$x_i y_i$	Prediction
	1	1.2	1	1.2	1.2
	2	1.8	4	3.6	1.86
	3	2.6	9	7.8	2.52
	4	3.2	16	12.8	3.18
	5	3.8	25	19	3.84
sum	15	12.6	55	44.4	
Avg	3	2.52	11	8.88	

Now, $\bar{x} = 3$, $\bar{y} = 2.52$, $\overline{x_i^2} = 11$, $\overline{xy} = 8.88$

$$a_1 = \frac{(8.88) - (3 \times 2.52)}{(11) - (3)^2} = 0.66$$

$$a_0 = 2.52 - (0.66 \times 3) = 0.54$$

$$\boxed{\begin{aligned} \therefore y &= a_0 + a_1 x \\ &= 0.54 + 0.66x \end{aligned}}$$



For, $x = 7$

$y = 0.54 + 0.66 \times 7$

$= 5.16$

Subject : Date :

Linear Regression Gradient Descent

Loss function $= \sum (y_{actual} - y_{pred})^2$

Mean Square Error (MSE) $= \frac{1}{n} \sum_{i=0}^{n} (y_i - \hat{y}_i)^2$

x_i	y_i
1	1.2
2	1.8
3	2.6
4	3.2
5	3.8

Lets assume that, $\boxed{y = 10x + 300}$

in $y = mx + b$,

$m = 10$, $b = 300$, learning Rate $= 0.001$

$$MSE = \frac{1}{n} \sum_{i=0}^{n} (y_i - \hat{y}_i)^2 \quad [n = \text{total row}]$$

$$= \frac{1}{n} \sum_{i=0}^{n} (y_i - [mx + b])^2$$

$$MSE = \frac{1}{5} \sum (1.2 - 310)^2 + (1.8 - 320)^2 + (2.6 - 330)^2 + (3.2 - 340)^2 + (3.8 - 350)^2$$

$$= \frac{1}{5} \sum (95357.4 + 101251.24 + 107190.76 + 113434.24 + 119854.44)$$

$$= \frac{1}{5} (537088.08)$$

$$= 107417.616$$

$$D_m = -\frac{2}{n} \sum_{i=0}^{n} ($$

Subject: | Date:

Problem Solving

$h_\theta(X) = \theta_0 + \theta_1 x$

Assume that, $\theta_0 = 1$ and $\theta_1 = 1.5$

Cost Function:

$$\frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)^2$$

Gradient Decent

$$\frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)^2$$

For, θ_0

$$\theta_0 = \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)$$

$$\theta_1 = \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right) \times x_i$$

X	Y	Pred-Y
7	11	11.5
2	4	4
17	25	26.5
9	15	14.5
4	7	7
11	16.5	17.5
12	19	19
6	10.2	10
1	2.3	2.5
3	5.1	5.5
2.5	4	4.75
3.8	7	6.7
5.6	11	9.4
14	22	22

Cost Function

$$\frac{1}{2 \times (14)} \sum \left([11.5-11]^2 + [4-4]^2 + [26.5-25]^2 + \cdots + [22-22]^2 \right)$$

$$\frac{1}{28} \times 7.2025$$

$$= 0.25366071$$

Subject: Date:

1st Iteration

Assume that $\theta_0 = 1$ and $\theta_1 = 1$

$$\theta_0 = \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)^2$$

$$\theta_0 = \theta_0 - \alpha \times \left[\frac{1}{14}\left((8-11) + (3-4) + (18-25) \cdots\cdots + (15-22)\right)\right]$$

$$\theta_0 = \theta_0 - \alpha \times \left[\frac{1}{14} \times -47.2\right]$$

$$\theta_0 = 1 - 0.01 \times (-3.3714)$$

$$= 1.0337$$

$$\theta_1 = \theta_1 - \frac{\alpha}{m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right) x_i$$

$$= 1 - \frac{0.01}{14}\left[7(8-11) + (3-4)\times 2 + (18-25)\times 17 \cdots\cdots\right]$$

$$= 1 - 0.01 \times (-33.6821)$$

$$= 1.336$$

x	y	Pred-y
7	11	8
2	4	3
17	25	18
9	15	10
4	7	5
11	16.5	12
12	19	13
6	10.2	7
1	2.3	2
3	5.1	4
2.5	4	3.5
3.8	7	4.8
5.6	11	6.6
14	22	15

Subject: Date:

~~Neural Network~~

Probability Problem

1. In the table below, you are given a dataset containing 9 rows and 3 feature. Using naive bayes determine the most likely value of Y if $X_1 = 1, X_2 = a, X_3 = q$.

