Mon Tue Wed Thu Fri Sat

Date: __/__/20

14 attr

$\frac{1}{14}$ → entropy closer to zero

iska info gain 0.94 ho jaye gy

$G(S, A = \text{outlook}) = 0.94 - 0$

→ gain ratio use karlein to solve this biasedness problem

post rule post pruning poora chap 3 yahi hoga

→ jitna shallower decision tree utna behtar

rule post pruning

→ stop cut off / prune off

after you have done your training, over-fitting kel baad rules likh lenge, aap unki pre condition ko nikaal dein

A ∧ B ∧ ~~C~~ → Class 1

ek ko nikaal kar dekhein kya asar aata.

validation error badhta toh you cant cut off C, otherwise you can remove

~~A~~ ∧ B ∧ C

~~A~~ ko bhi nikaal sakte

→ missing values

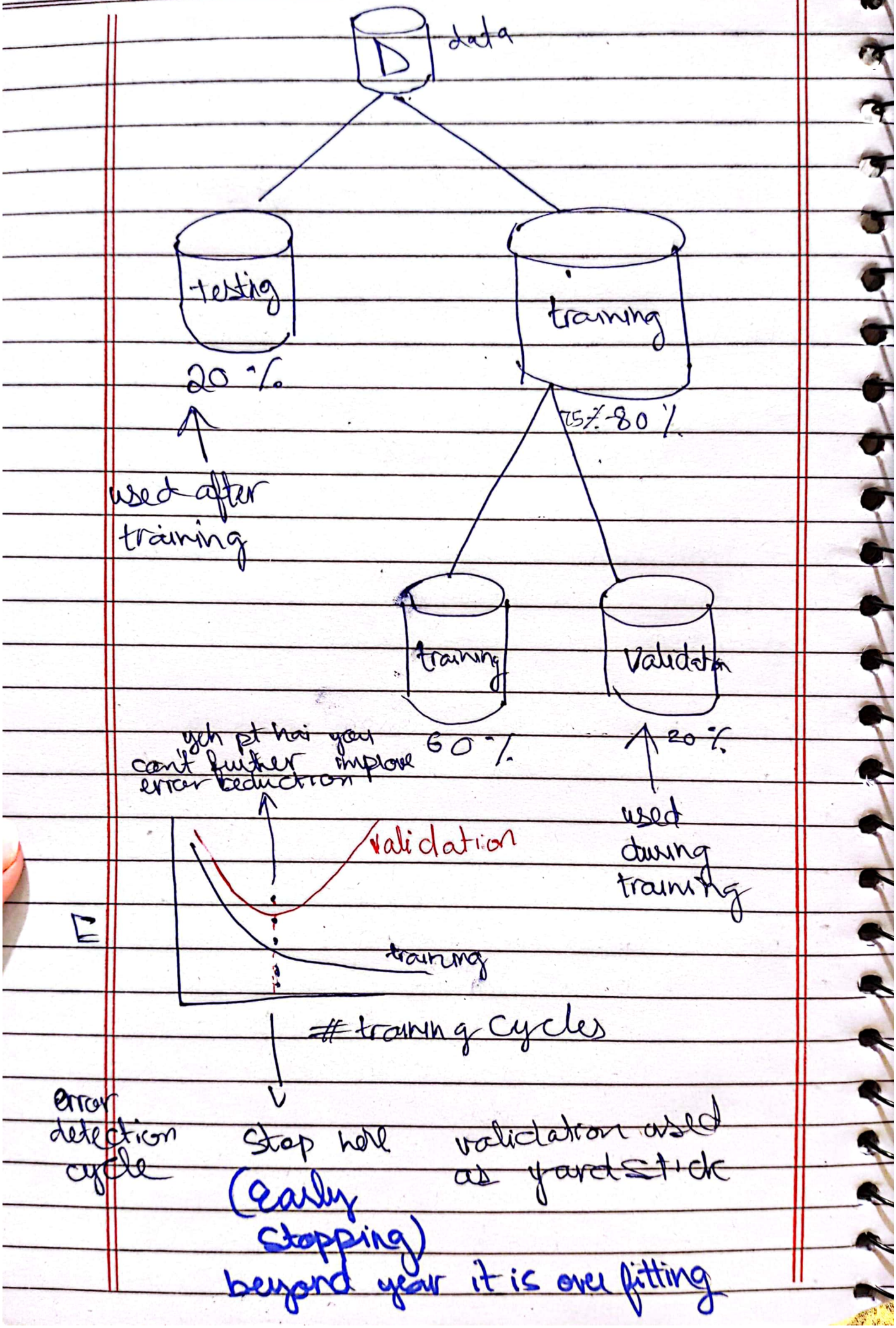
→ dt pure partitions karna chahta hai

ek attr 10 possible val ek ka 4 ek ka 2

↓

individual entropy

aray lekin aray ka chance purity toh uski kam zyada


D
data
testing
training
20 %
used after training
75%-80 %
training
Validation
60 %
20 %
used during training
yeh pt hai you can't further improve error reduction
validation
training
E
training cycles
error detection cycle
Stop here
(early stopping)
beyond your it is over fitting
validation used as yardstick

Date: __/__/____

Mon Tue Wed Thu Fri Sat

examples empty : no value against it
no, existence

ID3 → based best on entropy
(info gain)

issues associated with virtually any ML model:

→ overfitting (training perfect)
100 %
near 100 %

→ test error is high, apnay data set ka bohot learn kar liya but unseen data pe poor work karta

noise in data, model learned noise & regularities of data

high variance scenario

how to avoid over-fitting?

→ ID3 would always overfit, unless there are some checks to stop it from learning to more than some certain depth
ya toh 7-8 samples hain, usseh further split na karen
→ Validation set banatay, aik portion reserve for testing

agar saaray no bhi toh entropy = 0

Mon Tue Wed Thu Fri Sat Date: __/__/20

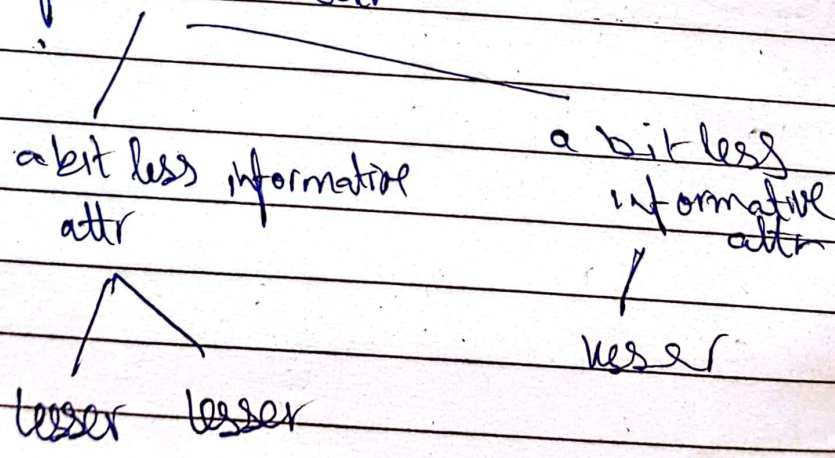
G(S, A=outlook) = 0.24

pure ki entropy 0, pure isse kehtay jiski sirf aik hi class aaye.

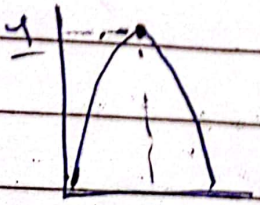
recursion seh khud call karkay thoray attributes keh saath yeh reduced set keh liye kaam karay gy

agar root node ka poochay toh sabka calculate karengay

Wana specific poochay toh aik ka hi.



for boolean classification toh entropy can be max 1 & min 0



max entropy 1 tab aati. jab dono classes ki prob same.

Date: __/__/20

Mon Tue Wed Thu Fri Sat

$$G_1(S, A=\text{outlook}) = 0.94 - \left(\frac{|S_{A=\text{sunny}}|}{|S|} \cdot E(S_{A=\text{sunny}}) + \frac{|S_{A=\text{overcast}}|}{|S|} \cdot E(S_{A=\text{overcast}}) + \right.$$

$$\left. 0.94 - \left(\frac{|S_{A=\text{Sunny}}|}{|S|} \cdot E(S_{A=\text{sunny}}) + \frac{|S_{A=\text{overcast}}|}{|S|} \cdot E(S_{A=\text{overcast}}) + \frac{|S_{A=\text{rain}}|}{|S|} \cdot E(S_{A=\text{rain}}) \right) \right.$$

$|S| = 14$

$|S_{A=\text{Sunny}}| = 5$

$|S_{A=\text{overcast}}| = 4$

$|S_{A=\text{rain}}| = 5$

$$E(S_{A=\text{Sunny}}) = -\left(\frac{2}{5} \log \frac{2}{5} + \frac{3}{5} \log \frac{3}{5} \right)$$

★ Saaray yes of overcast tah entropy = 0

$E(S_{A=\text{overcast}}) = 0$ since only one class exists thats why entropy = 0

Mon Tue Wed Thu Fri Sat

Date: __/__/20

decision trees used for taking decisions

given your data, temperature doesn't matter

Criteria for selecting attribute as root?

→ $G(S,A)$ is maximum

$$G(S,A) = E(S) - \sum_{v=1}^{M} \frac{|S_v|}{|S|} \cdot E(S_v)$$

$E(S)$: entropy

S bara set hai

$E(S_v)$: entropy

$$E(S) = -\sum_{c=1}^{2} P_c \log_2 P_c$$

$$= -\left(P_{yes} \log_p Y_{es} + P_{No} \log_p NO\right)$$

$$= 0.94$$

$c=1$ → means yes

$c=2$ → No

log 2 nhi toh $\log_{10} \frac{9}{14}$ karo

$P_{yes} = \frac{9}{14}$

$P_{No} = \frac{5}{14}$

conv

$$\frac{\log_{10} 9/14}{\log_{10} 2}$$

Mon Tue Wed Thu Fri Sat

Date: __/__/20__

AI

Kuch mein sirf input provide karay, & ML looks for patterns. categorical data

→ discrete labels { eg: sunny, rain, OC }

$S = \{ \langle x^1, y^1 \rangle, \langle x^2, y^2 \rangle, \ldots, \langle x^n, y^n \rangle \}$

n tuples in data set

x^i → table with attr & value

$x^i = \{ (a_j, v_j) \}$

$y^i = \{0, 1\}, \{yes, No\}, \{+, -\}$ Binary classification

$\{1, 2, \ldots k\}$ multiclass classification

2 tasks of supervised learning

- classification
- regression

Numeric values

$y^{(i)} \in \mathbb{R}$ } regression

$y^{(-1)}$

isko input normally numeric attributes but categorical bhi ho saktay but output is always numeric.

