GATE

CRASH COURSE

DS & AI

Algorithms

Greedy Method (Part 01)

(Lecture 6)

By - Aditya sir

Topics to be Covered

1.
2. Intro
3. Applications
4.

Telegram Link for Aditya Jain sir:

https://t.me/AdityaSir_PW

About Aditya Jain sir

1. Appeared for GATE during BTech and secured AIR 60 in GATE in very first attempt - City topper
2. Represented college as the first Google DSC Ambassador.
3. The only student from the batch to secure an internship at Amazon. (9+ CGPA)
4. Had offer from IIT Bombay and IISc Bangalore to join the Masters program
5. Joined IIT Bombay for my 2 year Masters program, specialization in Data Science
6. Published multiple research papers in well known conferences along with the team
7. Received the prestigious excellence in Research award from IIT Bombay for my Masters thesis
8. Completed my Masters with an overall GPA of 9.36/10
9. Joined Dream11 as a Data Scientist
10. Have mentored working professions in field of Data Science and Analytics
11. Have been mentoring GATE aspirants to secure a great rank in limited time
12. Have got around 27.5K followers on Linkedin where I share my insights and guide students and professionals.

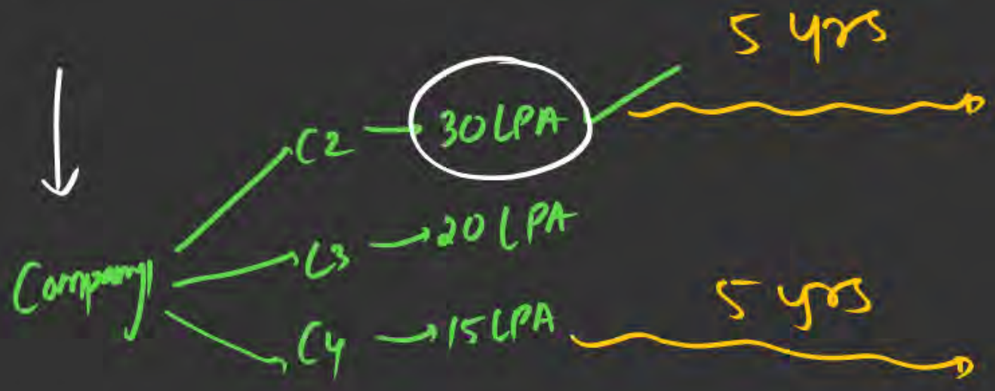
Topic : (Lecture Schedule)

3. Greedy Method

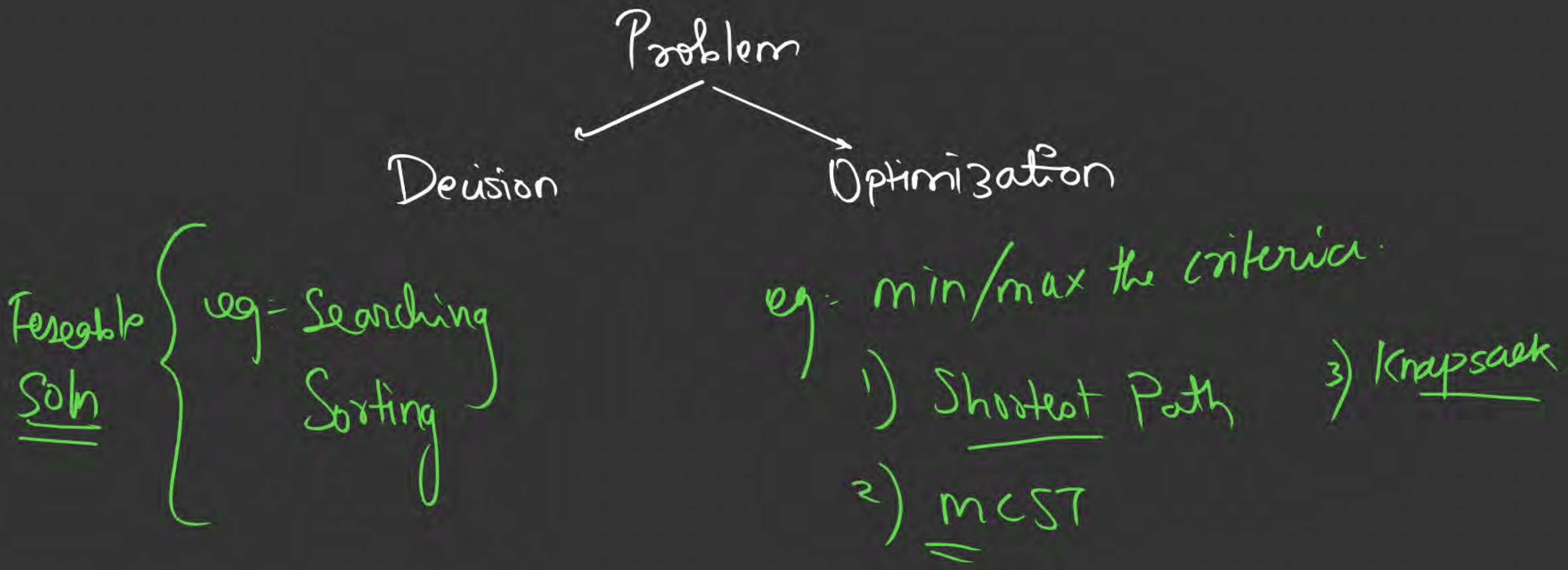
1. Introduction
2. Knapsack Problem
3. Job Sequencing with Deadlines
4. Optimal Merge Patterns
    1. Huffman Coding
5. Minimum Cost Spanning Trees (MCST)
    1. Prims Method
    2. Kruskal's Method
6. Dijkstras Shortest Paths Problem

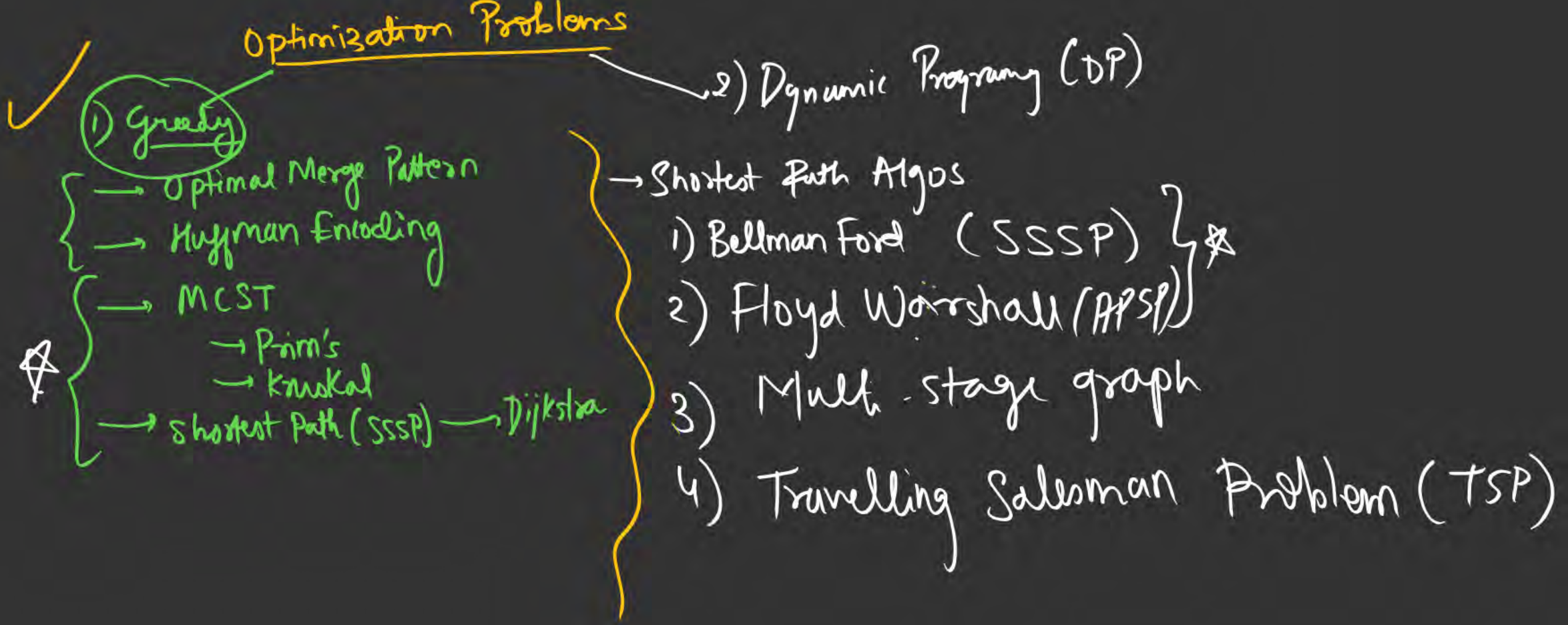
Merging

Greedy Algo:



local optimization


Problem
Decision
Optimization
Feasible Soln
eg:- Searching
Sorting
eg:- min/max the criteria.
1) Shortest Path
2) MCST
3) Knapsack