X_1	X_2	X_3	Y
1	a	p	0
2	b	r	1
3	b	p	1
3	c	q	1
2	c	r	0
1	b	q	1
2	a	p	0
3	a	r	1
3	b	q	0

$P(Y=0 \mid X_1 = 1, X_2 = a, X_3 = q)$

For $Y=0$, rows = 0, 4, 6, 8

For $Y=1$, rows = 1, 2, 3, 5, 7

$P(X_1 = 1 \mid Y=0) = \frac{1}{4}$

$P(X_2 = a \mid Y=0) = \frac{2}{4}$

$P(X_3 = q \mid Y=0) = \frac{1}{4}$

$P(Y=0) = 4/9$

$P(Y=1) = \frac{5}{9}$

Subject: Date:

$P(Y=0 \mid X_1=1, X_2=a, X_3=q) = \frac{4}{9} \times \frac{5}{8} \times \frac{1}{4} \times \frac{2}{4} \times \frac{1}{4}$

$= \frac{1}{72} \approx 0.0138$

Again,

$P(X_1=1 \mid Y=1) = \frac{1}{5}$ $P(X_2=a \mid Y=1) = \frac{1}{5}$

$P(X_2=q \mid Y=1) = \frac{2}{5}$

$\therefore P(Y=1 \mid X_1=1, X_2=a, X_3=q) = \frac{5}{9} \times \frac{1}{5} \times \frac{2}{5} \times \frac{1}{5}$

$= \frac{2}{225} \approx 8.889 \times 10^{-3}$

$0.0138 > 8.889 \times 10^{-3}$

$\therefore Y=0$ [Ans]

2. $X = \{A, B, C\}$. Assume $P(A) = 0.5$, $P(B) = 0.3$.

Determine $P(C)$ and $P(A \cup B)$

We know that,

$P(A) + P(B) + P(C) = 1$

or, $0.5 + 0.3 + P(C) = 1$

or, $P(C) = 0.2$.

$P(A \cup B) = P(A) + P(B) = 0.5 + 0.3 = 0.8$

Subject:	Date:

3. Assume 2 coins are tossed simultaneously.

Event A = The 1st coin comming up head.

Event B = The 2nd coin Comming up tail.

What is $P(A \cap B)$?

Sample Space, S = {HH, HT, TH, TT}

When the first coin is head, {HH, HT}

When the 2nd coin is tail, {HT, TT}

$\therefore P(A \cap B) = \{HT\}$ | Total space = 4, output = 1

$\therefore P(A \cap B) = \frac{1}{4} = 0.025$.

4. (i) Assume Event A & B are absolute independence and $P(B|A) = 0.5$. What is the value of x?

	A		A'	
	B	B'	B	B
C	0.1	0.2	0.2	Y
C'	X	0.1	0.1	0.

$P(B|A) = \frac{P(A \cap B)}{P(A)} \cdot \frac{P(B \cap A)}{P(A)} = 0.5$

Now,

$\frac{P(A|B)}{P(A)} = 0.5$

From the table, $P(B \cap A) = 0.1$

$\therefore P(A) = \frac{0.1}{0.5} = 0.2$

$P(A) = P(B) + P(B')$

$0.2 = 0.1 + X + 0.2 + 0.1$

$\therefore X = -0.2$

Subject: Date:

given that.

$P(B|A) = 0.5$

$P(A\cap B) = P(C\cap A\cap B) + P(C'\cap A\cap B)$

$= 0.1 + X$

$P(A) = P(C\cap A\cap B') + P(C\cap A\cap B') + P(C'\cap A\cap B) + P(C'\cap A\cap B')$

$= 0.1 + 0.2 + X + 0.1$

$= 0.4 + X$

$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{0.1 + X}{0.4 + X} = 0.5$

$0.5 = \frac{0.1 + X}{0.4 + X}$

or, $0.2 + 0.5X = 0.1 + X$

or, $0.2 - 0.1 = X - 0.5X$

or. $0.1 = 0.5X$

or, $X = 0.2$

(b) using $X = 0.2$ determine, Y.

$0.1 + 0.2 + 0.2 + Y + 0.2 + 0.1 + 0.1 + 0.1 = 1$

or, $1 + Y = 1$

or, $Y = 0$

$\therefore y = 0$

Subject :	Date :

ⓒ Determine $P(A|B\cap C)$

$$P(A|B\cap C) = \frac{P(A\cap B\cap C)}{P(B\cap C)}$$

Here,
$P(A\cap B\cap C) = 0.1$
$P(B\cap C) = 0.1+0.2$
$= 0.3$

$$= \frac{0.1}{0.3} = 0.333$$

	Pandamic		No Pandamic		Total
	Online	Offline	Online	Offline	
Public University	0.142	0.037	0.165	0.072	
Private University	0.102	0.145	0.217	0.118	
~~Private University~~					
Total					1

ⓓ Is pandamic Conditionality Independent of Public University Given that Online Class?

For Conditional Independence:

$$P(A|B,C) = P(A|C)$$

$\underbrace{}_{\text{given that part}}$

$$P(A|B,C) = \frac{P(A\cap B\cap C)}{P(B\cap C)}$$

$$P(A|C) = \frac{P(A\cap C)}{P(C)}$$

$$P(\text{Pand}|\text{Pub, Online}) = \frac{P(\text{Pand}\cap\text{Pub}\cap\text{online})}{P(\text{Pub}\cap\text{online})}$$

$$= \frac{0.142}{0.142+0.165}$$

$$= 0.46254.$$

$\boxed{\text{Not conditionally Independent.}}$

~~P(Pub~~ $P(\text{Pand}|\text{online}) = \frac{P(\text{Pand}\cap\text{Online})}{P(\text{online})} = \frac{0.142+0.103}{0.627}$