Mon Tue Wed Thu Fri Sat Date: __/__/20

probability = 1 if all yes : message is 0.

adhi yes adhi no, 1 bit to encode yes, other bit to encode no.

$\frac{1}{2} \log_2 \frac{1}{2}$

$\log 2^{-1}$

$\log_2 \frac{1}{2} = \log_2^1 - \log_2^2$

$= 0 - 1$

$= -1$

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

toh negative sign iss liye lagaday keh positive ho jaye

$-\left[\left(\frac{-1}{2}\right) + \frac{1}{2}\right] = 1$

$\text{Entropy}(S) = -\left[\left(\frac{9}{14} \log_2 \frac{9}{14}\right) + \left(\frac{5}{14} \log_2 \frac{5}{14}\right)\right]$

$= -(-0.94)$

$= 0.94$

id3 algorithm
iterative dorcemizer version 3

id syllabus : 3, 4, 12 → 12.1 + decision tree numericals, mcqs, short conceptual questions

Gain(S, outlook) = 0.246

Entropy : degree of purity ko number assign karta hai

2 classes hain (binary) : yes or no

$$Entropy(S) = -P_{\oplus} \log_2 P_{\oplus} - P_{\ominus} \log_2 P_{\ominus}$$

probability of seeing a yes

$P_{\oplus} = \frac{9}{14}$ 9 → yes, 5 → no

$P_{\ominus} = \frac{5}{14}$ probability of seeing a no

$\log_2 P_{\oplus} = \log_2 \frac{9}{14}$

log kyun liya gaya
& saath no. kyun multiply kiya
& why the minus sign

entropy in physics : measure of disorder

why $\log_2$ used?

↳ How many bits do you need to encode the qty

Same classification → so you need message with 0, don't need any bit to encode this message

→

In a set of decision trees, it selects a simple decision tree

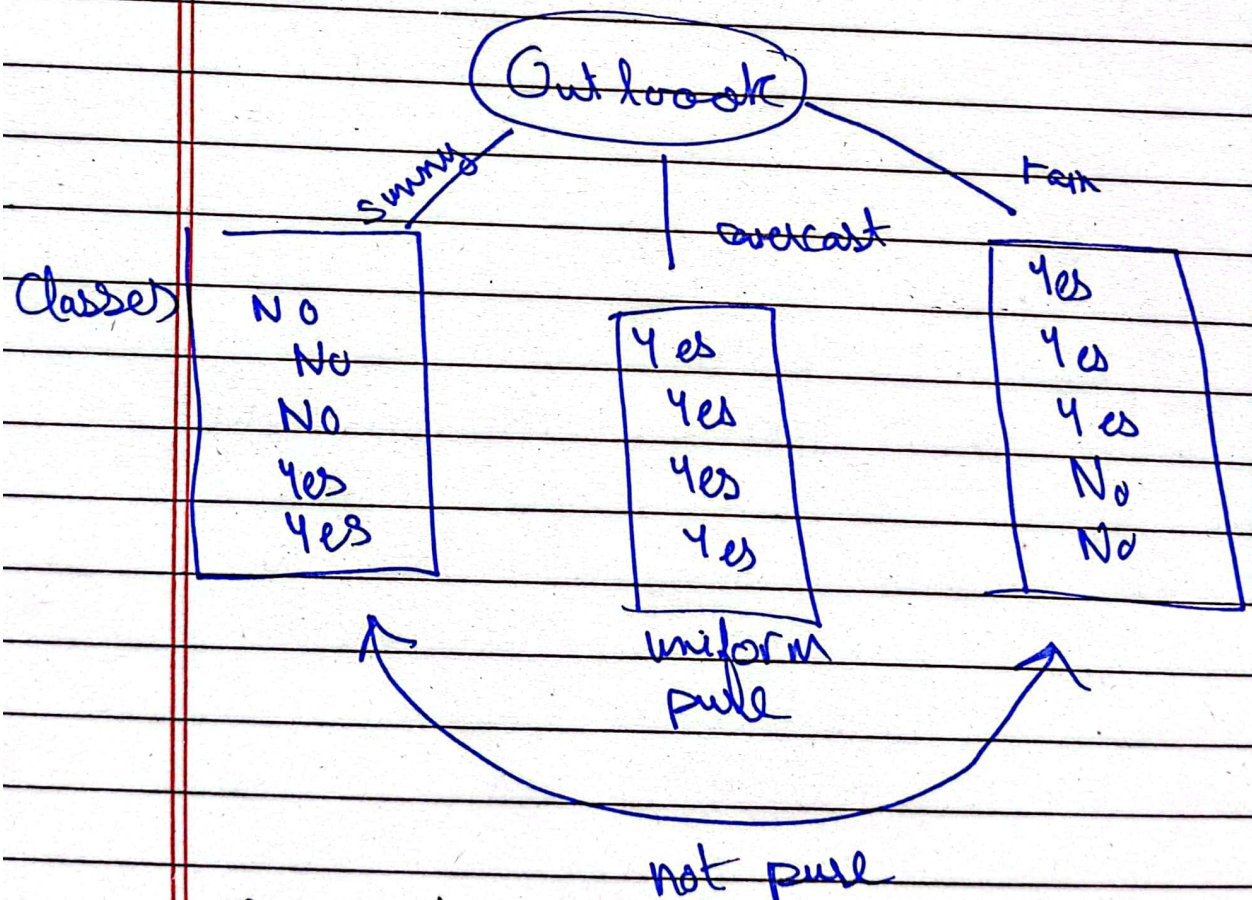
→ aik hi node wala sabsay kam level wala decision tree ho sakta

(yes) → Simplest possible decision tree

→ shallowest, with least attributes

→ rarely possible

Sabsay zyada uniform classes ki distribution hain



aisi root node select karni jisse partitions as pure as possible hoonay

Mon Tue Wed Thu Fri Sat

Date: __/__/20__

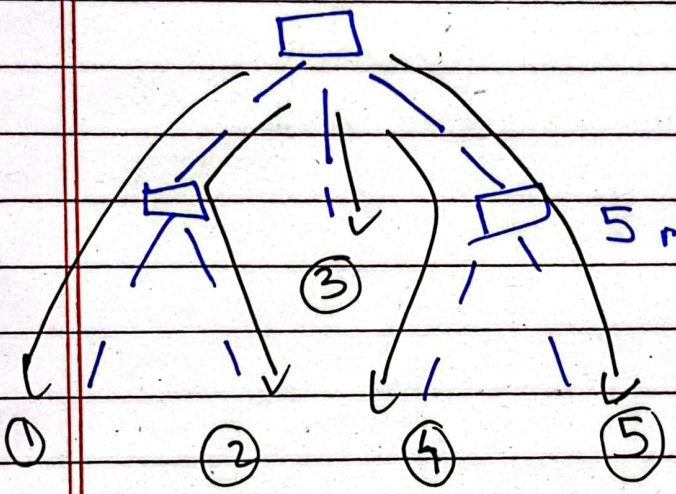
Decision Trees

→ categorical data (no numeric attribute)

ch3 of another book

Weatherseh tell kai tennis khelta.

decision tree → temperature dt mein included nhi hai



how to decide konsa attribute use karna konsa nhi, kisko root node banana.

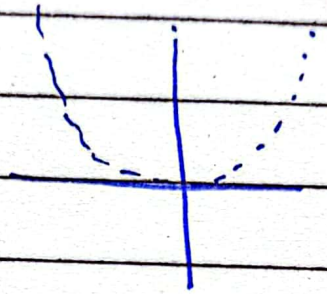
agar temperature numerical data hai toh you want to make into categorical data wrap interval bana saktay, there are other systematic methods as well.

→ recursive algo (data mein aik attribute select kartay data uss bunyaad per split karta

I/P O/P

x	f(x)
-1	1
0	0
2	4
3	9
⋮	⋮

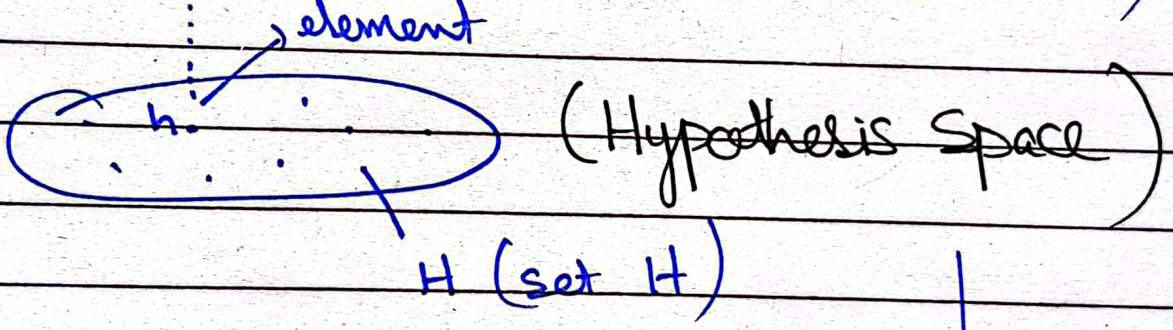
$= x^2$



$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

h(x)
Program

g → unknown function (generated your data)



h is closest to g

We cannot access g function

Consistently (pooray data ka explain kalday)

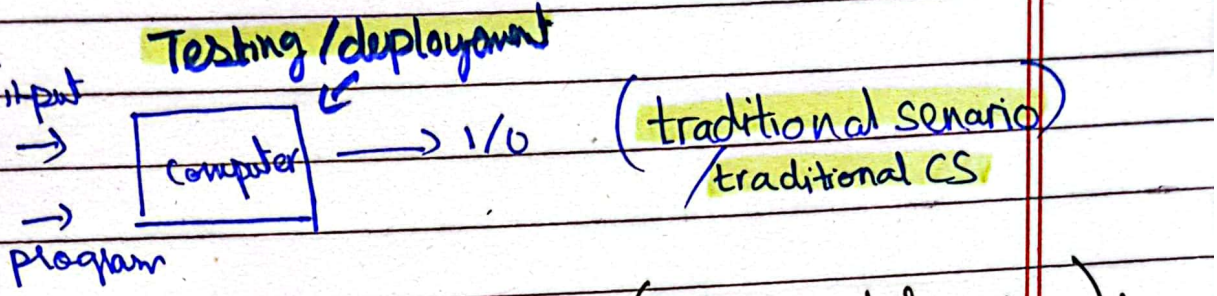
Set of all possible polynomials in this case