Optimization Problems
1) Greedy
→ Optimal Merge Pattern
→ Huffman Encoding
→ MCST
→ Prim's
→ Kruskal
→ shortest Path (SSSP) → Dijkstra
2) Dynamic Programming (DP)
→ Shortest Path Algos
1) Bellman Ford (SSSP)
2) Floyd Warshall (APSP)
3) Multi-stage graph
4) Travelling Salesman Problem (TSP)

1) Optimal Merge Pattern (OMP)

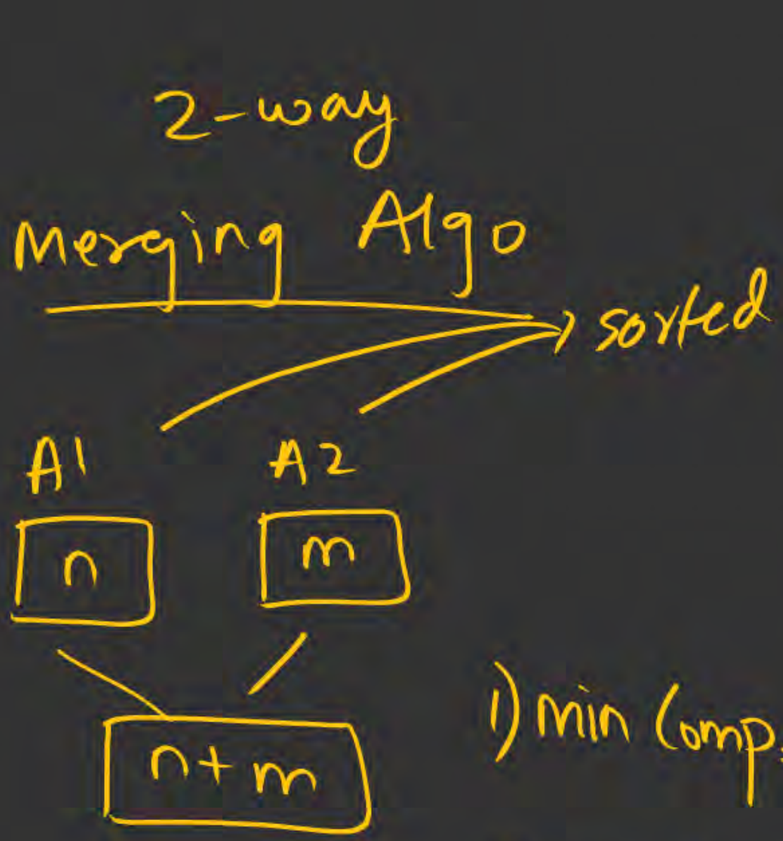
↳ Merging Algo Application

Problem Statement:

↳ Given some Sorted Files having few records each, we have to merge them into a single sorted file with minimum total Record Movements.

By default: Merging 2 files at a time

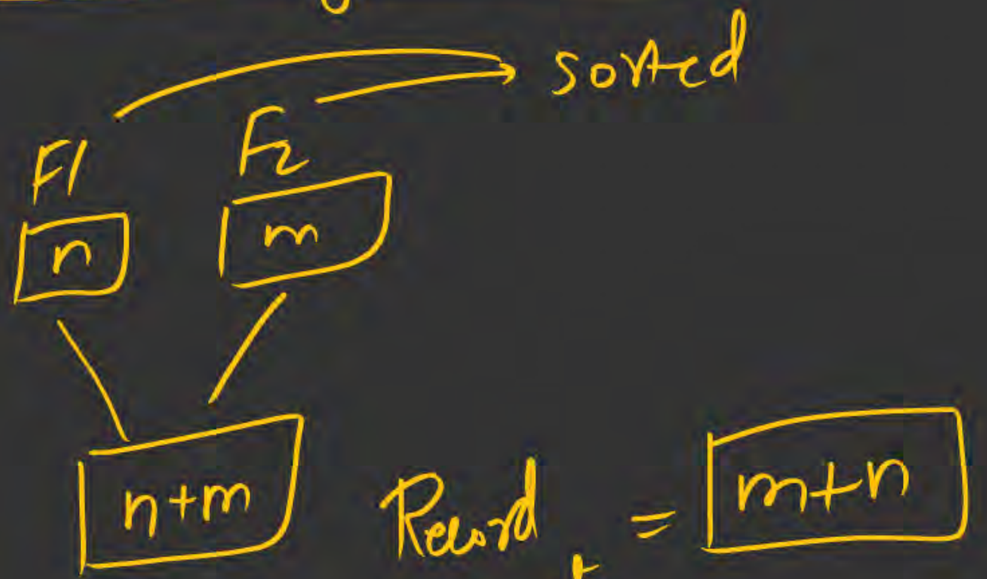
2-way Merging Algo



1) min Comp = min(n, m)

2) max Comp = m+n-1

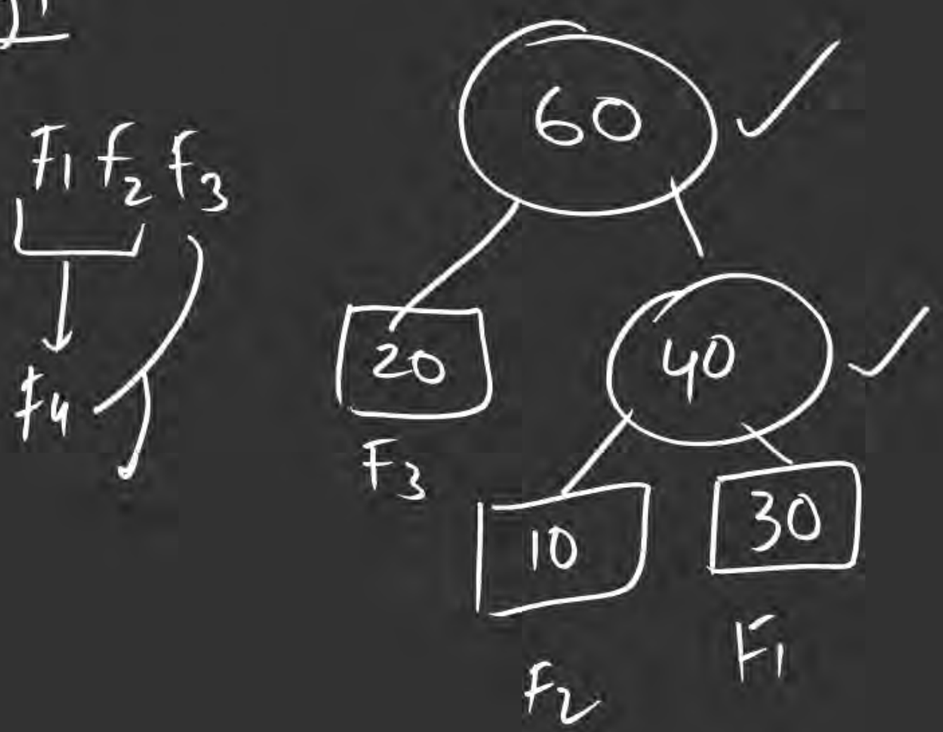
Optimal Merge Pattern



Record movements = m+n

eg: F1 = ~~30~~
F2 = ~~10~~
F3 = 20 , 40

way 1



60 X

Total Record movements
= 40 + 60
= 100

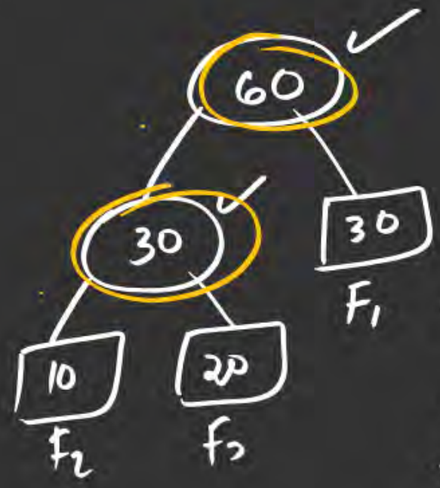
$f_1 = 30$

$f_2 = 10$

$f_3 = 20$

way 2

$f_2 f_3 f_1$



Total Rec movements

$= 30 + 60$

$= \underline{90}$

Total Record movements = sum all internal nodes $= \Sigma d_i q_i$

d_i = depth of File F_i

q_i = no. of Rec in File F_i

$= 2 \times 10 + 2 \times 20 + 1 \times 36$

$= 20 + 40 + 30 = 90$

optimal Soln;