$= 0.391$

Subject : Date :

(b) Is Private Uni Independent of Online Class?

$$P(\text{Priv} \cap \text{Online}) = P(\text{Priv}) \cdot P(\text{online})$$

$P(\text{Private} \cap \text{online}) = 0.103 + 0.217 = 0.32$ —— (a)

$P(\text{Priv}) = 0.103 + 0.146 + 0.217 + 0.118 = 0.584$

$P(\text{Online}) = 0.142 + 0.103 + 0.165 + 0.217 = 0.627$

$P(\text{Private}) \cdot P(\text{online}) = 0.584 \times 0.627$

$= 0.366168$ —— (b)

Not Independent

(c) Find the marginal probability of offline Class

$P(\text{Public} \Rightarrow (0.037 + 0.146 + 0.072 + 0.118)$

$= 0.373$

Subject: Date:

ⓐ $h_w(x) = w_1 x + w_0$

here compare the equation with $y = mx + c$

$m = w_1$ and $c = w_0$

w_1 = slope of the line.

w_0 = the intercept point.

ⓑ given that,

$$Loss(h_w) = \sum_{j=1}^{N} (y_i - (w_1 x_j + w_0))^2$$

$$\frac{\partial Loss}{\partial w_1} = -2 \sum x_j (y_j - h_w(x_j))$$

$$\frac{\partial Loss}{\partial w_0} = -2 \sum (y_i - h_w(x_j))$$

Initially $w_1 = 15$, $w_0 = 0$

Iteration	h(w)	loss	x	y	w_1	w_1
1	$-410.8x - 132$	-66	132	425.8	-132	-410.8
2	$31495.3x - 11726.2$	5959.08	-11918.16	-31906.096	11786.16	31495.3

$h_w(x) = 15x + 0$

$h_{w_1}(x) = 15(1.6) = 24$

$h_{w_2}(x) = 15(2.0) = 30$

$h_{w_3}(x) = 15(2.5) = 37.5$

$h_{w_4}(x) = 15(3.0) = 45$

$h_{w_5}(x) = 15(3.5) = 52.5$

Here, error = Actual − Prediction

$e_1 = (30 - 24) = 6$

$e_2 = (27 - 30) = -3$

$e_3 = (24 - 37.5) = -13.5$

$e_4 = (22 - 45) = -23$

$e_5 = (20 - 52.5) = -32.5$

Subject : Date :

$$\frac{\partial Loss}{\partial w_0} = -2\sum (y_i - h_w(x))$$

$$= -2(6-3-13.5-23-32.5)$$

$$= -132$$

$$\frac{\partial Loss}{\partial w_1} = -2\sum x_i (y_i - h_w(x))$$

$$= -2\left[(1.6\times 6) + (2.0\times -3) + (2.5\times -13.5) + (3\times -23) + (3.5\times -32.5)\right]$$

$$= -2\left[9.6 - 6.0 - 33.75 - 69 - 113.75\right]$$

$$= -2(-212.9)$$

$$= 425.8$$

Now,

$$w_0 = w_0 - \alpha \frac{\partial Loss}{\partial w_0}$$

$$= 0 - 1\times(132)$$

$$= -132$$

$$w_1 = w_1 - \alpha \frac{\partial Loss}{\partial w_1}$$

$$= 15 - 1\times(425.8)$$

$$= -410.8$$

Final equation : $-410.8x - 132$

Subject : Date :

Again,

$h\omega(x) = -410.8x - 132$

$h\omega_1(x) = -410.8(1.6) - 132 = -789.28$

$h\omega_2(x) = -410.8(2.0) - 132 = -953.6$

$h\omega_3(x) = -410.8(2.5) - 132 = -1159$

$h\omega_4(x) = -410.8(3.0) - 132 = -1364.4$

$h\omega_5(x) = -410.8(3.5) - 132 = -1569.8$

Error = Actual - Prediction Value

$e_1 = (30 + 789.28) = 819.28$

$e_2 = (27 + 953.6) = 980.6$

$e_3 = (24 + 1159) = 1183$

$e_4 = (22 + 1364.4) = 1386.4$

$e_5 = (20 + 1569.8) = 1589.8$

$$\frac{\partial Cost}{\partial \omega_0} = -2\sum(y_i - h\omega(x))$$

$$= -2(819.22 + 980.6 + 1183 + 1386.4 + 1589.8)$$

$$= -11918.16$$

$$\frac{\partial Cost}{\partial \omega_1} = -2\sum x(y_i - h\omega(x))$$

$$= -2[(1.6 \times 819.28) + (2.0 \times 980.6) + (2.5 \times 1183) + (3.0 \times 1386.4) + (3.5 \times 1589.8)]$$

$$= -2[1310.848 + 1961.2 + 2957.5 + 4159.2 + 5564.3]$$

$$= -31906.096$$

Hamid Book Bainding

Subject: Date:

Now,

$$w_0 = w_0 - \alpha \frac{\partial Loss}{\partial w_0}$$

$$= -132 - 1(-11918.16)$$

$$= 11786.16$$

$$w_1 = w_1 - \alpha \frac{\partial Loss}{\partial w_1}$$

$$= -410.8 - 1(-31906.096)$$

$$= -410.8 + 31906.096$$

$$= 31495.296$$

Final equation: $31495.296x + 11786.16$

(c) from question 2;

$$31495.296x + 11786.16$$

For $x = 3$, $y = 31495.296(3) + 11786.16$

$$= 106272.048 \text{ mpg}$$

Actual = 22 mpg

Predict = 106272.048 mpg

difference = (106272.048 − 22) mpg

$$= 106250.048 \text{ mpg}$$

Here learning rate $\alpha = 1$ causes a misslead of prediction. The rate must in range of 0.01 to 0.001 for better output.