Mon Tue Wed Thu Fri Sat

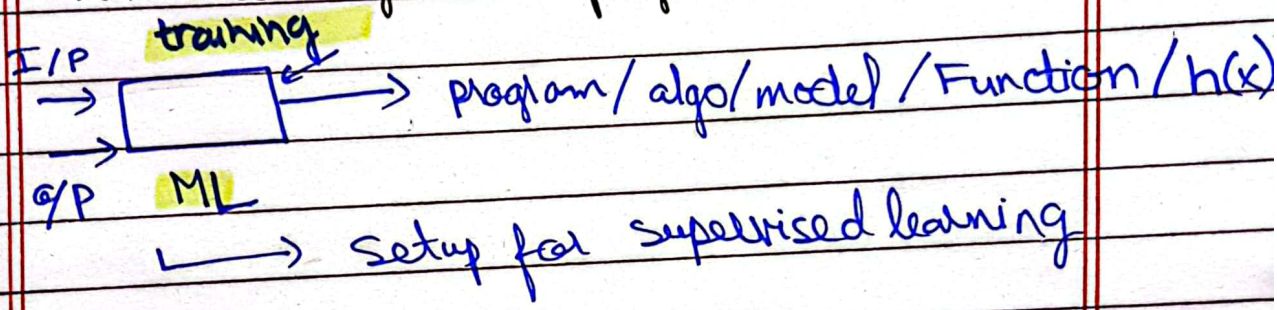
Date: 5/11/2024

AI

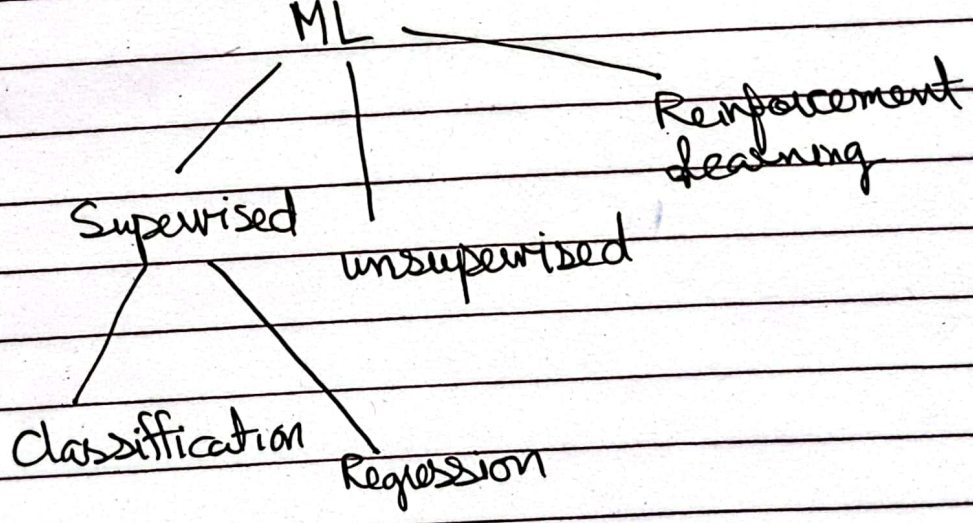
ML - Supervised learning mainly



Machine learning ki category (Supervised Learning):



→ setup for supervised learning



In Supervised learning (ML)

Learning : is performance improvement in a sufficient task given sufficient data.

Date:

$f_i = \frac{1}{c_i}$ fitness; jiski cost kam

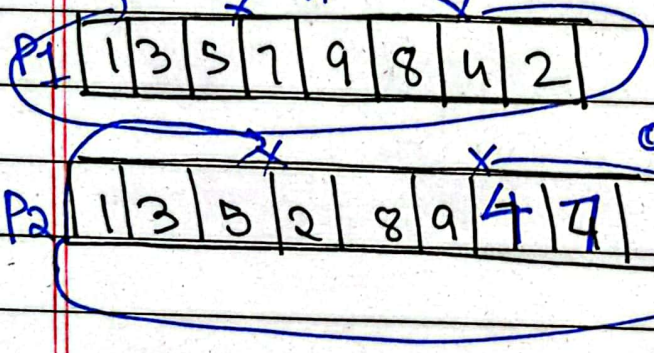
fitness
selection
crossover → single parent not allowed
mutation

invalid sol:
Sab ko visit nhi kar paye ga, we cannot visit one node twice

order crossover

define cut points

6 → start



4 2 1 3 5 7

cut points keh as it is likh den

4 7 1 3 5 2 8 9

C1 | 4 | 1 | 3 | 7 | 9 | 8 | 5 | 2 |

C2 | 4 | 1 | 3 | 2 | 8 | 9 | 5 | 7 |

mutation: 2 random cities ko pakar kar swap karein

Date:

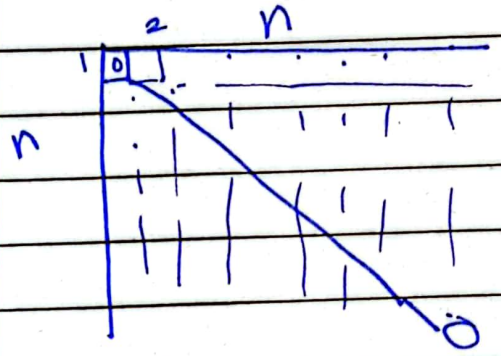
totallogy

homesha hi false hoga

guarantee karta sol aa jaye gi, koh loop chalti jaye ga sol exists → please confirm it

TSP problem

n cities to visit
host per start karna ga return to it
n-1 nodes → one time visit
1 node → 2 times visit

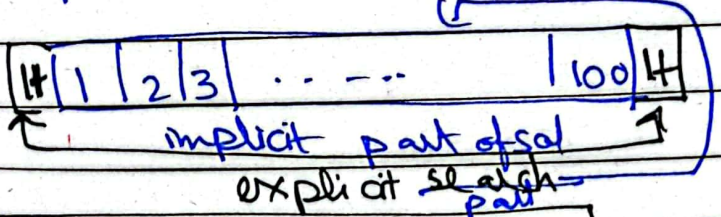


not symmetric always

give best loop

n! solutions → agar brute force
n factorial

Representation kesay?



nearest neighbor heuristic is good if not genetic algo.

Selection:
roulette's wheel

25 dafa crossover karkay 50 create ho jayen gay child.

We set limit for crossover.

Crossover
→ single point bhi ho sakta
→ multipoint bhi ho sakta (define parts)
→ uniform bhi ho sakta

can generate index randomly and bachon ki bhi wahi pass wahhi index mein hon.

uniform; probability assign karen 0-1 keh beteen.
phir fill karatay bit by bit, uniform probability. fair coin

Introduction selection
crossover
mutation bhi basi asaan hai.
replacement
25% weak nikaal doo, define your replacement, full, partial

★ Soch lein and implement it.

Date:

1) CNF Satisfiability Prob
Conjuctive normal form

boolean expression ~~A=B~~

$\neg a \vee b$ → 0, 1's → a, b

a=0
b=0; $\neg a \vee b$ = True

boolean algebra seh nikli

not in cnf form
disjunction

$\neg a \wedge b$ = true
a=0
b=1
conjunction

can be arbitarily long:

$(\neg a \vee b) \wedge (a \vee d \vee \neg e) \wedge (\neg c \vee f) \ldots\ldots$

~~clauses~~

CNF is conjunction of disjunction (clauses)

exponential growth
3 var 2^3

hard problem

genetic algo for CNF problem

1	0	0	1	1	0
a	b	c	d	e	f

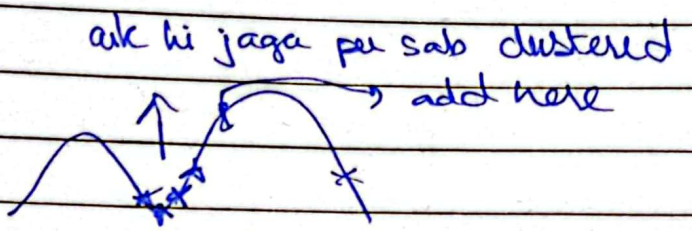
goal iska True karna poora eq.

fitness: $\frac{\text{no. of true clauses}}{\text{total no. of clauses}}$

Date:

If there is only one hill, greedy search or hill climbing search is the best approach, would give optimal solution

Genetic algo works on multiple solutions simultaneously



diversity ka kya farda?
like sab fitness ratio almost same aa rahi hai
jab stagnant ho jaye tab stop

As the number of size increases, so brute force is not feasible

there is a method to madness
search meta heuristic
Hill climbing mein best ka le kar chalen.

neural networks mein aa jaye gi hill climbing ho jai

12.1 It is not necessary for knapsack problem to use genetic algo, you can just use hill climbing

Date: AI

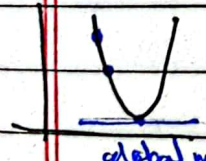
tournament selection

Knapsack 0-1 → for genetic algorithm problem

It works
it helps transfer fitness traits (characteristics) parent to offspring

Q. Why we do crossover in genetic algo?
Q. What is the relationship of genetic algo & hill climbing search?
Q. What is the benefit of using mutation operator?
→ it creates diversity (variety)

→ genetic algo ~~is a parallel version of hill climbing~~.
hill climbing mein ek possibility.
Faida kya agar 1 banda hai ya 100 bandey any benefit?
(take a step in the direction of the goal)
we have multiple hills.
hills ka kya muhasbat with ga algo?
→ fitness ~~ratio/quality~~
landscape



min find karna
Ab yeh optimization prob. We have found minimum using differentiation.
put $\frac{dy}{dx} = 0$

Date:

CS2 10 111000
CS3 10 11 01100

$\frac{200+150+250+120}{5+15+4+7} = 23.22$

$\frac{200+150+250+160}{5+15+4+8} = 23.75$

weight is exceeding in both cases
31, 32

so for our sake of purpose, we ignore the weight was 30 & set it to 35kg

Ab weakest population ko hum replace kardein jin CS2 & CS3 sah

CS-2, CS-9 ko replace kardein

if invalid sol (weight zyada):
done an existing sol or initialize randomly a CS-5

go with the flow of algorithms

otherways for selection:
tournament selection

$r_2 \in [0,1] = 0.5$ CS-2

CS-3 & CS-2 per gene/ classo

	0	1	2	3	4	5	6	7
CS2	0	0	1	1	0	1	0	0
CS3	1	0	1	1	1	0	0	0

define type of crossover, then apply it

point

Single pt. crossover

CS2' 0 0 1 | 1 0 0 0

~~0 1 0 0 1 0~~

CS3' 1 0 1 | 0 1 0 0

question pertinent

bahut zyada mixing nhi krni

→ preserve wali chees khatam ho jaye gi

Mutation

→ Pick a random location (gene) & flip it.

Let the nature take its course.

→ 2-3% of gene

16 genes haia ton aik random location per change krtay, first sel randomly CS sol then bits within. like random search.