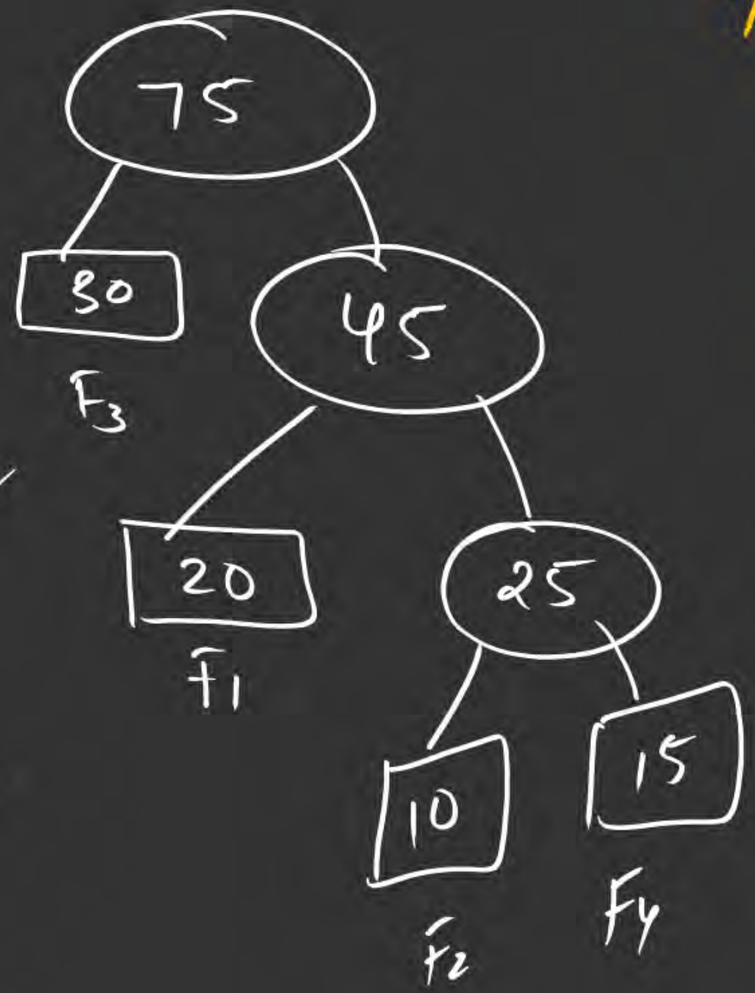
20, 25, 30

$f_1 = 20$ — 3

$f_2 = 10$ — 1 ✓

$f_3 = 30$ — 4

$f_4 = 15$ — 3 ✓



Total RM

$= 75 + 45 + 25 = 145$

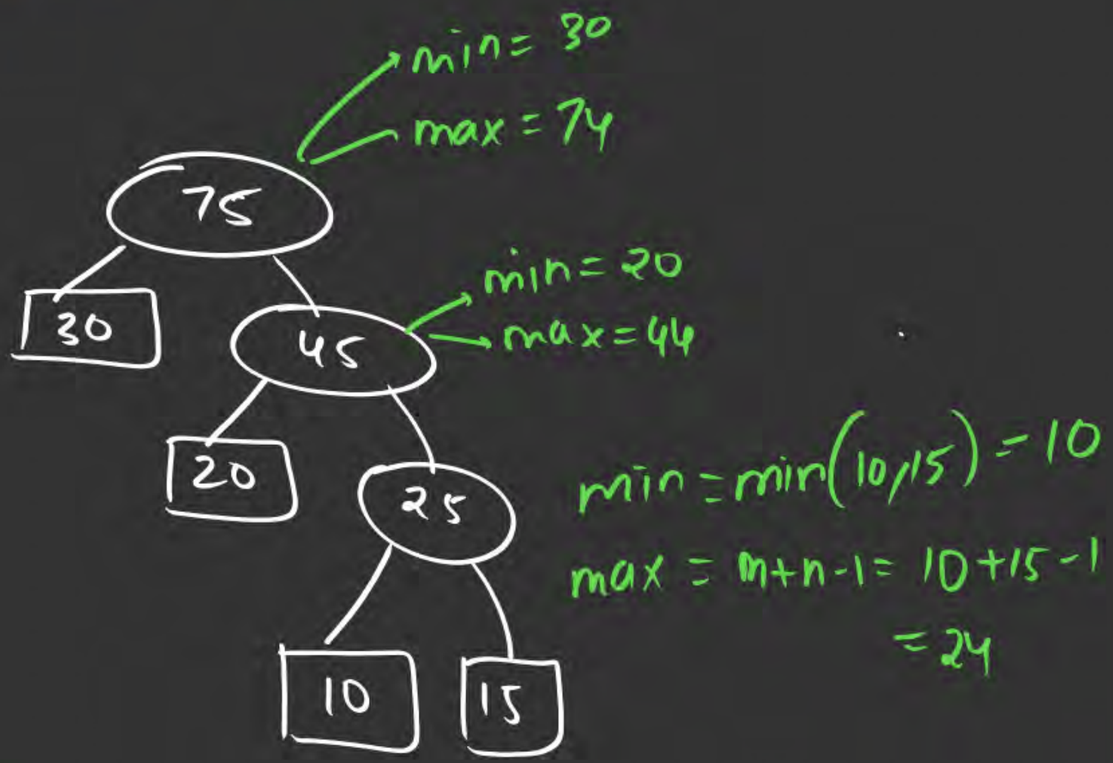
$= \sum d_i q_i$

$= 30 \times 1 + 20 \times 2 + 10 \times 3 + 15 \times 3$

$= 30 + 40 + 30 + 45$

$= \underline{145}$

Total Key-Comp in worst Case



$min = min(10,15) = 10$

$max = m+n-1 = 10+15-1$

$= 24$

Key-Comparisons

Total min

$= 10 + 20 + 30 = 60$

Total max

$= 74 + 44 + 24$

$= (75-1) + (45-1) + (25-1)$

$= 145 - 3 = \underline{142}$

Min Record movements?

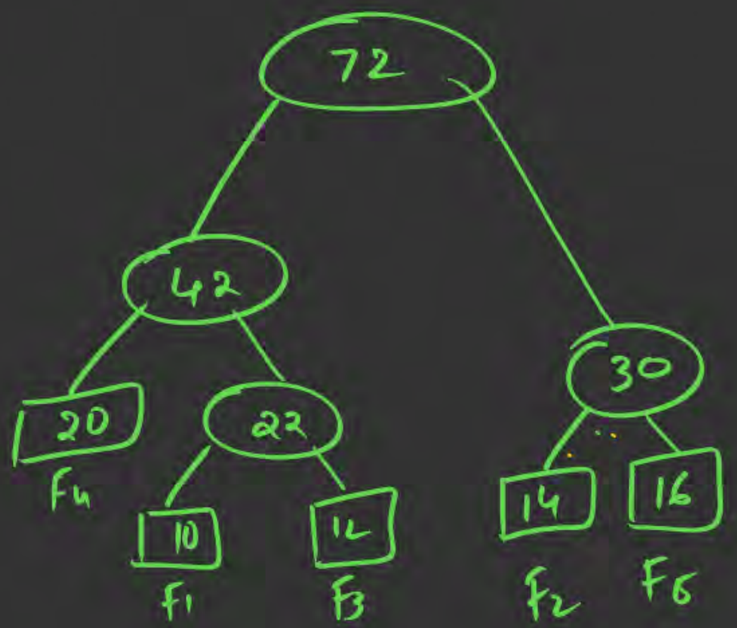
$f_1 = 10$ — (1) ✓✓

$f_2 = 14$ — (3) ✓

$f_3 = 12$ — (2) ✓✓

$f_4 = 20$ — (5)

$f_5 = 16$ — (4)