Subject : Date :

(d)

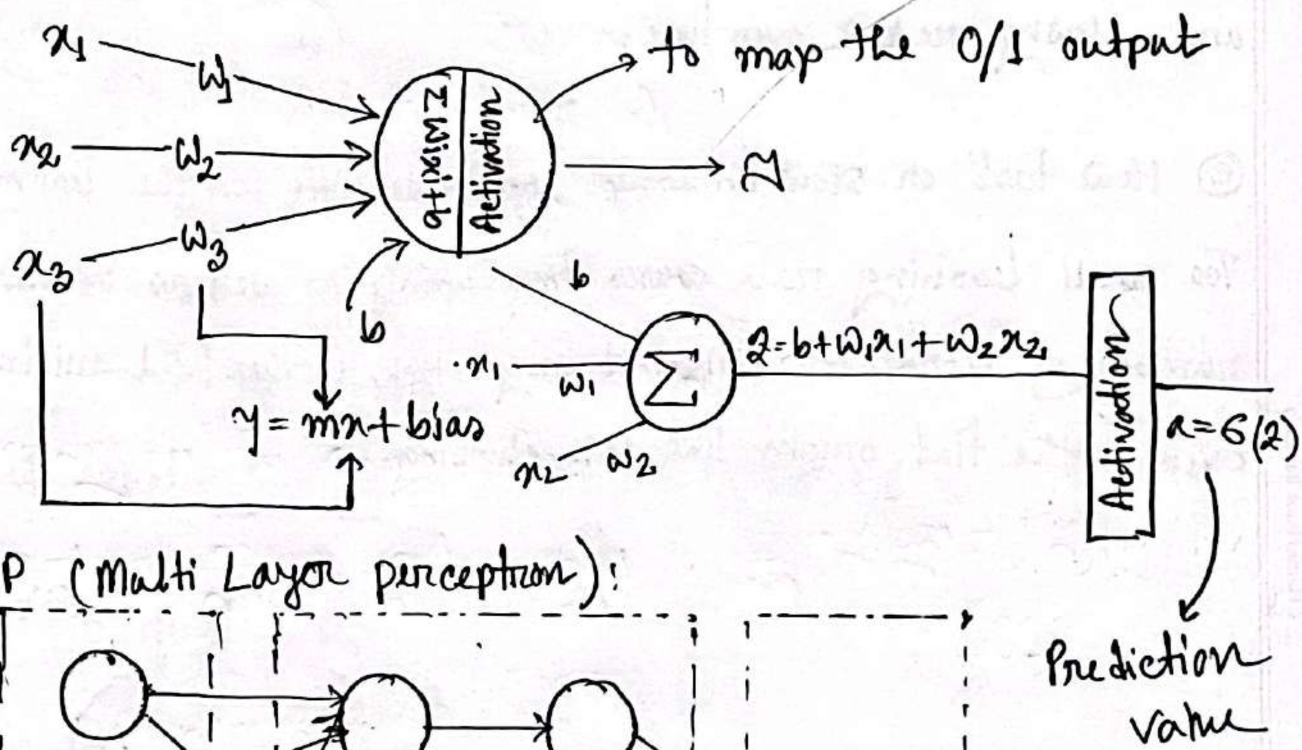
Large learning rate causes of divergence of the output. Also a large learning rate may cause a bad prediction. On the other hand, a decaying learning rate start with a high value to jump quickly and slowly reduce over-time.

(e) How fast or slow a model learn, depends on the learning rate. Too small learning rate causes tiny update as well as increase the number of iteration. Also, it may stock in the local minima or even in the flat origin like hill climbing.

