Exercise :

Using ID3, compute the decision tree with Boolean formulas, (AND, OR, XOR, ...).

$n = 2$

⇒ for AND ,

$2^n = 2^2 = 4$.

A	B	A . B
0	0	0
0	1	0
1	0	0
1	1	1

Sample,
$S : [1+, 3-]$ (1T, 3F)

Entropy $E(S) = -1/4 \log_2 1/4 - 3/4 \log_2 3/4 = 0.811$

Selection of attribute that max the Information gain (I)

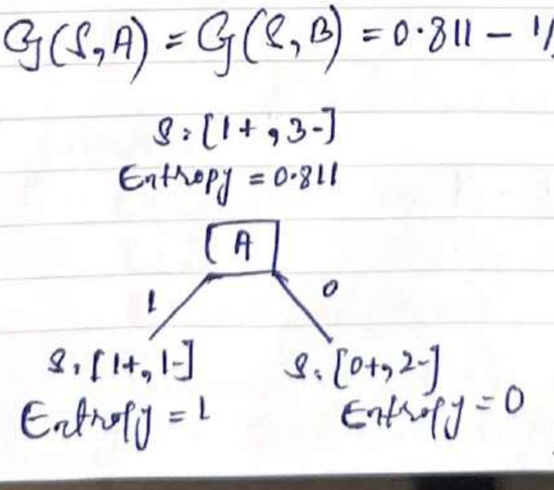
$S_{A=0} : [0+, 2-]$ $E(S_{A=0}) = -2/2 \log_2 2/2 = 0$

$S_{A=1} : [1+, 1-]$ $E(S_{A=1}) = -1/2 \log_2 1/2 - 1/2 \log_2 1/2 = 1$

$S_{B=0} = S_{A=0}$ $E(S_{B=0}) = E(S_{A=0})$.

$S_{B=1} = S_{A=1}$ $E(S_{B=1}) = E(S_{A=1})$

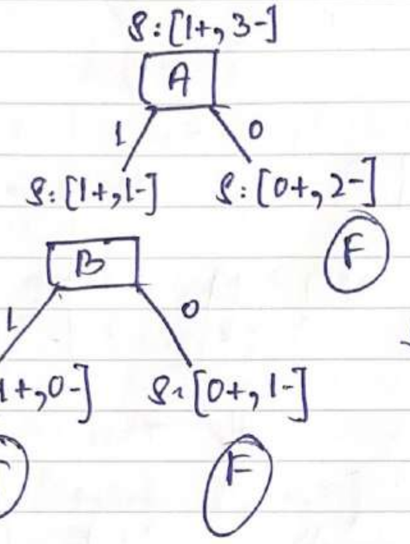
$I = G(S, A) = G(S, B) = 0.811 - 1/2 \cdot 1 = 0.311$.



$\boxed{G(S, A) = 0.311}$

As the info gain is same for both (A & B), we can select any attribute.

We choose A.



Only B remains

Same class e ⇒ leaf with 'e' as a label.

⇒ For OR gate:

A	B	A+B
0	0	0
0	1	1
1	0	1
1	1	1

Sample S: [3+, 1-]

Entropy $E(S) = -3/4 \log_2 3/4 - 1/4 \log_2 1/4$

$= 0.811$

Attribute Selection:

$S_{A=0}: [1+, 1-]$ $\quad E(S_{A=0}) = -1/2 \log_2 1/2 - 1/2 \log_2 1/2 = 1$

$S_{A=1}: [2+, 0-]$ $\quad E(S_{A=1}) = -2/2 \log_2 2/2 = 0$

$S_{B=0} = S_{A=0}$ $\quad E(S_{B=0}) = E(S_{A=0})$

$S_{B=1} = S_{A=1}$ $\quad E(S_{B=1}) = E(S_{A=1})$

Information gain: $G(S, A) = G(S, B) = 0.811 - 1/2 \, 1 = 0.311$

S: [3+, 1-]

Entropy = 0.811

$\boxed{A}$

1 | 0

S: [2+, 0-]   S: [1+, 1-]

E = 0   E = 1

G (S, A) = 0.311

Information gain is more so we select A.

S: [3+, 1-]

$\boxed{A}$

1 | 0

S: [2+, 0-]   S: [1+, 1-]

T

$\boxed{B}$ Only B remaining

1 | 0

S: [1+, 0-]   S: [0+, 1-]

T   F

⇒ for XOR:

A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0

$S: [2+, 2-]$, $E(S) = -2/4 \log_2 2/4 - 2/4 \log_2 2/4 = 1$

Selection of attributes.

$S_{A=0} : [1+, 1-]$ $\quad E(S_{A=0}) = -1/2 \log_2 1/2 - 1/2 \log_2 1/2 = 1$
$S_{A=1} : [1+, 1-]$ $\quad E(S_{A=1}) = -1/2 \log_2 1/2 - 1/2 \log_2 1/2 = 1.$

$S_{B=0} = S_{A=0}$ $\quad E(S_{B=0}) = E(S_{A=0})$
$S_{B=1} = S_{A=1}$ $\quad E(S_{B=1}) = E(S_{A=1}).$

Info gain $\quad G(S, A) = G(S, B) = 1 - 1/2\, 1 - 1/2\, 1 = 0.$

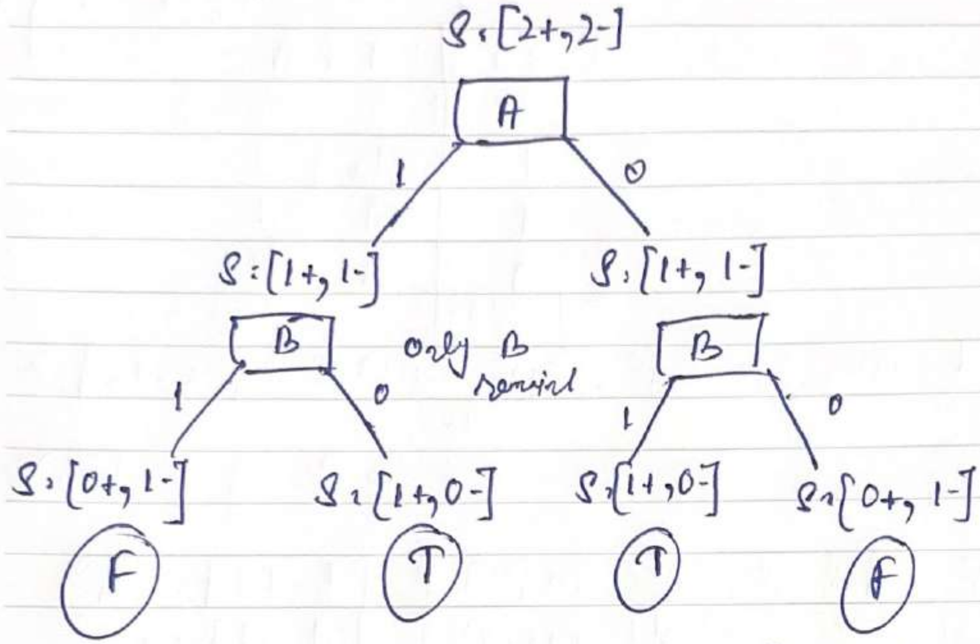
$S: [2+, 2-]$
$E = 1$

$\boxed{A}$

1 / \ 0

$S: [1+, 1-]$ $\quad S: [1+, 1-]$
$E = 1$ $\quad E = 1$
$G(S, A) = 0$

Same info gain, we select A then.



Same class c that leaf labelled 'c'.

□