RM

$= (72 + 42 + 30 + 22)$

$= 72 + 72 + 22$

$= 144 + 22$

$= 166$ ✓

2) Huffman Encoding

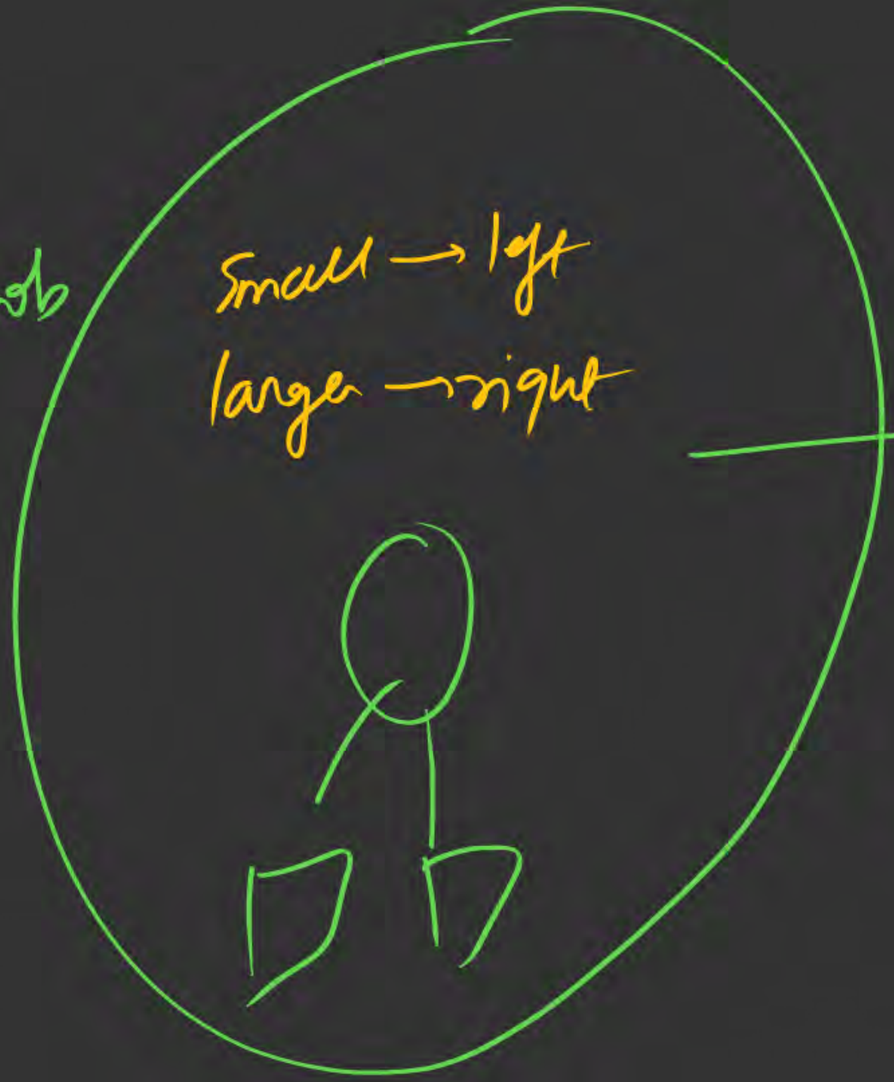
- non-uniform encoding
- Optimal

| uniform (Binary encoding)

2 bits →

11
10
01
00

Char	Freq/Prob
A	f_1
B	f_2
C	f_3
D	f_4



→ Encoding

left → 0

Right → 1

Question

#Q. A message is made up entirely of characters from the set X = {P, Q, R, S, T}. The table probabilities for each of the characters is shown below:

Characters	Probability
P	0.22
Q	0.34
R	0.17
S	0.19
T	0.08
Total	1.00

If a message of 100 characters over X; is encoded using Huffman coding, then the expected length of the encoded message in bits is _____.


P → 0.22
Q → 0.34
R → 0.17
S → 0.19
T → 0.08
1
0
1
0.41
0.59
0
1
0
1
0.19
0.22
0.34
0.25
S
P
Q
0
1
0.08
0.17
T
R
Encoding
P → 01
Q → 10
R → 111
S → 00
T → 110

Avg no. of bits/char

$$= \sum_{i=1}^{n} d_i q_i$$

(HW)

Avg no. of bits/char = Sum of internal nodes

$= 1 + 0.41 + 0.59 + 0.25$

$= 1 + 1 + 0.25$

$= \boxed{2.25}$ ✓

Total bits for a text with 100 " $= 2.25 \times 100$

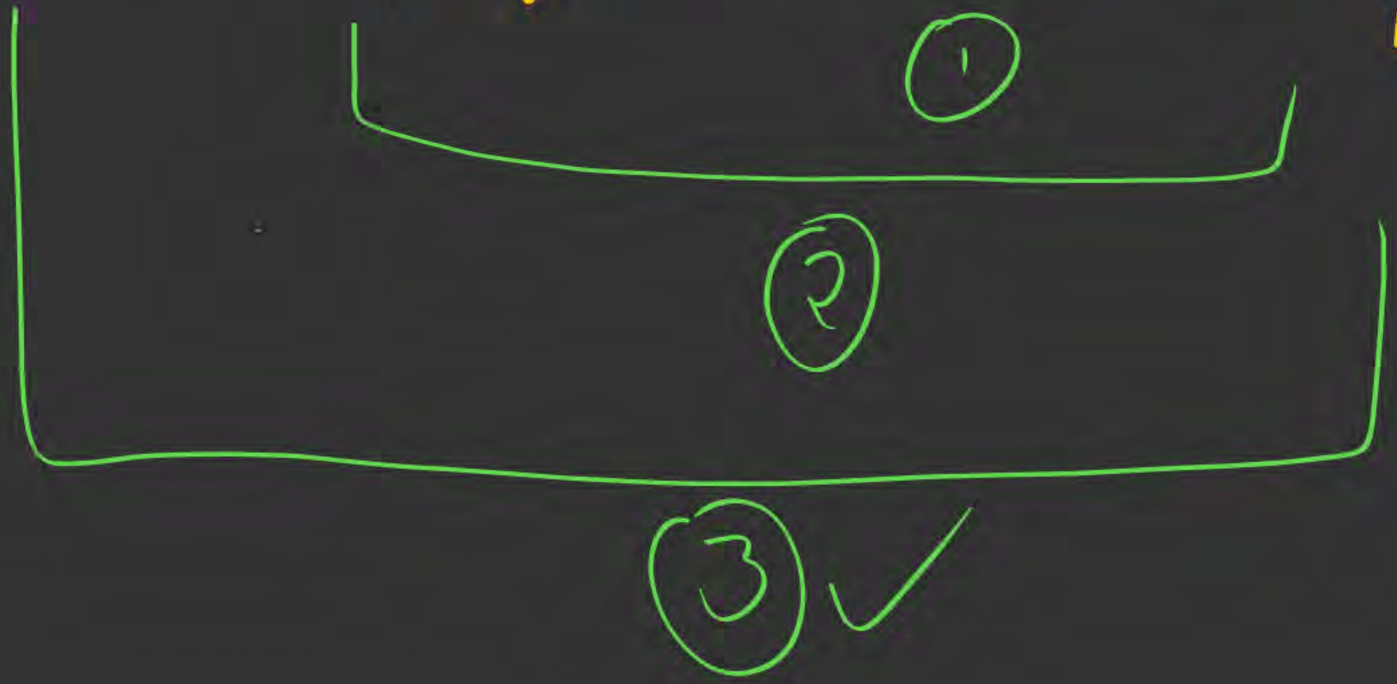
$= \boxed{225}$

3) Minimum Cost Spanning Tree (MCST)

(weight)

mwST

MST



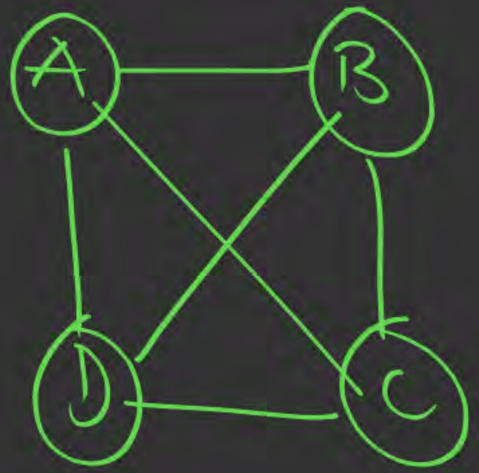
E ^ C E

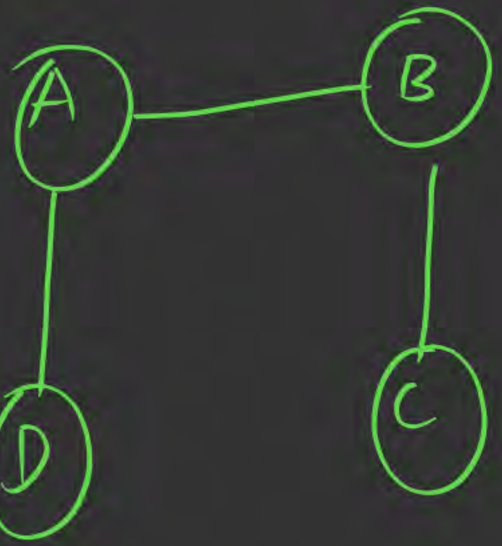
(V, E')