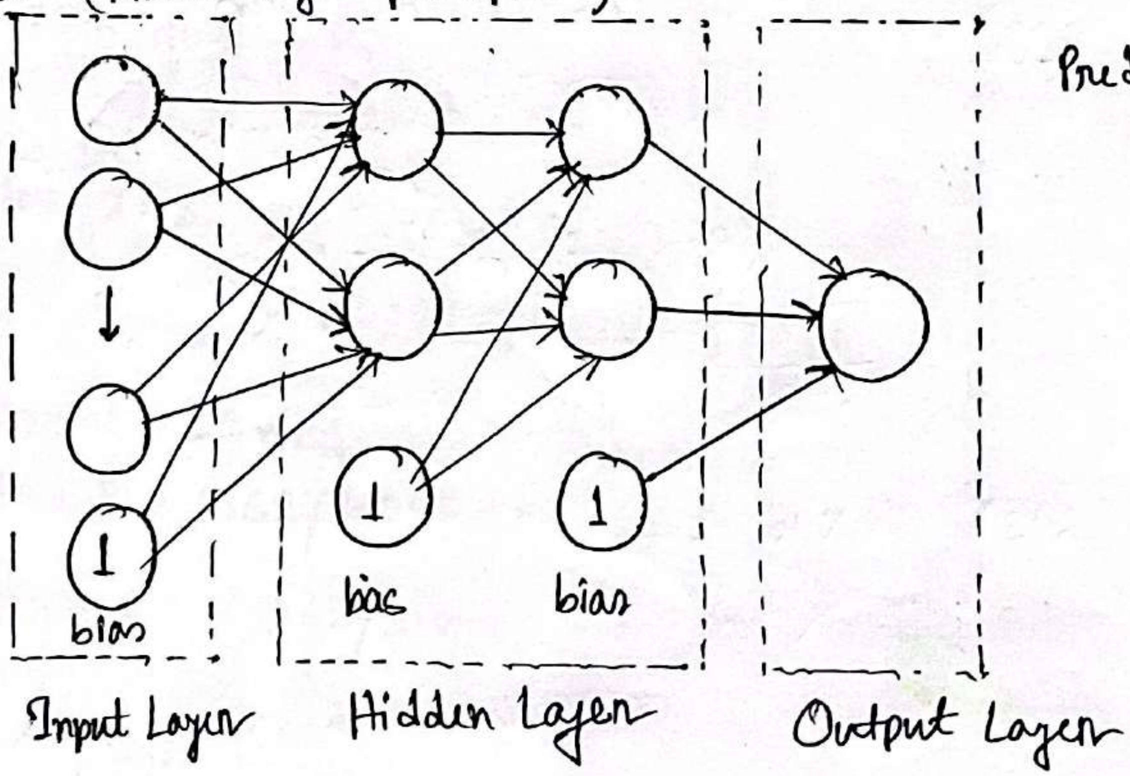
Neural Network

What is a perceptron?

One layer neural network that can classify things into two parts.



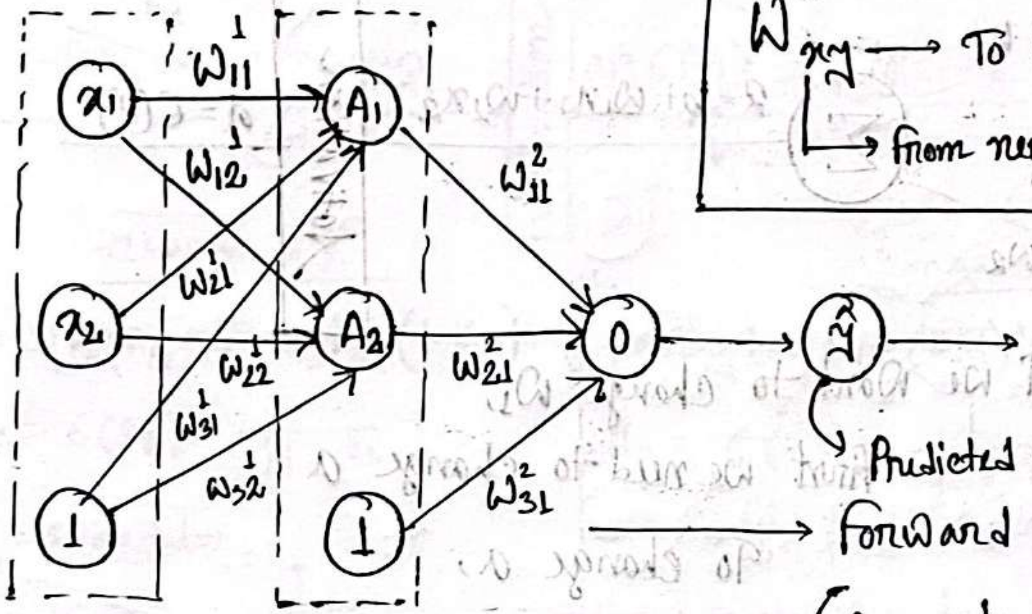
MLP (Multi Layer perceptron):



Subject: Date:

Forward Propagation

W^{2}_{xy} → Layer number (2); → To neuron (y); → from neuron (x)



Layer 1 Layer 2

Predicted value

→ Forward Propagation

$$\sum_{i=1}^{n} w_i x_i + b = \begin{cases} A_1 = w^1_{11}x_1 + w^1_{21}x_2 + w^1_{31} \\ A_2 = w^1_{12}x_1 + w^1_{22}x_2 + w^1_{32} \\ O = w^2_{11}\sigma(A_1) + w^2_{21}\sigma(A_2) + w^2_{31} \end{cases}$$

$$\underbrace{\sigma(x) = \frac{1}{1+e^{-x}}}_{\text{Sigmoid Function}}$$

Also called as the activation function → $\boxed{A_1/A_2}$ A_1 or A_2 or ---- A_n