Iteration = generation

Termination conditions:
(i) desired fitness threshold has been achieved
(ii) Defined # of generation have expired

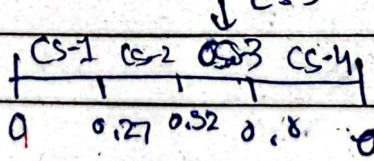
$\frac{f_i}{\sum f_i}$

	Fitness ratio
$f_1 = 22.14$	$22.14/87.63 = 0.27$
$f_2 = 20.74$	$= 0.25$
$f_3 = 23.22$	$= 0.28$
$f_4 = 21.53$	$= 0.24$
$\Sigma = 87.63$	

also called probability but better to call it fitness

Roulette wheel Selection

Cumulative fitness
0.27
0.52
0.8
0.99



selected

$r_1 \in [0,1] = 0.74$

$W_{max} = 30$ kg

Item	Value (1000PKR)	weight
1	200	5
2	100	10
3	150	15
4	250	4
5	120	7
6	160	8
7	50	12
8	40	6

0-1 Knapsack Problem (Solution via Genetic Algorithm)

	0	1	2	3	4	5	6	7
CS-1	1	0	0	1	1	0	1	0

CS-2 ~~0 0 0 1 1 0 1 0 0~~ 0 0 1 1 0 1 0 0

CS-3 1 0 1 1 1 0 0 0

CS-4 0 0 1 1 0 1 1 0

Fitness

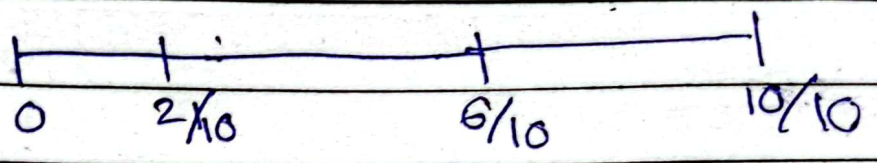
$f1 = \frac{200 + 250 + 120 + 50}{5 + 4 + 7 + 12} = 22.14$

$f2 = \frac{200 + 250 + 150 + 160}{5 + 4 + 7 + 6} = 20.74$

~~$f3 = 23.22$~~

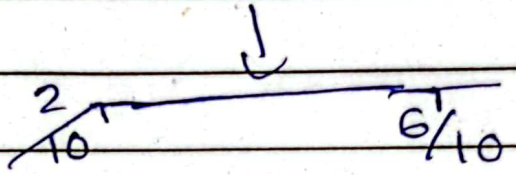
$f4 = 21.53$

Date:



roulette wheel logic:

0 seh 1 keh darmyaan random no. generate krain gaey



random no ka kya probability hai
→ uniform

gaussion bhi kaain gaey

selected ka tag laga dein gaey.
uskay upar crossover ki line hogi
and tab tag 0 ho jaey gey.

When to stop:

1. Jab desired fitness achieve ho jaye
2. Predefined no. of generations expires
3. Stagnation

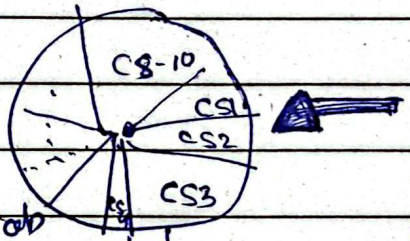
Can deteriorate but uss ay evolution mein improve hi hota

t = t+1

aik aur baar chalay ga for another time (second iteration.

Selection:

1. Fittest %
2. Roulette wheel selection
   gambling wheel
   Lets assume fair roulette wheel hai



unequal slice hai.
fittest slice ka portion ziyada toh

	Fitness	Prob Fitness ratio	
CS-1	2	2/10	It is fair, as it favours fitter ones more but there is an element of uncertainty
CS-2	4	4/10	
⋮	⋮		
CS-10	3	3/10	

$\sum_{i=1}^{10} f_i = 10$

Fitness ratio/Prob	Cumulative fitness ratio
2/10	2/10
4/10	6/10
⋮	⋮
3/10	10/10

Date:

crossover point can be selected at middle.

crossover
after creating children :
→ ab mutation (agani (random change)

bits flip kardein in knapsack

→ mutation is a random change
randomly

% genes mutation
1 gene 5-6 %
2 genes 10.% or more

⇒ 100 genes mein 2 ya 3 genes generate no sakti nayi.

diversity, variety
Stagnation

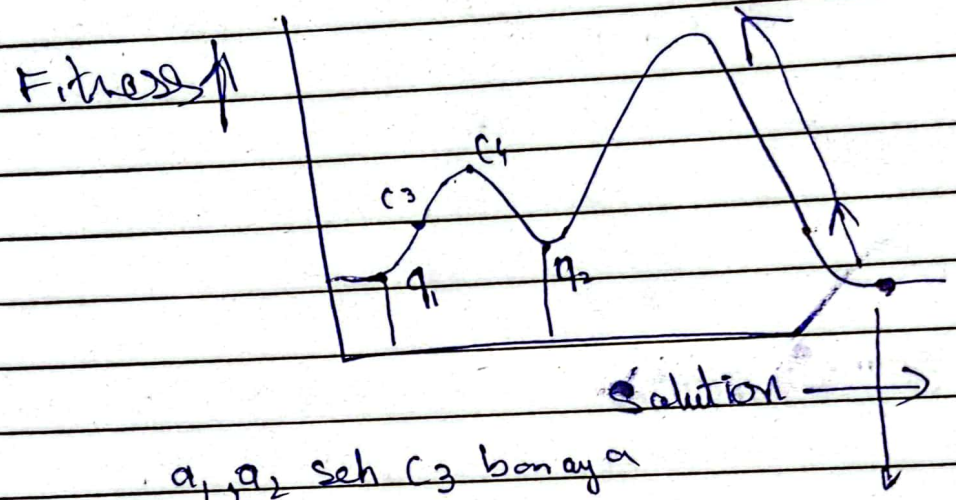
— replace P(t) with some offsprings
10 ki population
2 children create
ab population 12.
bit some rakhnay keh liye
you have to remove some
based on fitness.
add
partially replace
total replacement ⇒ crossover karna padega all old parents die

Date:

- Select a pool of fit solutions (parent)

top 4, top2

agar sabay elite hon toh we can be stuck after a while, and improve nhi hoga jis ki dja



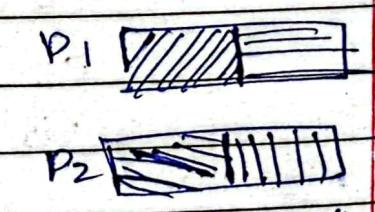
q1, q2 seh C3 banaya
C3 seh C4

isko ignore kiya toh agar explore karay toh kuch pata global max per pohunchay

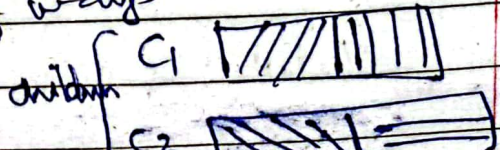
ak random
ak such toppers

- apply genetics ops to create offsprings

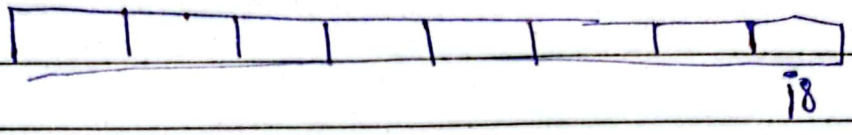
crossover
mutation



one of the ways to create, mix children



Date:



Item	Value	Weight
i_1	v_1	w_1
i_2	v_2	w_2
⋮	⋮	⋮
i_8	v_8	w_8

$$\sum_{i=1} w_i \leq w_{max}$$

TSP is a description of problem not only for actual travelling salesperson

fitness (CS-1)

$$= \frac{v_1 + v_4}{w_1 + w_4}$$

CS-1 [1 | 0 | 0 | 1]

$w_1 + w_4 \leq w_{max}$

Can generate illegal solution as well.

CS - 1

⋮

CS-10

- Validate every x_i time

(To remain legal chahiye, toh illegal ko skip & generate next till requirement fulfill)

Date:

$S = 2(b)^{d/2} - 1$ even

$S = b^{(d+1)/2} + b^{(d-1)/2} - 1$ odd

Genetic Algorithm (P(t))

representation of solution

$P(t) = \{ x_1^t, x_2^t, \ldots, x_n^t \}$

genetic_algo (P(t)) {

→ t = 0

→ initialize P(t) [initialize yahan hoga framework

while termination criteria isn't satisfied { (generation per iterate karenge)

- evaluate every x_i^t (find its fitness)
- select a pool of fit solutions (parents)
- apply genetic operators to create children (offsprings)
- Replace P(t) with some offsprings

} - t = t+1

}

dk if were reached Solution

knapsack mein soln is array of 1s, 0s of len = amt of object.