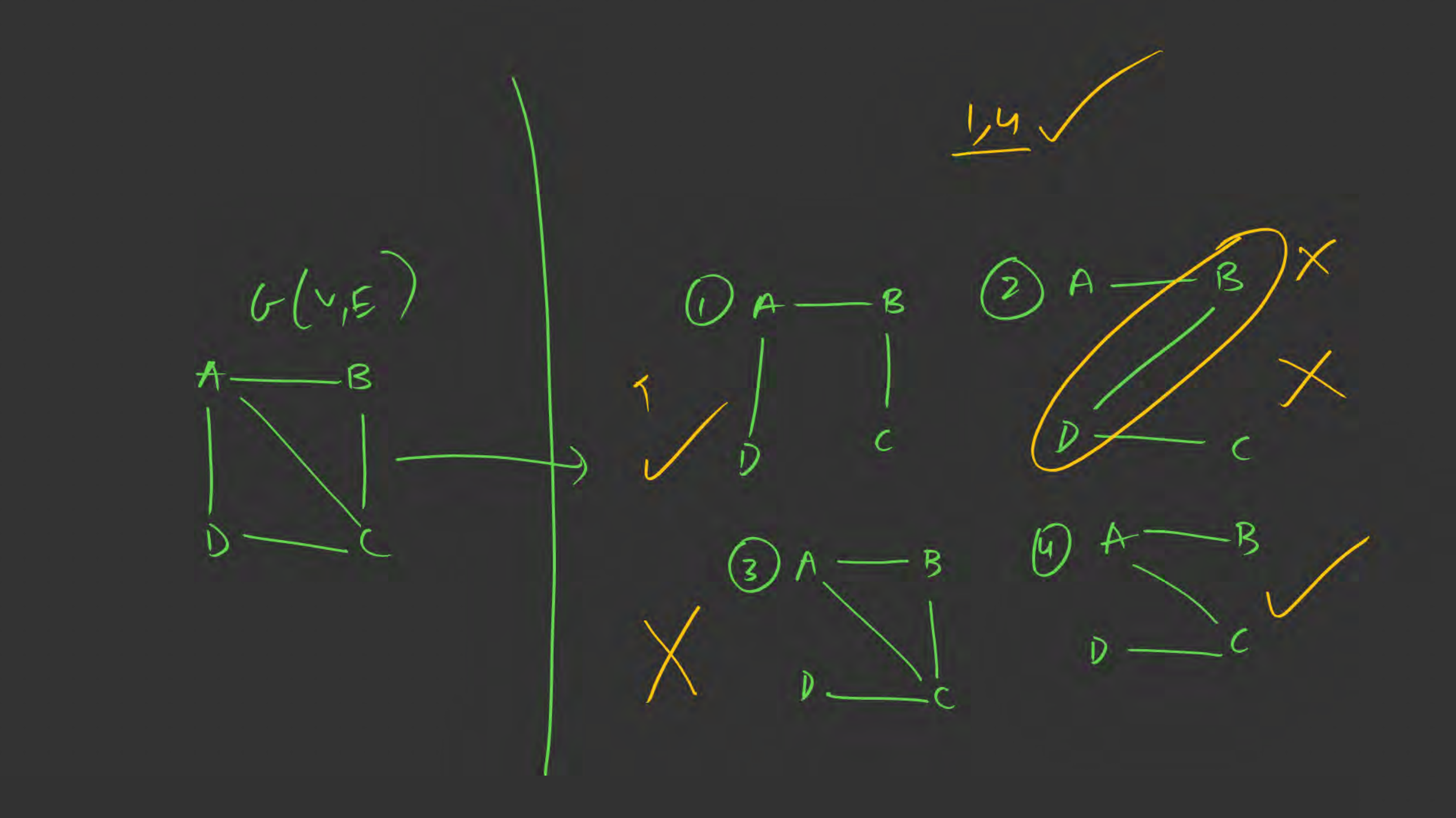
Graph : G(V, E)

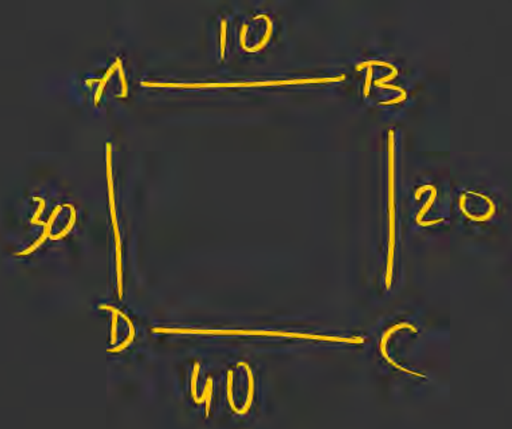
Spanning Tree (T)