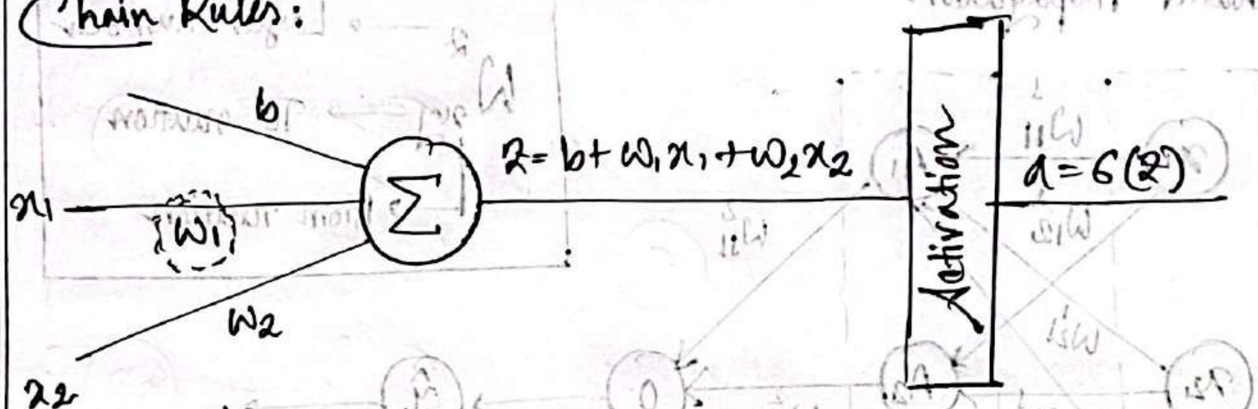
$$\hat{y} = \sigma \cdot \begin{bmatrix} w^2_{11} & w^2_{21} & w^2_{31} \end{bmatrix} \cdot \sigma \begin{bmatrix} w^1_{11} & w^1_{21} & w^1_{31} \\ w^1_{12} & w^1_{22} & w^1_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

$$\hat{y} = \sigma \cdot w^2 \cdot \sigma\, w^1 x$$

- w^1 → Weight of Layer 1
- w^2 → Weight of Layer 2
- σ → Sigmoid Function

Subject: Date:

Chain Rules:



If we want to change w_1,
first we need to change a
To change a,
we need to change z

According to the chain rules:

$$\frac{dL}{dw_1} = \frac{dz}{dw_1} \times \frac{da}{dz} \times \frac{dL}{da}$$

$$\underbrace{L = -y\log a - (1-y)\log(1-a)}_{\text{LOSS}}$$

Why this dependency?

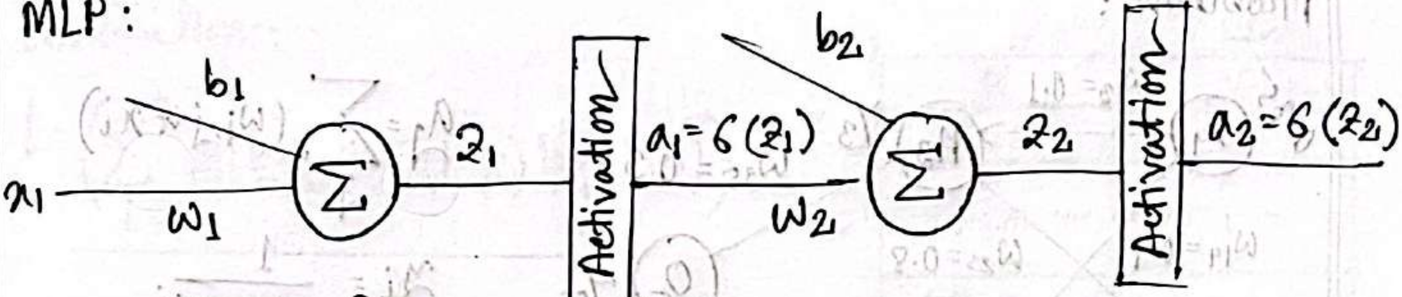
$$L = -y\log a - (1-y)\log(1-a)$$

$\hookrightarrow a = \sigma(z) = \dfrac{1}{1+e^{-z}}$

$\hookrightarrow z = b + x_1 w_1 + x_2 w_2$

Subject : Date :

MLP :



Change W_2

$$L = -y \log a_2 - (1-y)\log(1-a_2)$$

$$a_2 = \sigma(z_2) = \frac{1}{1+e^{-z_2}}$$

$$z_2 = a_1 w_2 + b_2$$

$$a_1 = \sigma(z_1) = \frac{1}{1+e^{-z_1}}$$

$$z_1 = x_1 w_1 + b_1$$

$$\frac{dL}{dw_2} = \frac{dz_2}{dw_2} \times \frac{da_2}{dz_2} \times \frac{dL}{da_2}$$

$\boxed{W_1}$

Change W_1

$$L = -y \log a_2 - (1-y)\log(1-a_2)$$

$$a_2 = \sigma(z_2) = \frac{1}{1+e^{-z_2}}$$

$$z_2 = a_1 w_2 + b_2$$

$$a_1 = \sigma(z_1) = \frac{1}{1+e^{-z_1}}$$

$$z_1 = x_1 w_1 + b_1$$

$$\frac{dL}{dw_1} = \frac{dz_1}{dw_1} \times \frac{da_1}{dz_1} \times \frac{dz_2}{da_1} \times \frac{da_2}{dz_2} \times \frac{dL}{da_2}$$