knapsack problem

bool vals

initialized randomly

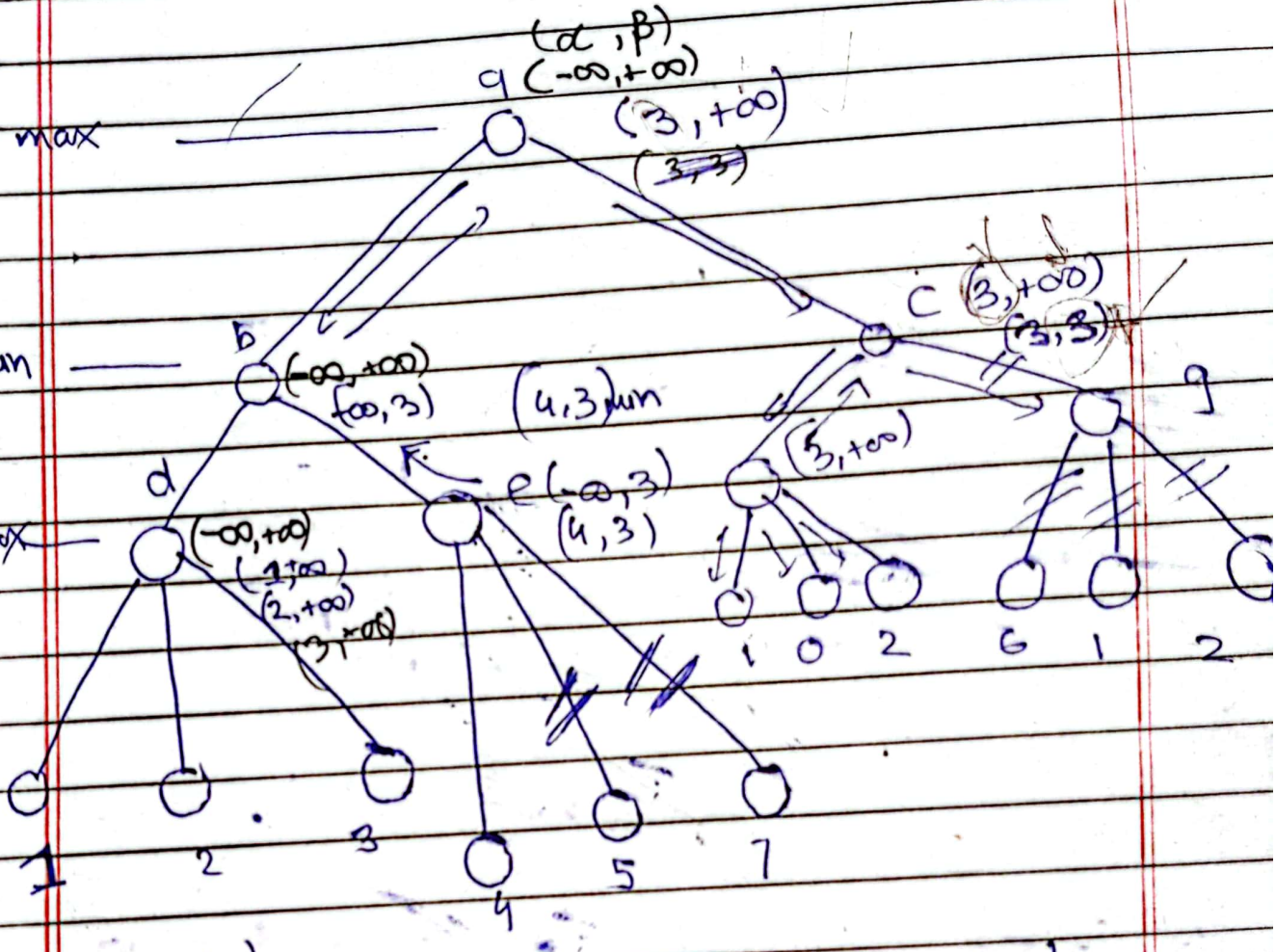
x_i^t = | 0 | | 1 | | | . | | | 1 |

i_1 i_2 ... i_n

Date: 24/10/24

AI

↙ generated on the fly



max (-∞, 1)

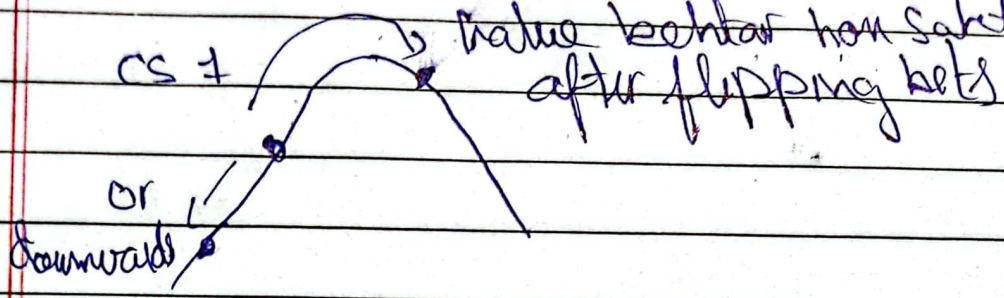
ply depth
wo max level
jahan tak aap
apna game
tree generate kar
saktay

6 nodes haven't been explored

$d \leq \beta$

Date:

3.b) Mutation
→ flip bits



can discard some

gives good result with such simplicity
TSPs ka 4 generations
(iterations)
until you meet the termination criteria.

Solve alphabeta pruning, Monotonicity
bhi aayega ji.

=Chromosome theory
(read)
John Holland

Date:

2) Fitness criteria

$\frac{V_1}{W_1}$ $\frac{\text{value of sol 1}}{\text{weight of sol 1}}$

$W_1 \leq W$ ↑

jiska value ziyada wow ziyada acha ratio achi toh wow acha

$\frac{V_n}{W_n}$ $\frac{V_8}{W_8}$ $\frac{}{110}$ illegal val ×

3) Improving Solution Quality

→ isman randomness kaafi, but there is a method to randomness | random search
→ better than random
→ make changes in existing solution

operators:

3(a) crossover
→ mix solution keh pieces

take 1 part from parent 1
& other part from parent 2

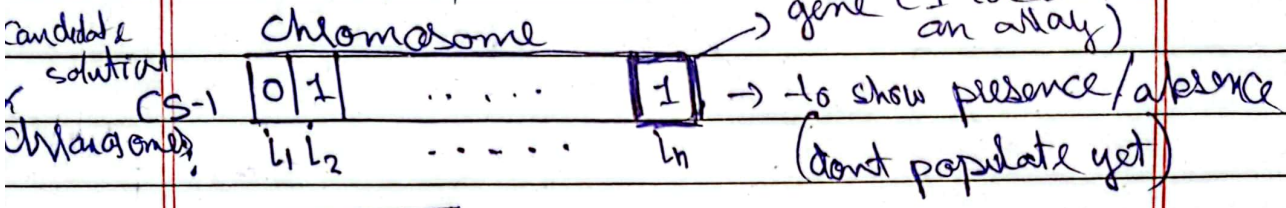
→ adaptability
→ inheritance mein usko intelligence mili gayi
→ inheritance is biased
→ fitter parents, fitter kids

Date:

Genetic Algorithm

→ inspired by evolution of living things.
→ not just biologically inspired (like neural network)

1) Representation of problem solution (what is the formal of solution? X, Y, Z)

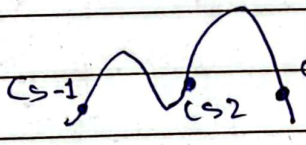


CS-n → we look at the army of solutions.

One solution isn't proposed, a set of solution is proposed.

timetable ki sol kasi? table, visual representation

baghair sochay samjhay initialize karein

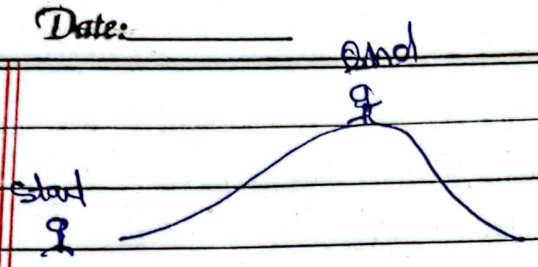


instead of following one path, we follow many paths!

initialized
so many solutions,
parallel computing
survival of the fittest

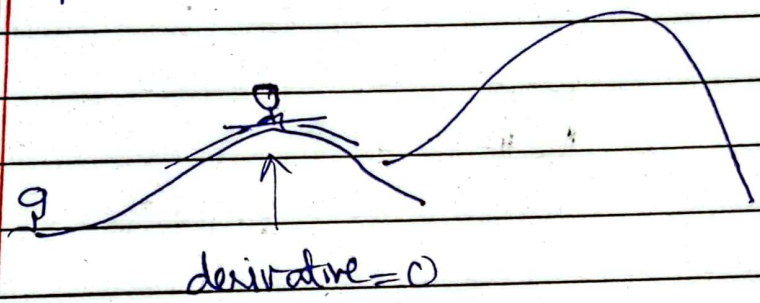
represent n solutions

Date:



hill climbing
one hill
↳ greedy approach

You can always climb a hill (single) using greedy approach but not optimal for more than 1



kya optimal timetable hai? we can try our best, it is NP hard problem
↳ factorial complexity

TSP: har city ka aik data visit kar kay apnay starting node pe terminate kar ga, and best wali hoga jo min cost.
⇒ can't guarantee optimal sol.
→ $(n-1)!$ time complexity
→ circuitboard pe bhi apply hota.

minimum spanning tree (like TSP) but doesn't end on start node

Date: 22/10/24

quiz 2
Thursday
24th Oct
morocity
seh owstad
till alpha
beta pruning

Genetic Algorithm

read 12.1 article book

→ not for normal problems, it is for NP hard problems

Knapsack problem

optimization problem

knapsack hai articles isstarhan rakhna keh capacity mein rehtay huey rakhein ya bag mein itnay wazan ka rakhna ya usssay ziyada nhi. Bag nhi bharab karna.

→ 0-1 knapsack problem

yes or no, lein gaey ya nhi. Apni space ko optimize karna

→ Fractional knapsack problem

- empaphous things (no shape)
- → You can add things in fractions as well, not necessarily complete (powdered form)
- → can be solved using greedy approach (optimal sol (best sol) for knapsack 0-1)
- → powder, dust
- → optimal sol by greedy approach → hard

0-1 knapsack can be solved using greedy approach but not optimal, but fractional knapsack is optimal by greedy approach

IBM deep blue programmer

Date:

kitni nodes by mini max in ex & alpha beta pruning?

26 15 $\frac{15}{26}$

best case : $S_E = 2b^{d/2} - 1$ (even ply depth)

↓

no. of nodes examined / checked / generated

$S_O = b^{(d+1)/2} + b^{(d-1)/2} - 1$ (odd ply depth)