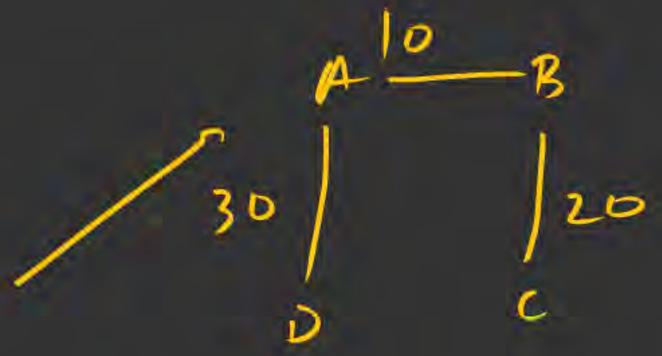
↳ Acyclic



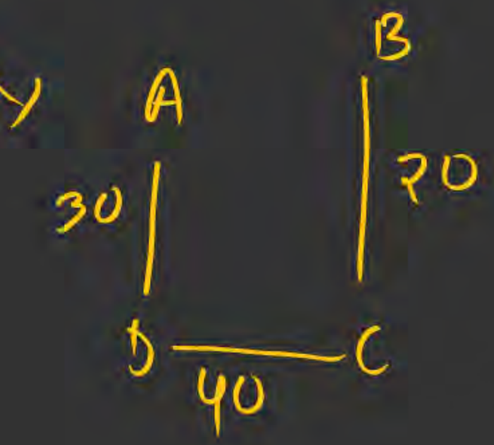



1,4 ✓
G(V,E)
A
B
D
C
1
A
B
T
D
C
2
A
B
X
X
D
C
3
A
B
X
D
C
4
A
B
D
C
✓

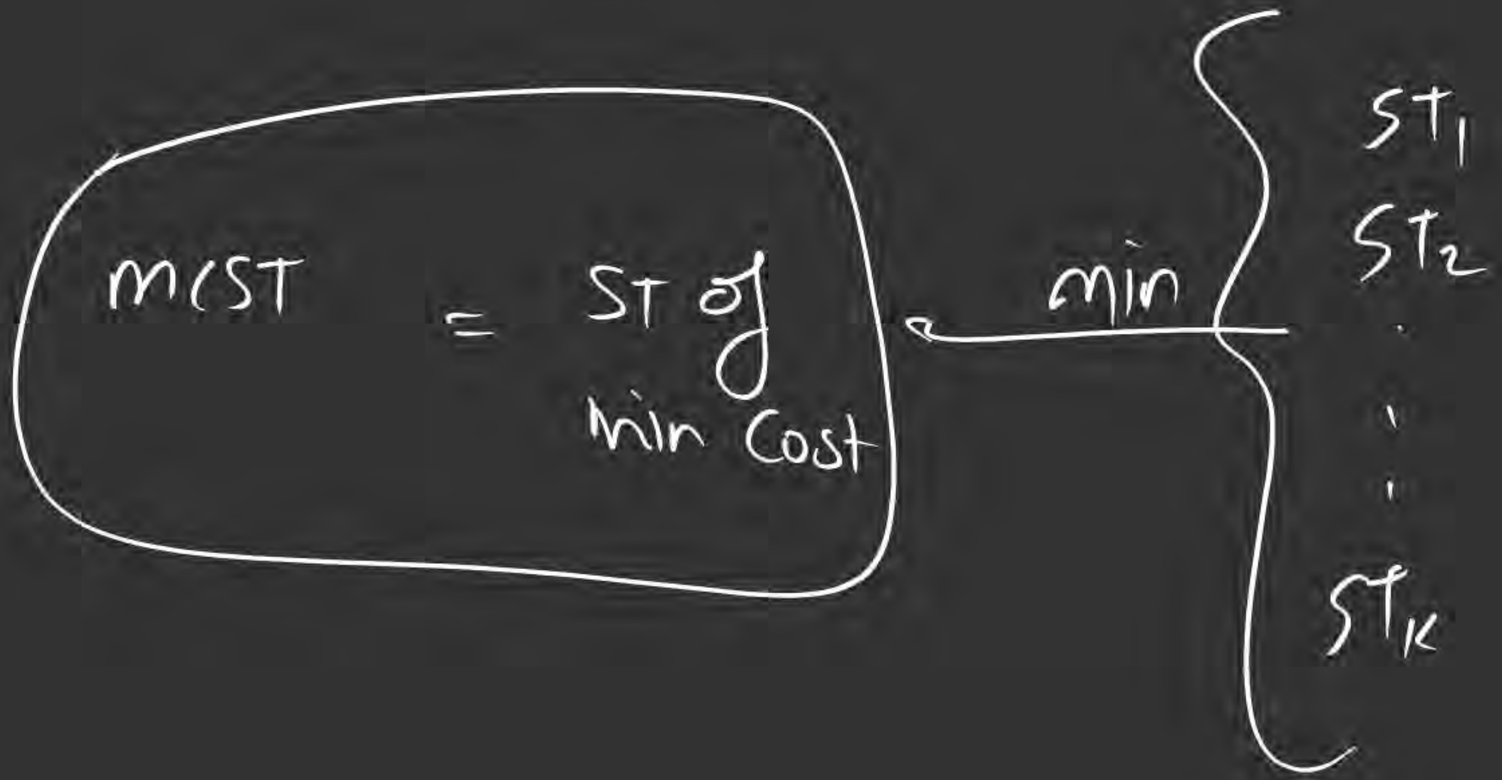




Cost = 10 + 30 + 20 = 60



Cost = 30 + 20 + 40 = 90

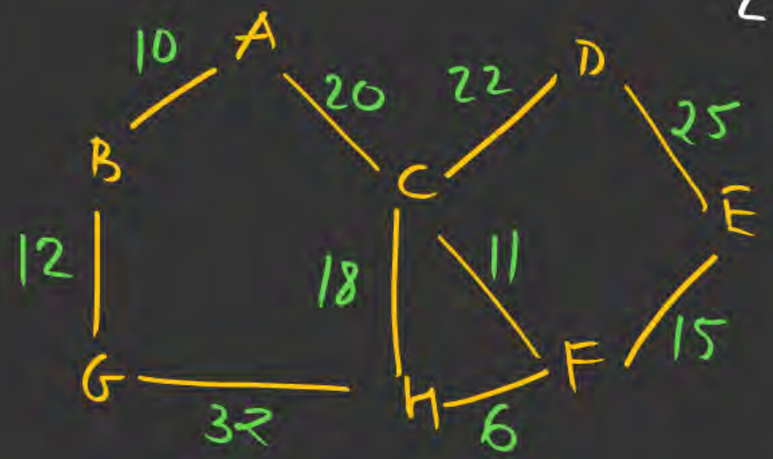

MCST
=
ST of
Min Cost
min
ST_1
ST_2
ST_K

Construction of MCST (optimization)

- 1) Prim's
- 2) Kruskal

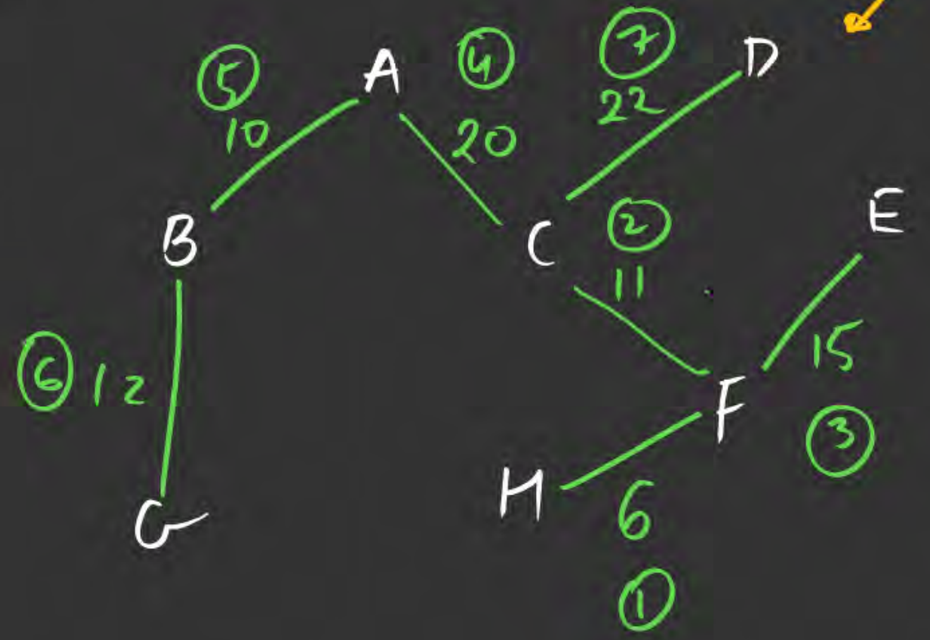
} Greedy

Graph: G(V, E)



Prims

MCST

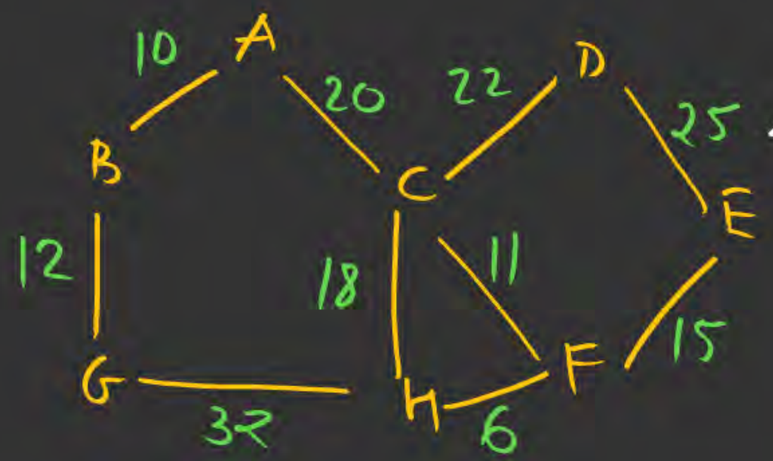


Cost of MCST = 6 + 11 + 15 + 20 + 10 + 12 + 22

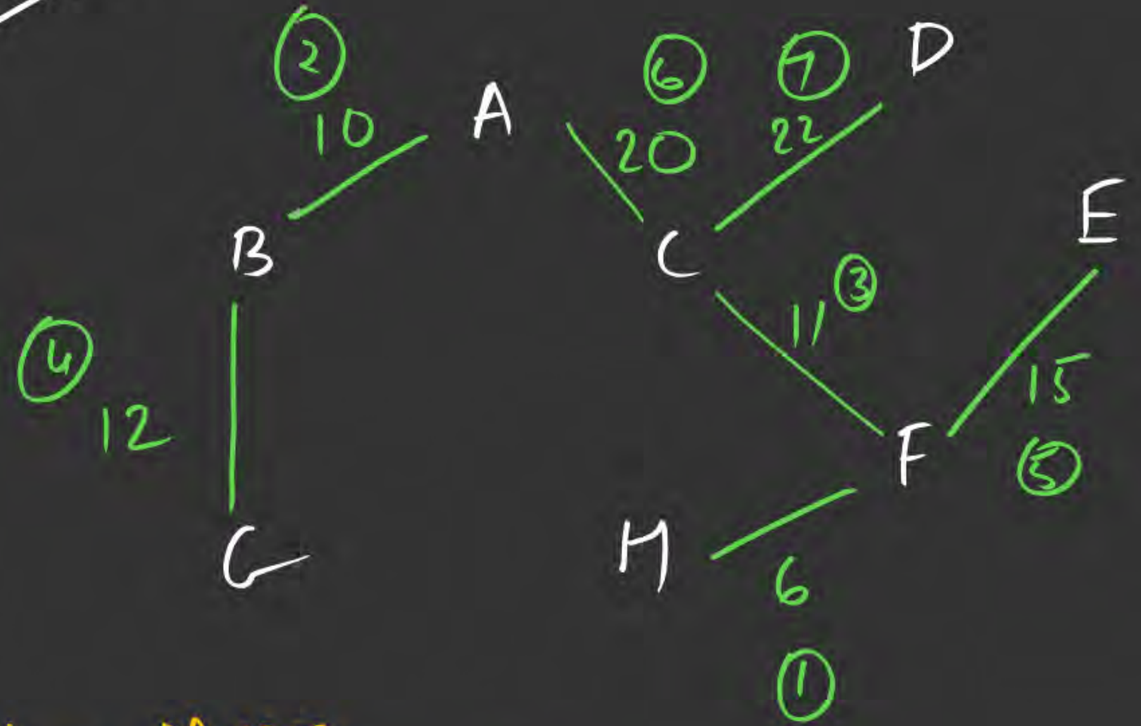
= 32 + 30 + 34

= 96

Graph: G(V, E)