$\boxed{W_2}$

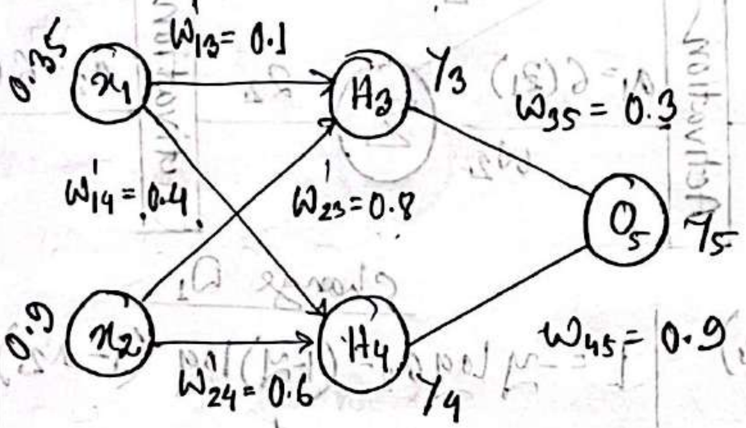
Here w_1 and w_2 are independent. Change in w_1 doesnot effect in w_2. But does effect a_1 and z_1.

For W_1, W_2 & W_3:

$$\frac{dL}{dw_1} = \frac{dz_1}{dw_1} \times \frac{da_1}{dz_1} \times \frac{dz_2}{da_1} \times \frac{da_2}{dz_2} \times \frac{dz_3}{da_2} \times \frac{da_3}{dz_3} \times \frac{dL}{da_3}$$

Subject: Date:

Problems:



$$a_j = \sum_i (w_{i,j} * x_i)$$

$$y_i = \frac{1}{1+e^{-a_i}}$$

$Y_5 = 0.5$

$$a_1 = (w_{13} \times x_1) + (w_{23} \times x_2)$$
$$= (0.1 \times 0.35) + (0.8 \times 0.9)$$
$$= 0.755$$

$$y_3 = \frac{1}{1+e^{-0.755}}$$
$$= 0.68$$

$$a_2 = (w_{14} \times x_1) + (w_{24} \times x_2)$$
$$= (0.4 \times 0.35) + (0.6 \times 0.9)$$
$$= 0.68$$

$$y_4 = \frac{1}{1+e^{-0.68}}$$
$$= 0.6637$$

$$a_3 = (w_{35} \times Y_3) + (w_{45} \times Y_4)$$
$$= (0.3 \times 0.68) + (0.9 \times 0.6637)$$
$$= 0.801$$

$$Y_5 = \frac{1}{1+e^{-x}} = \frac{1}{1+e^{-0.801}} = 0.69 \text{ [output]}$$

Error = Y target − Y pred = (0.5 − 0.69) = −0.19

Subject : Date :

Loss function:

$$L = \frac{\left(\sum_{i=1}^{3} -y_i \log_{ai} a_i - (1-y_i)\log(1-a_i)\right)}{3}$$

X	Y
0.1	0
0.2	0
0.3	1
0.4	

Assume that,

$W = 0.7$ and $b = 0.1$

$z = 0.7x + 0.1$ $[y = mx + b]$

$$L = -0\log\left(\frac{1}{1+e^{-(0.7\times0.1+0.1)}}\right) - (1-0)\log\left(1 - \frac{1}{1+e^{-(0.7\times0.1+0.1)}}\right)$$

$$- \boxed{} - \boxed{}$$

$$= -\log\left(1 - \frac{1}{1+e^{-(0.7\times0.1+0.1)}}\right) - \log\left(1 - \frac{1}{1+e^{-(0.7\times0.2+0.1)}}\right)$$

$$- \log\left(\frac{1}{1+e^{-(0.7\times0.3+0.1)}}\right)$$

$$= -\log(1-0.54) - \log(1-0.55) - \log 0.57$$

$$= 0.928$$

Avg Loss value $= \left(\frac{0.928}{3}\right) = 0.309$

Subject : Date :

$$\frac{dL}{dW} = \frac{\left(\sum_{i=1}^{n}\left(\frac{1}{1+e^{-x}} - y_i\right) x_i\right)}{3}$$

$$W = W - \alpha \frac{dL}{dW}$$

$$\frac{dL}{db} = \frac{\left(\sum_{i=1}^{n}\left(\frac{1}{1+e^{-x}} - y_i\right)\right)}{3}$$

$$b = b - \alpha \frac{dL}{db}$$

$$\frac{dL}{dW} = \left(\left(\frac{1}{1+e^{-(0.7\times0.1+0.1)}} - 0\right)0.1 + \left(\frac{1}{1+e^{-(0.7\times0.2+0.1)}} - 0\right)0.2\right.$$
$$\left. + \left(\frac{1}{1+e^{-(0.7\times0.3+0.1)}} - 1\right)0.3\right)/3$$

$$= \frac{(0.054 + 0.112 - 0.127)}{3}$$

$$= 0.013$$

$$\frac{\partial L}{\partial b} = \left(\left(\frac{1}{1+e^{-(0.7\times0.1+0.1)}} - 0\right) + \left(\frac{1}{1+e^{-(0.7\times0.2+0.1)}} - 0\right)\right.$$
$$\left. + \left(\frac{1}{1+e^{-(0.7\times0.3+0.1)}} - 0\right)\right)/3$$