$O(b^{d/2}) = O(\sqrt{b})^d$

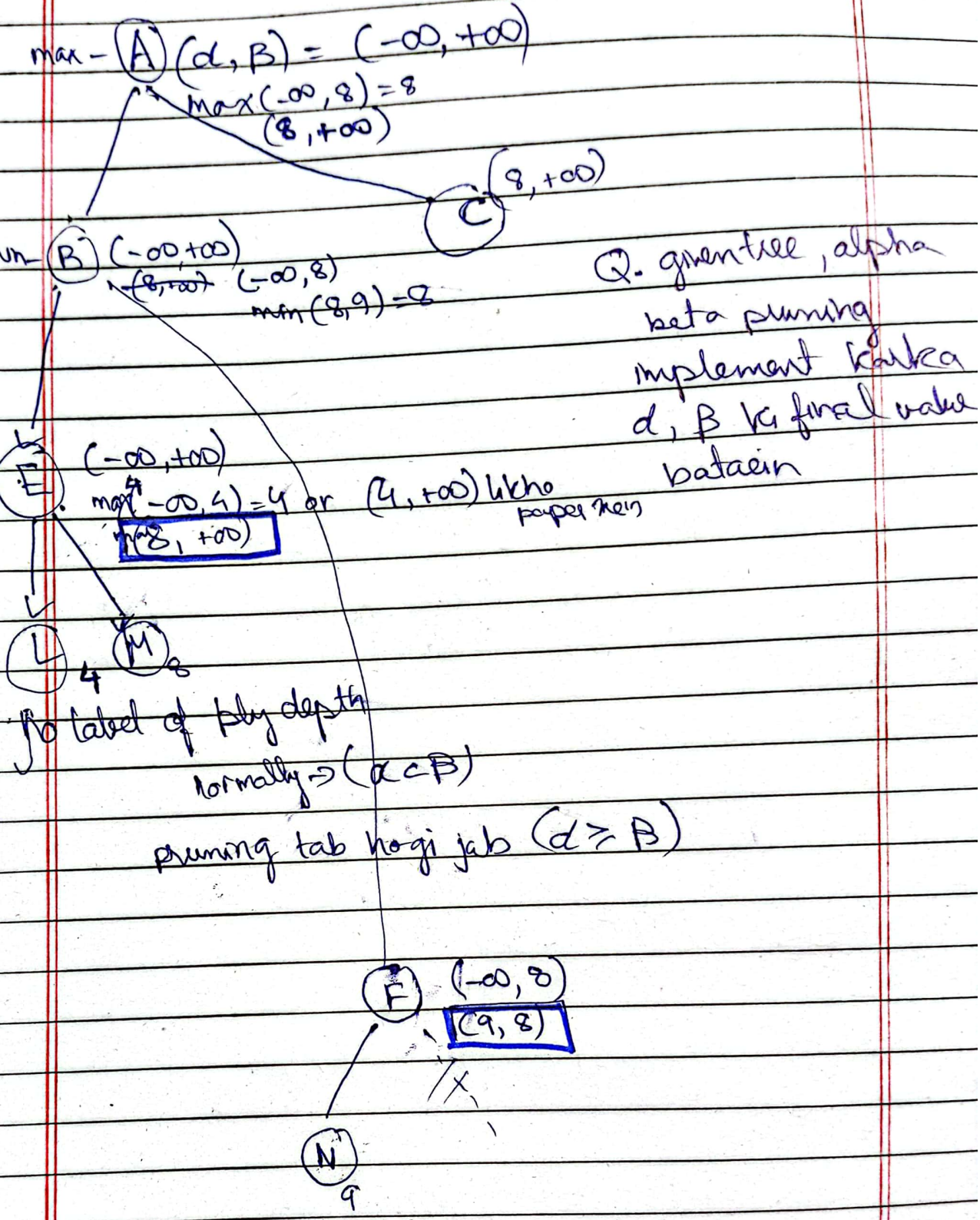
$S_E = 2b^{d/2} - 1$

$S_O = b^{(d+1)/2} + b^{(d-1)/2} - 1$

best case $= b^{d/2}$

worst case $\frac{1-b^{d+1}}{1-b}$

Date:



Q. given tree, alpha beta pruning implement karke α, β ki final value bataein

No label of ply depth

normally ⇒ (α < β)

pruning tab hogi jab (α ≥ β)

~~not generated~~

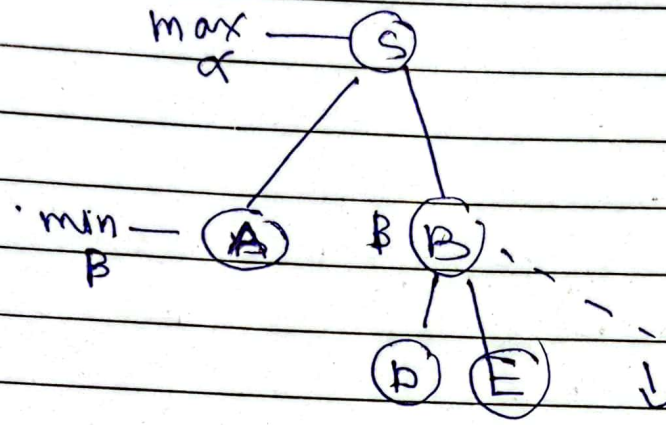
⊙ (just for reference)

$O(bd) \rightarrow O(\sqrt{b}\,d)$

abhi bhi linear hi hai

Mini Max ko alpha beta pruning seh kesay implement karein

(S) root node keh saaray bachay generate hota, since wo leaf node nahi thaey.



↓ wasn't generated since it was 20 already, min less than A's min

$\alpha(i) \geq \beta(j)$ pruning condition

min-nodes. $\beta(j) \leq \alpha(i)$ α-cutoff

max-nodes: $\alpha(j) \geq \beta(i)$ β-cutoff

alpha cutoff on min nodes
beta cutoff on max nodes

associate alpha, beta for each node

α → inc monotonically (non-decreasing)

b → dec monotonically (non-increasing)

Date:

Singular extension heuristic

→ aik individual node ko further down tak check kartay.

* playdepth kesay kartay & tracking kesay karna

Mini Max is exhaustive search
agar bfs way mein chalatay tah
mem $O(b^q)$
time $O(b^q)$
agar dfs way:
mem: $O(bd)$
time $O(b^d)$

Node S kah perspective seh dekhna:

(B) ko further explore karna useless, since wrna adversary ka node, wrna hamesha apko kam pe hi le kar aaye ga.

A → min is 100
B → min 20 → dont explore further.

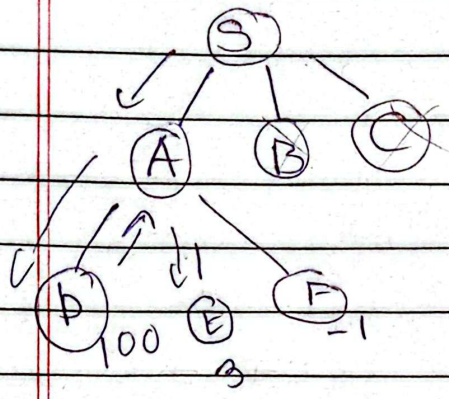
deepblue: Bseh further dekhta tha.
5 depth level
alpha beta pruning
max min

Date: 17/10/24

AI

agar sirf aik bacha daalein stack mein & explore to depth, then baad mein sibling generate karein.

* know the diff b/w this func logic & DFS ka normal logic



A = min (100, 3, -1) = -1

B = min (G, H) = 5

C = min (2, 9) = 2

→ B is best, maximum advantage

plywood

Faida & Nuqsan of More ply depth:

→ more memory usage

→ Zyada achi analysis

→ time zyada lga

horizon effect

→ can't see beyond the ply depth

Quiescent states: jismein aik dum seh pieces kam ho jayein ya ziyada

Date:

M T W T F S

if $g(n) = 0$

$f(n) = h(n)$ best first search

jon sirf distances ki basis pe chaltay, un algorithms ko **greedy algorithms** kehtay hain.

irrevocable = hill climbing algo

greedy algo cannot guarantee optimal path (they don't back track and are irrevocable).

(Syllabus of quiz 1)

mini max algo

ch 4

Date:

To find 'G_2' first

$f(n) > f(G_2) = C_2$ — (ii)

$C_2 < f(n) \leq C_1$ (from (i) & (ii))

$\Rightarrow C_1 > C_2$ (by transitivity)

But $C_2 > C_1$

∴ our assumption is contradictory.

∴ Therefore A^* is always optimal.

$f(n) = g(n) + h(n)$

If $h(n) = 0$

$f(n) = g(n)$ → if not constant it is uniform cost search / Dijkstra's algo / Branch & bound

If constant it is BFS

Date: M T W T F S

rule of induction:

deduction is based on logic, so it can't be wrong.

flawed reasoning

- we can infer, it is all about ML, and it is all about capturing intelligence

best-first search is a framework heuristic O kardain to bfs ban jaye ga.

→ jiski cost zyada

Optimality of A* (formal proof)

proof by contradiction:

Assume that A* finds a suboptimal goal before it hits the optimal one.

$f(G_1) = C$; $f(G_2) = C_2$ (suboptimal) $\Rightarrow C_2 > C$

$f(n) \leq C$ (by admissibility) → (1)

$h_1(n) \leq h_2(n) \leq h_3(n) \leq h_4 \leq h^*(n)$

(0) euclidean

Date: M T W T F S

~~h_2 h_3 h_4~~

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

underestimate ← → overestimate

no. of nodes explored increases.

↳ koi bhi solution explore kar saktay, but kam nodes ziyada bhtar toh seh sel $h(4)$

Infomedness :

↳ it is more infomed

$h_1(n) \leq h_2(n)$

↳ bigger is better in the underestimate

memory takes less space

HW:

Q koi aur heuristic bana saktay jo manhattan distance seh bhi ziyada informed hou?

Date:

M T W T F S

Admissible {oracle} bata rahaa

admissibility: {yehmatlab min. cost actual path ki / actual cost

$h(n) \leq h^*(n)$

estimate

heuristic hamesha underestimate hoy.

→ It is always an underestimate of the true (actual) cost to the goal.

overestimate kyun nhi kartay?

kaplon ki sale

underestimate hoy	overestimate hoy
toh ziyada sastay kaplay le saktay	toh ~~gol~~ nehnga hi milega

how will we know our heuristic underestimates?

problem relaxation problem

constraint ko follow kartay huay

heuristic underestimate hi hogi
heuristic design
problem ki restriction ko follow ya relax karkay woh admissable heuristic.

Date:

M T W T F S

5, 5, 5, 6

↑

top

agar h(n) zero toh hum goal state pe pohunch gaye.

f(n)

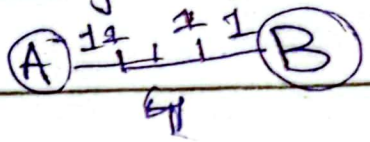
total evaluation function

problem relaxation problem

Date: M T W T F S

bfs, dfs → for unweighted graphs

but way around?

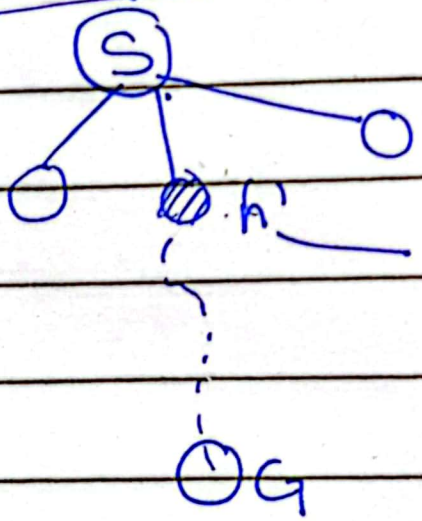


weighted graphs → in assignment

cost in 8 puzzle: moves

Ignore the weights

A* algo:



g(n) = 2

h(n) = 4 → eg (estimate

f(n) = 6

g(n) + h(n).