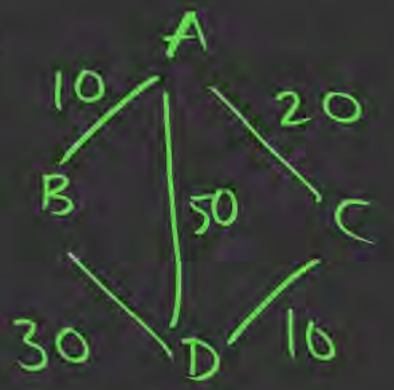


Kruskal MCST

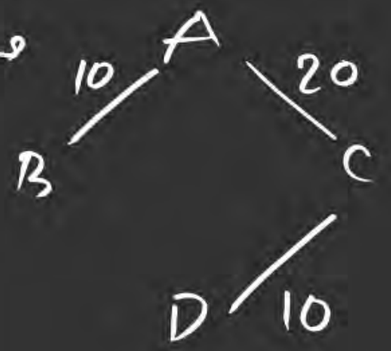


Cost of MCST = 96

eg2 :



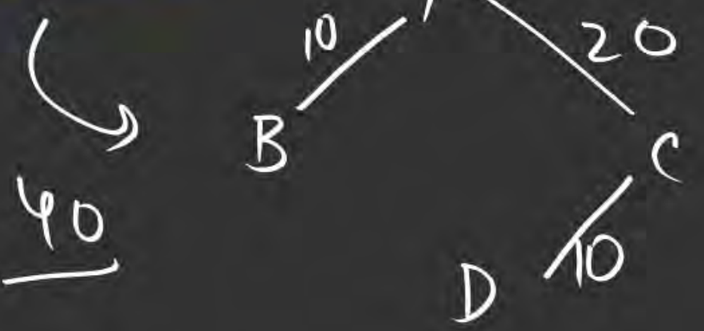
Prims



MCST

$\Rightarrow$ 40

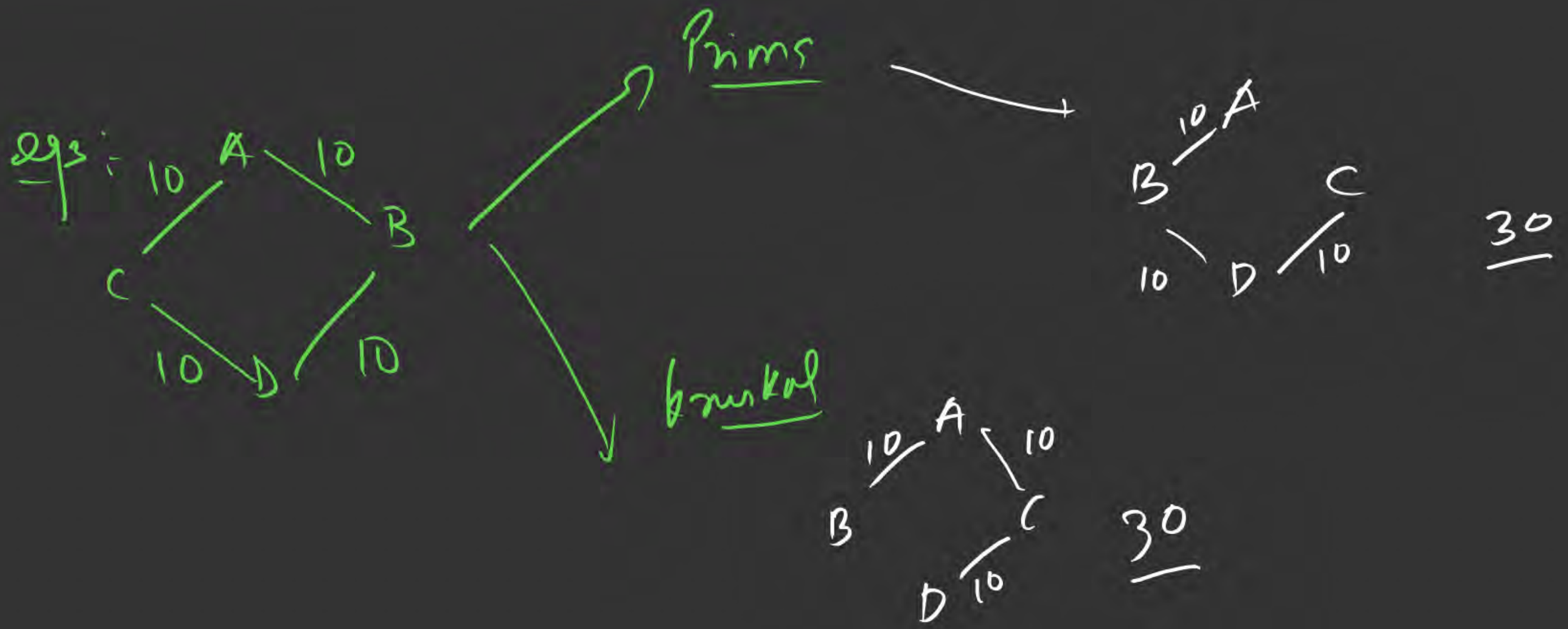
Kruskal



40

G(V, E), |V| = n

MCST edges = (n-1)


Eg:
A
10
10
B
C
10
D
10
Prims
B
10
A
C
10
10
D
30
Kruskal
10
A
10
B
C
D
10
30

Question

n=4

#Q. Let G be a complete undirected graph with 4 vertices and edge weights are {1, 2, 3, 4, 5, 6}. The maximum possible weight that a minimum weight spanning tree can have is ____.