Subject: | Date:

$$= \frac{(0.54 + 0.56 - 0.42)}{3}$$

$$= 0.23$$

if learning rate, $\alpha = 1\%$.

updated weight, $w = w - \alpha \frac{dL}{dw}$

$$= 0.7 - 1(0.013)$$

$$= 0.687$$

updated bias, $b = b - \alpha \frac{dL}{db}$

$$= 0.1 - 1(0.23)$$

$$= -0.13$$

Old $z = 0.7x + 0.1$

new $z = 0.687x - 0.13$

L for new z will be, $L = 0.86$

Avg Loss value $= \frac{0.86}{3}$

$= 0.2866$ [new]

Avg Loss value $= 0.309$ [previous]

Conclusion error/Loss slightly reduce from 0.309 to 0.2866.

Subject: Date:

LOSS function:

Logistic Regression:

$$\frac{\partial Loss}{\partial W_j} = (y_{pred} - y_{act}) \cdot x_j$$

$$\frac{\partial Loss}{\partial b} = (y_{pred} - y_{act})$$

$$\frac{\partial L}{\partial W} = \sum_{i=1}^{n} \left(\frac{1}{1+e^{-x}} - y_i \right) x_i$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n} \left(\frac{1}{1+e^{-x}} - y_i \right)$$

Update weight:

$$W = W - \alpha \frac{d\,Loss}{dW}$$

Update bias:

$$W = W - \alpha \frac{d\,Loss}{db}$$

Sigmoid function: $\frac{1}{1+e^{-x}}$

$$\frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right) = \frac{e^{-x}}{1+e^{-x}} \rightarrow \frac{1}{1+e^{-x}} e^{-x}$$

$$\sigma(x) \quad (1 - \sigma(x))$$

or, $\sigma'(x) = \sigma(x) \cdot (1 - \sigma(x))$

Scanned with CamScanner

Subject: | Date:

Gradient Descent Loss function:

$$L = \frac{1}{n}\sum_{i=1}^{n}(Y_i - (mx+c))^2$$

$$\frac{\partial L}{\partial m} = -\frac{2}{n}\sum x_i (y_{actual} - y_{pred})$$

$$\frac{\partial L}{\partial c} = -\frac{2}{n}\sum (y_{actual} - y_{pred})$$

Actual Data Table

↳ Create an equation and predict the output

↳ Calculate the loss function

↳ Update w an bias

↳ $w = w - \alpha \frac{dL}{dw}$

$b = b - \alpha \frac{dL}{db}$

MES/SSR:

Actual Table

↳ Prediction

↳ Find Loss $\frac{\sum(\text{actual} - \text{predict})^2}{n}$

↳

Subject: Date:

$$E(\text{Decision}) = -P(\text{yes})\log_2 P(\text{yes}) - P(\text{No})\log_2 P(\text{No})$$

$$= -\frac{9}{14}\log_2\left(\frac{9}{14}\right) - \frac{5}{14}\log_2\left(\frac{5}{14}\right)$$

$$= 0.940.$$

Student.

$E(\text{student} \mid \text{yes}) =$

$$E(\text{Decision}) = -P(\text{yes})\log_2(\text{yes}) - P(\text{No})\log_2 P(\text{No})$$

$$= -\frac{4}{10}\log_2\left(\frac{4}{10}\right) - \frac{6}{10}\log_2\left(\frac{6}{10}\right)$$

$$= 0.971$$

x_1

$$E(\text{yes} \mid a_1) = -\frac{4}{5}\log_2\left(\frac{4}{5}\right) - \left(\frac{1}{5}\right)\log_2\left(\frac{1}{5}\right) \cancel{- \frac{5}{5}\log_2()}$$

$$\cancel{E(\text{No} \mid a_1) = -\frac{5}{5}}$$

Subject : Date :

$$E(D) = -P(Y)\log_2 P(Y) - P(N)\log_2 P(N)$$

$$= -\frac{4}{10}\log_2(4/10) - \frac{6}{10}\log_2(6/10)$$

$$= 0.971$$

$X1$:

$$\underset{a_1}{E(X1)} = -P(Y|X_1)\log_2 P(Y|X_1) - P(N|X_1)\log_2 P(N|X_1)$$

$$= -\frac{4}{5}\log_2(4/5) - \frac{1}{5}\log_2(1/5) = 0.72$$

$$E(a_2) = -P(Y|a_2)\log_2 P(Y|a_2) - P(N|a_2)\log_2 (N|a_2)$$

$$= -\frac{0}{5}\log_2(0/5) - 5/5\log_2(5/5) = \cancel{0.361}\ 0$$

$$IG = 0.971 - 0.72\,P(a_1) - 0.361\,P(a_2)$$

$$= 0.971 - 0.721\ 5/10 - \cancel{0.361}\ 5/10\ 0$$

$$= 0.610$$