→ agar always equal to zero

toh A* algorithm degenerate hai

Ka bfs ban jaye ga

Date: A1

M T W T F S

Lecture #5

min max algo

A* algorithm (Best first search + admissable heuristic)

→ most optimal path

→ complete algorithm

→ Competitor to Dijstra's algo

→ uses admissable heuristics

$f(n) = g(n) + h(n)$

$g(n)$ → true/actual cost from start node to 'n'

$h(n)$ → estimated cost from node n to the goal

A algorithm

→ (best first search + not admissable search)

= A algorithm

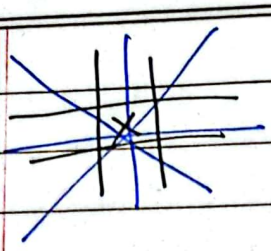
~~→ selective search~~

- A star is optimally more efficient than Dijstra's algorithm
- $f(n)$ → gives total evaluation gives estimate of distance b/w start state & goal state

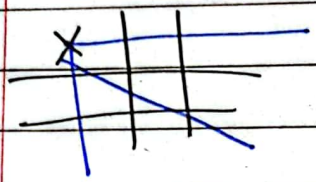
start state: $g(n) = 0$

goal state: $f(n) = 0$

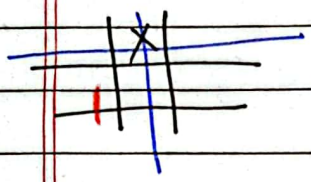
Date:



④ possible successes



③ possible successes



② possible successes

The no. of winning lines you have available

vs

The no. of winning lines your rival has available

agar deficit mein toh select the less deficit ones.

Yeh heuristic negative values bhi de saktay.

quiz next lec: ch (3-4), slides

next sem next: 2 lab → heuristic function use along with priority queue.

→ Blind search implement karna hoga next lab mein.

Assignment

imp → output of pseudocode, Code, description, sab aa'sakti quiz mein later on

A* → best fs, always be an underestimate

greedy → won't normally backtrack

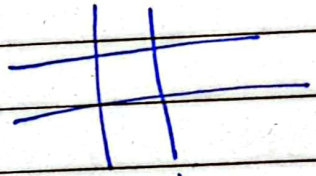
Date:

obstacles zigyada hi hon
trail will be zigzag.
→ Optimal path kyun hoga? kyun
keh wo hamesha heuristic
underestimate dega, robot will try
since uska heuristic hamesha hi
underestimate hoga.

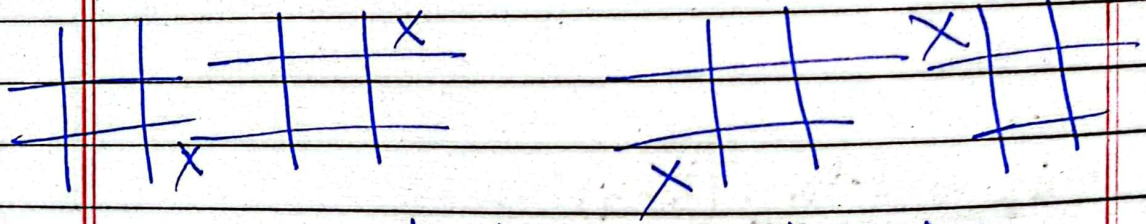
Zero-Sum games

→ profit at the expense of someone
else's loss.

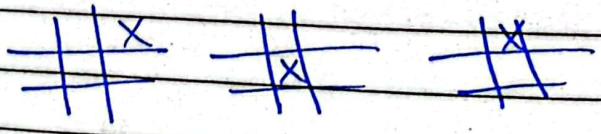
tick tack toe



states 9! ban sakti at max
but ismein 2 log hain, alternatively
dono chalein gaey, toh it is not
the same as 8 puzzle problem



yeh chalein barabar hain,
since we only need to complete a
line of 3.



3 possible states

Date:

Best-First Search

→ priority queue used here.
priority kiski?
→ from heuristics, the number comes.
• kesay implement kartay priority queue?
tree, heap, min, max heap.
• linked list seh bhi kar saktay, make priority, sort on the basis of number associated.
• top per hoga woh pehlay aaay ga

~~ab~~

Comparison with Uninformed search:

→ won't exhaust all possibilities
→ won't be exponential in memory
→ time in the worst case, linear at worst is exponential unfortunately
→ ~~optimal search?~~
→ optimistic path deti, ~~ap~~

heuristic apka underestimate hona chahiye

Robot in a maze

heuristic kesay banaeyen gaey?
→ distance (sld) maintained rahay
→ jis state mein kam distance hoga
us state mein jaye chahe

Date:

2	8	3
1	6	4
7		5

1	2	3
8		4
7	6	5

out of place tiles = 4

Somehow 4 moves keh distance per hain

variable → nested loop in x direction & in y direction

underestimate of the actual # moves

heuristic function

blank is not a tile

heuristic

$h(n)$ = sum of distances out of place = 5

manhattan distance

$x_1 - x_2$ coordinates

$y_1 - y_2$ coordinates

Euclidean heuristics distance

$$d = \sqrt{(x_2 - x_1)^2 + (y - y_1)^2}$$

+ve real number

euclidean will be lower than manhattan distance

underestimate

$e_n \leq a_n$

estimated number of moves

actual moves

Date: 3/10/24 AI Lecture #5

(till today syllabus) quiz in next lec
descriptive quiz & assignment
mcqs

ch4 → heuristics search

rule of thumb → rules jo overtime seekhay

→ doctor keh paas jaatay wou has test paform nhi kaltay, mechanic bhi saari ~~dekhtay~~ possibilities nhi dekhti.

rule of thumb → ml hi hai, machine khud seh pick up kaati.

Once you provide it with heuristic & provide a method to use it, it will do selective search.

$h: b \longmapsto \mathbb{R}$

board ki state ko real number per map kati

* be able to write a heuristic function.

heuristic → is a rule of thumb.
→ it is not supposed to be right always, it can make mistakes.

Date:

$N = b^0 + b^1 + b^2 + \ldots + b^d$

$N = (d+1)b^0 + db^1 + (d-1)b^2 + \ldots + 2b^{d-1} + 1b^d$

IDS

kitni baar generate hua hai

$N_{DFS} = b_0 + b^1 + \ldots + b^d$

→ sirf constant ka farq hai tab itna koi farq nhi hai

$N_{IDS} \rightarrow O(b^d)$ → time complexity

$N_{BFS} \rightarrow O(b^d)$

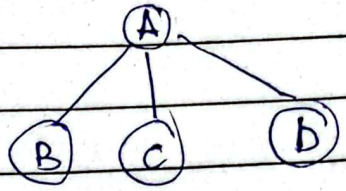
time ∝ no of nodes

IDDFS → DFS hi hai but level save nhi karnay, have to store depth info global variable hoga jo persist karay ga across iterations.

Heuristic search

hill climbing
gradient descent
gradient ascent

Read physical symbol system hypothesis
analytically can't prove
→ impirically prove ho rahi

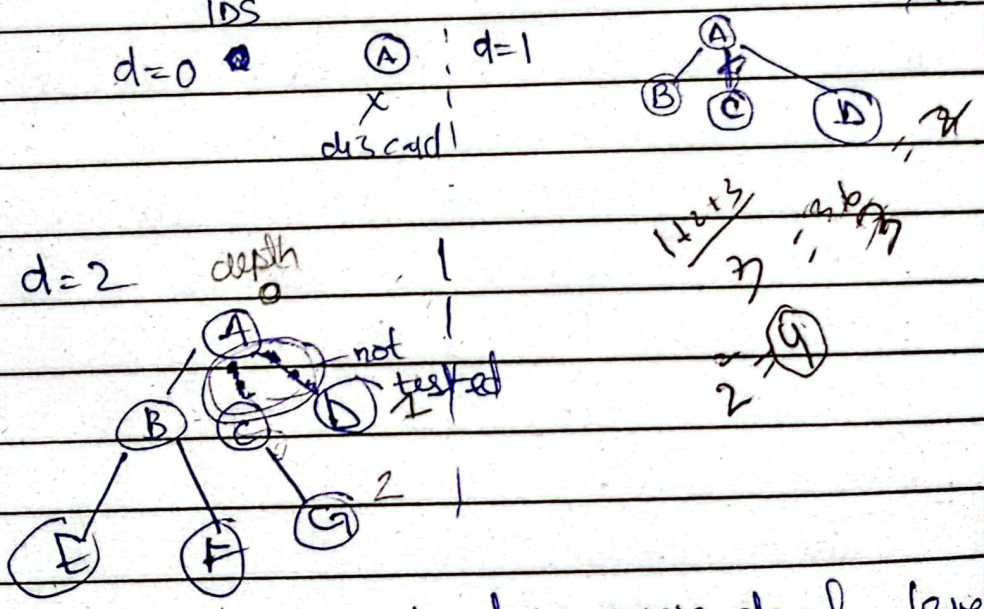


generate & test goal hai kahi bhi

		Space complexity
DFS	incomplete search	linear vs exponential
BFS	complete search	

DFID / IDS — depth-first with iterative deepening / iterative deeping search

	Space	Time
IDS	O(bd)	$O(b^d)$



har iteration mein chal level by level hi rahay hain

dobara kyun bana rahay? wasteful nazar aa raha hai but we want to avoid storing / occupying extra memory

Date:

every problem is not solvable (halting problem)

DFS is very ~~time~~ memory efficient (linear behaviour)

BFS seh path dhoondnay keh liye predecessor list seh kartay rahen.

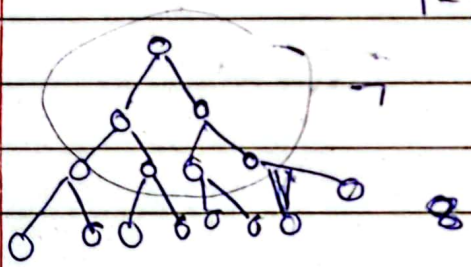
kitnay nodes hain?

$N_T = N_0 + N_1 + \ldots + N_d$ no. of nodes at depth ...

$N_T = b^0 + b^1 + b^2 + \ldots + b^d$

$N_T = \frac{1 - b^{d+1}}{1 - b}$ total no. of nodes in a tree till depth 'd'.

$\frac{1 - 2^{2+1}}{1 - 2} = \frac{-7}{-1} = 7$



chess → 35 branching factor bht no sahi

aik level add kardetay toh pichlay saaray levels keh sum seh ziyada nodes hain.

No. of nodes grow exponentially.

BFS → hamesha optimal output degi.

You have not left any stone unturned.