6X

ans = max → { MCST$_1$, MCST$_2$, ⋮, MCST$_k$ }

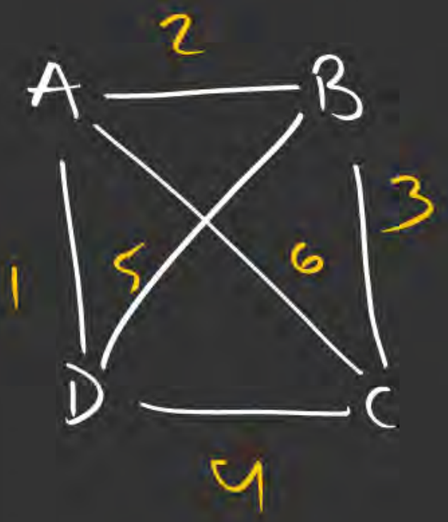
Ans: 7 ✓

Complete graph

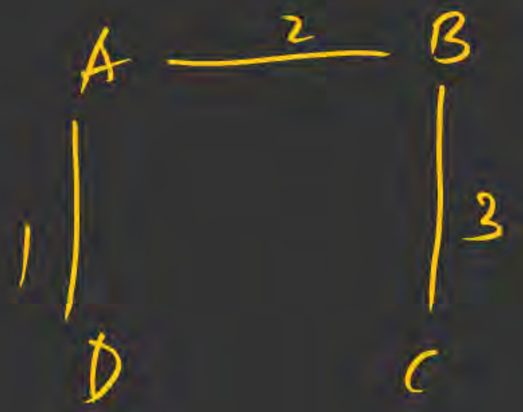
$n = 4$

$e = \frac{n(n-1)}{2}$

$= 6$

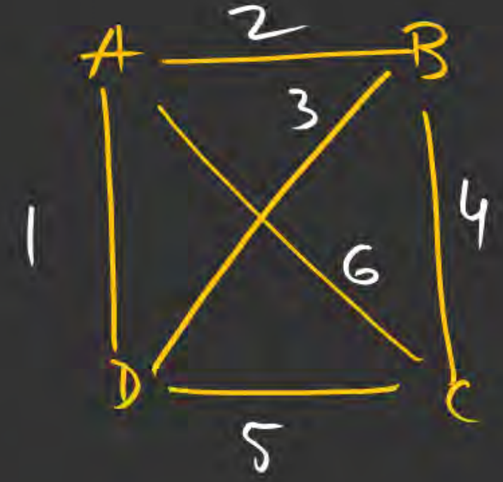


MCST

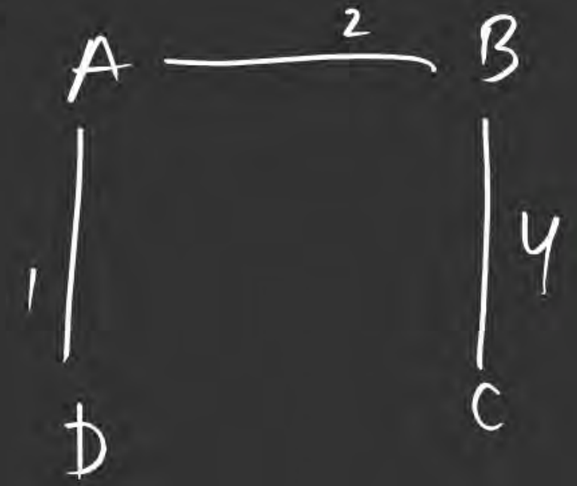


cost = 6

Case 2

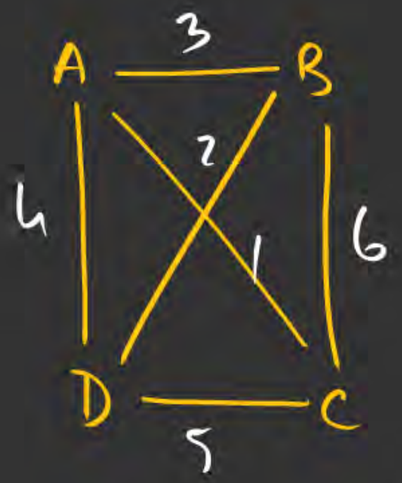


mcst

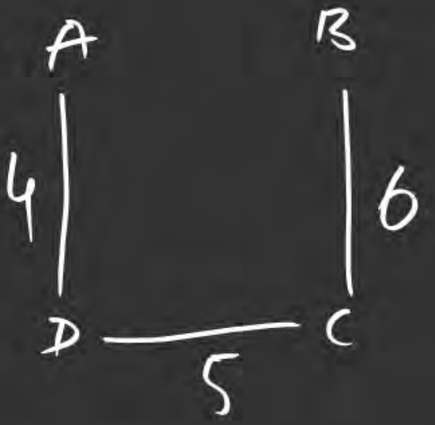


cost = 7

Case 3

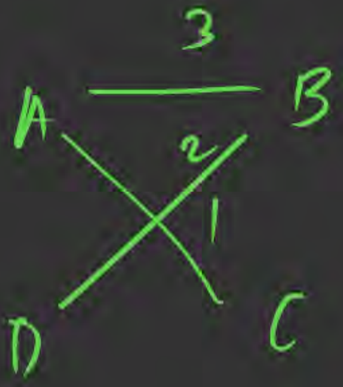


mcST ✗



Cost = 15

mcST



Cost = 6

Question

#Q. Let G a connected undirected graph of 100 vertices and 300 edges. The weight of a minimum spanning tree of G is 500. When the weight of each edge of G is increased by five, the weight of a minimum spanning tree becomes ____.

Soln: G → |V| = n = 100

|E| = 300

mst = 500

+5

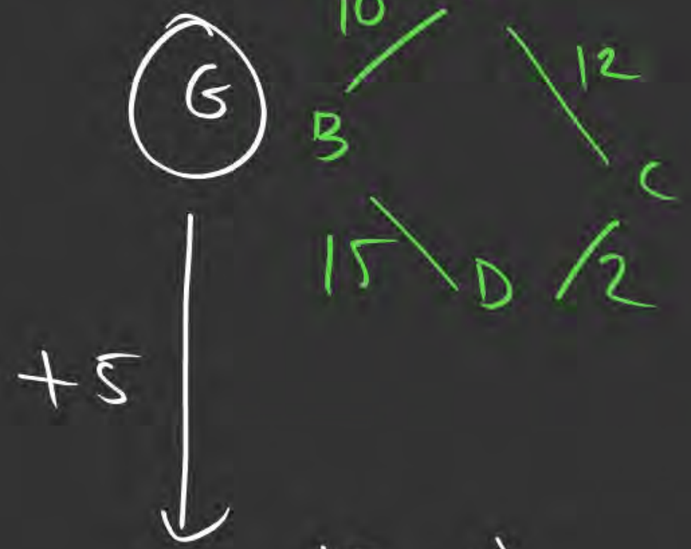
New Cost = 500 + (n − 1) × 5

= 500 + (100 − 1) × 5

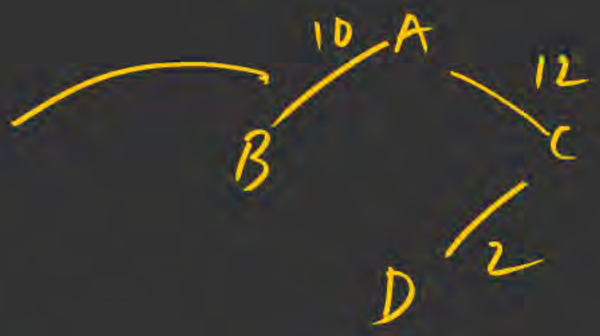
= 1000 − 5

= 995

Eg:-

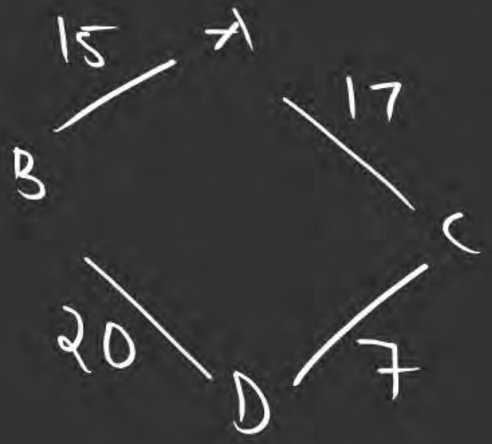


MCST

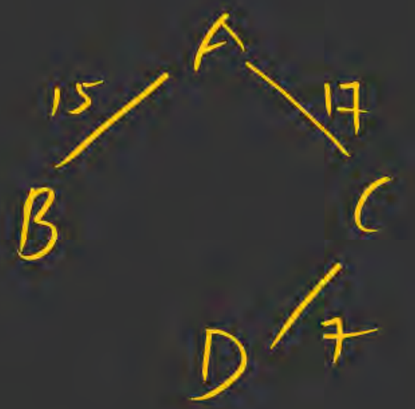


cost = 10+12+2

= 24



MCST

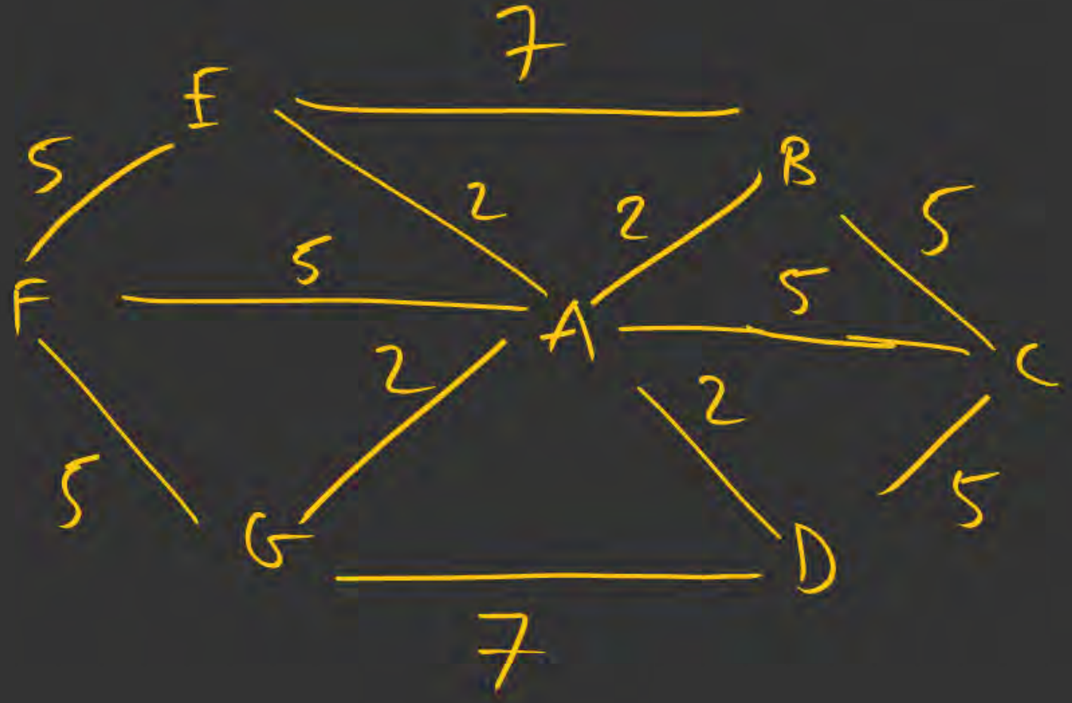


cost = 15+17+7

= 24+(3)×5

= 39

H.W



→ How many MCST? are possible

Thank

THANK

Keep Hustling!