BFS is a complete search, agar ho ga goal toh BFS dhoond lega

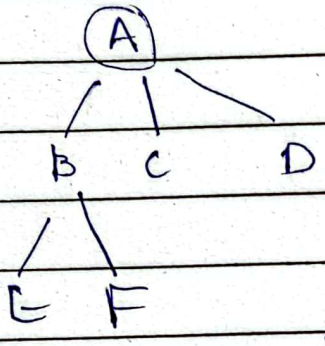
DFS, agar levels ziyada toh guarantee nhi hai keh goal mile ga ya nhi, kam levels mein no issue

Date: 1/10/24

AI lecture #4

Informed

backtracking search is essentially dfs search



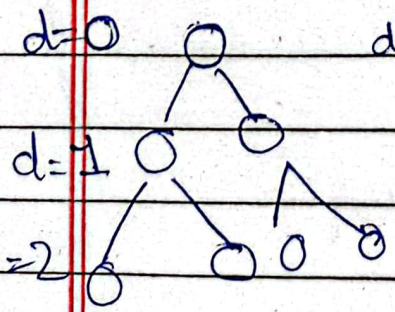
open	closed
~~A~~ ~~B~~ C D	[A]
~~A~~ ~~B~~ C D E F	B A]

branching factor

$\frac{3+4+5}{3} = \frac{12}{3} =$

avg branching factor

		Space	Time
BFS	(optimal complexity)	$O(b^d)$	$O(b^d)$
DFS	(Memory efficient)	$O(bd)$	$O(b^d)$



depth level

Date:

Historic person hai aapka ancestor, which direction will you check? From you to historic person or historic person to you?

→ First one as bf sirf 2 hai, while in the other case it could be too much, unkay descendants bohut hotay hain toh agay bf bohut zyada ho sakta.

2 types of search:

1) data-driven search (forward chaining)
2) goal-driven search (backward chain

→ disease diagnose, saaray test kaiwaa, dekha kya hua wa

→ starts off with the goal, kis tareeqay seh goal password.

Doctors use goal-driven reasoning. Kya malaria tohnhi? Symptoms kya hain?

→ less time consuming

Branching factor: aik node keh kitnay on average children hain.

→ Baaz problems mein 4 states like 8-puzzle problem.

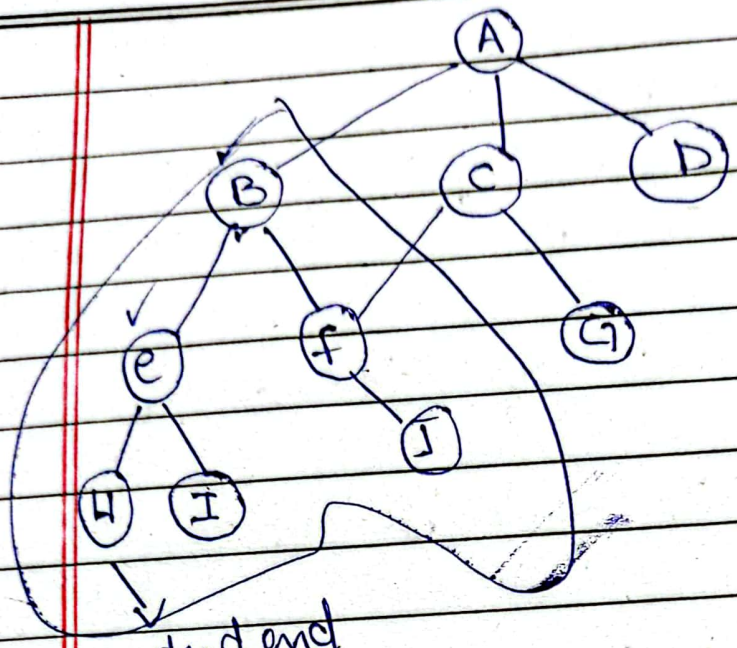
avg bf: $\frac{4+3}{2}$

goal driven

Kya aap Start state seh goal state per jaana chahen gaey ya goal State seh start state jaana chahen gaey? Dono hi same in case of 8 puzzle problem.

Date:

Search tree



Directed graph with repetition?

C mein jaakay aap F create karein gaey but wont add it.

A → C → G

NSL, SL → stack

Hamein graph diya jata hai, toh kya yeh graph hamein bhi pehla bana banaya milta?

→ Nhi hamnay graph nhi banaya, explicitly nodes create nhi kiye, implicitly define kar rahay graph.

→ Actual creation nhi kartay graph keh, possibilities ka tree bana rahay, and checks lagatay.

Invision as a state space graph

Date:

In algorithms where you can backtrack it is called an revocable algorithm.

BFS → complete algo
DFS → uncomplete algo

Backtrack, DFS → behan bhai

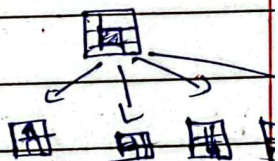
Systematic → not random

Backtracking

3 lists that keep track of algo:

1) SL (state list)

2) NSL (new state list)



3) DE (dead ends)

4) CS (current state) pointer

node under examination

we check yeh goal hai, pointer hai yeh

start state mein multiple nodes bhi hou saktay.

Date:

Blind Search Algorithm

→ Systematic search karni, problem-solving karni without explicit problem solving knowledge.

→ Brute force hi ho jati

→ Uninformed search

Heuristic : explicit problem solving knowledge

4 algos:

1 & 2) Backtracking & DFS
→ mein path bhi store kar rahay hotay but dfs mein nhi karta.

3) BFS
→ orthagonal

4) DF with iterative deeping

Search strategies:

(i) Complexity O( )
search algorithm maximum kitna time ya space le sakta

(ii) Completeness
Agar goal exist karta, complete algorithm guarantees keh goal achieve karay ga.

(iii) Optimality
Algorithm that returns best solution, best is what you define, like in terms of cost, moves according to your criteria.

(iv) Irrevocability
An algorithm in which when you can't undo or backtrack.
eg: Hill climbing search (weight search edge)
Is irrevocable

Date: 26/9/24

AI Lecture #3

2 things required for intelligent problem solving:
1) representation (Knowledge representation) KR
2) search

you need a formal structure, chahe structured, unstructured db, but you need it in some form.

→ sequence of action we perform to reach our required goal.

In 2 cheezon ko interplay algorithms kehlatay.

Random search → koi random permutation aa jaye aur woh goal state toh search random

Systematic search → hamyeh karein gaey

DFS (page by page search)

heuristic & DFS in nature

Kabhi kabar insaan random search bhi karta

Weak AI follower -> sir

Itni bhi AI taraqi kalay, insaan hamesha upar hai

Physical symbol system hypothesis:

Physical symbol system eg computer has necessary & means for sufficient means for general intelligent action.

Modern AI

8 puzzles

3x3 puzzle hai toh kitni 9! = combinations possible. there are possible 9! permutations.

OOP object class -> puzzle state ki

graph mein nodes & edges hain search keh andar implement karna

When you search a graph:

you actual develop or generate on the go, a search tree

control repetitions through parity checks.

phir ismein intelligence kesay? exhaustive hai yeh, toh aap heuristic search karein gay,

Date:

physical system hai yahan pei?
↳ computer is a physical symbol system.
→ algebra, logic
→ input hain, output hai

→ insaan aik ~~system~~ physical system hai.
but kya physical symbol system hai? insaan?

Insaan aik physical system but kya woh aik physical symbol system hai?

neurologist have shown keh kesay memory store hoti, konsay regions kya store kartay, language area.

Brain cannot feel pain, surgery brain ki jaagay huay hoti hai

→ Insaan physical symbol system hai.

emotions kya?

There is more to humans then physical symbol system that is a sense of spirituality.

Date:

Sentience

Weak AI

John Searl

Mind, brains, program — rebuttal 1980 turing test

↓

chinese room,
doesn't know chinese, but resources hain, work karna samjh kar karta.
He thinks kabhi bhi strong AI nhi ho sakta chatgpt bhi pattern patching kati

Rational bomb
ATM | SE time
error
peechay banda bithaya paisay denay.

System

Physical Symbol Hypothesis

its just a hypothesis, a statement that can be proven true or false

Intelligence = KR + search (for solutions)

1. using symbols

symbol → nishaan / representation
koi language use karlein

Date:

Kya turing test reliable hai?

intelligence cannot be defined specifically

self-awareness kya hai?

chatgp keh context mein sensxious?

Sentience - cognitive state jis mein aapkay pass self-awareness hai.

STRONG AI

Machine = Human

Sentience

meaning self-awareness hai, aapno environment, apnay baray mein jante.

WEAK AI

AI = helpful tool for humans

dirty bombs

smart bombs

Date: 24/9/24 AI Lecture #2

AI - big body of knowledge

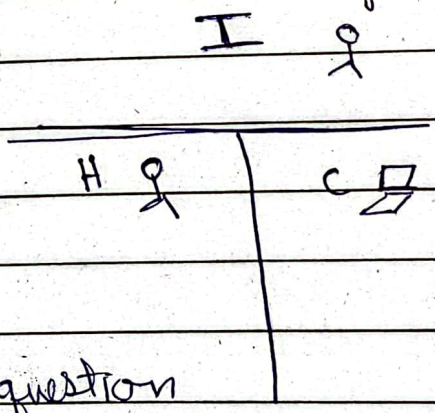
Ch3 → seh start

AGI

node 2D array class

AI ko computational point of view seh path rahay.
main focus : understanding
no definition in exams like AI

Turing Test kya tha?



text ke question
likhna, aur text per reply.
Pehchaana keh computer kousa hai.

Alan turing :

Machine ko pehchaan nhi sakun
gay, toh yeh intelligence hai

Date:

clustering reinforcement learning
practical applications.

Is chatgpt really intelligent?
- how can we check
turing test
judge keh human keh maching, by answer
test

Dr. Hugh Lubner

LLMs aanay keh baad that turing test was done.
There are many other turing test.

movie: A beautiful mind

ch2, 3 sah baat shuru hoti course mein
DFS, BFS karkay aao

19/8/24

Advanced Stats

Lecture #1

4 assignments:
research paper relevant etc

Date: _____

chair
kursi kursi, how do you classify?

how you do that? you dont know

how machines does that? you should know

yeh billi kesay hai?

Alan turing
John Von Neuman

Arthur Samuels checkers program 1958? C

Perception couldnt do XOR (at that time?)

back propogation algorithm MLP

ask neuron
ask perception or many

Will shift to machine learning in this course
in this course? → conventional machine learning
→ conventional AI

Deep networks 

feature engineering
kursi per banda hai, kya yeh kursi hai?
no

Alex Net no need to specify banda
convolution network

watch: video of
AI chat bots can converse with each others

automated reasoning	evolution of approaches to learning
heuristic knowledge	not deep learning in this course.

Date: 19/8/24

AI

Lecture #1

computing core course

↳ CS-203 (3+1=4) sentience

15 marks → quiz

5 marks → sem proj

2+3

↓ assignments

class participation

prerequisites: OOP

objects not numbers but a combination

Characteristics of Intelligence:

→ problem-solving

→ learning to make better decisions

called AI (John Maccard) 1956 — list lost in stupid parenthesis

tentative: 1942 ↓ artificial neuron banana — 2024

ability to learn

AGI

e
n
e
r
a
l

Artificial Super Intelligence

Super Intelligence

Micen (expert system)