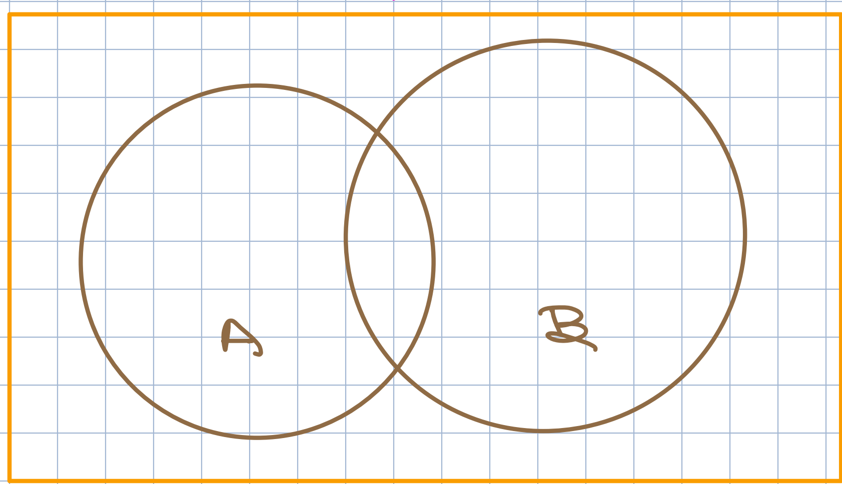
Agenda

- Conditional Probability
- Multiplication Rule
- Marginal and Joint probability
- Tree Diagram Approach
- Law of Total probability
- Baye's Theorem

Recap



$$A \cup B = A + B - A \cap B$$

$$\Downarrow$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Question What will be $P(A \cap B)$ if A and B are disjoint Events?

Answer:

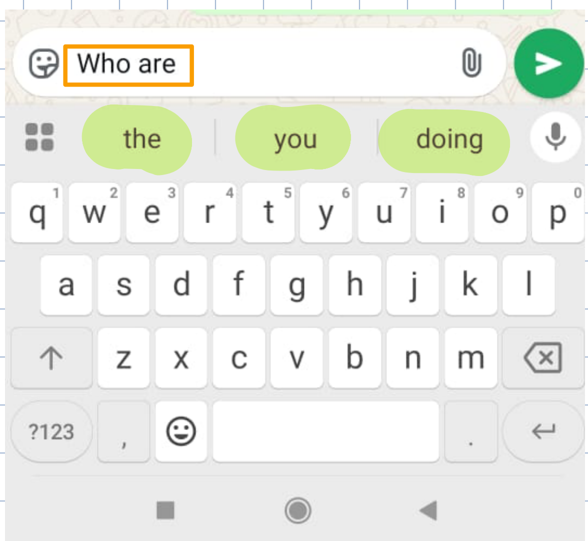
$$A \cap B = \{\ \}$$

$$\Downarrow$$

$$P(A \cap B) = 0$$

$$P(A \cup B) = P(A) + P(B)$$

Conditional Probability



Who are → the
→ you
→ doing

Q: ~ 100,000 words in English (Assume)

The keyboard app shortlisted 3 words out of 100 K words

$P(x_i \mid x_1 = \text{How})$

1

are >> 0
is >> 0
... ...
... ...
cond

100k

# of occurrence

| are | is | |

$P(x_i \mid (x_1 = \underbrace{\text{How}}_{T_1} \text{ and } x_2 = \underbrace{\text{are}}_{T_2}))$

1 → 100k

...

$$\boxed{P(A \mid B)}$$

⇒ Probability of A given that B has occurred

* A ⇒ Prediction
* B ⇒ Words you have typed
  Tokens

$$P(A \mid B) \neq P(A)$$

$$P(\text{are} \mid \text{How}) \neq P(\text{are})$$

Experiment

Experiment : Sum of 2 Dices

Dice2 output

Dice1 output	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$P(D_1=2) = \frac{6}{36}$

$P(D_1+D_2 \le 5) = \frac{10}{36}$

$P\left[\begin{array}{l} D_1 = 2 \text{ and} \\ D_1 + D_2 \le 5 \end{array}\right]$

$P(D_1=2 \cap D_1+D_2 \le 5) = 3/36$

Sample Space $\Rightarrow D_1 + D_2$

What is $P(D_1=2)$ given that $D_1+D_2 \le 5$

$P(D_1 = 2 \mid D_1 + D_2 \le 5) = 3/10$

Event that has already occured

2	3	4	5
$\boxed{3 \quad 4 \quad 5}$			
4	5		
5			

$D_1 = 2$

$\leftarrow D_1 + D_2 \le 5$

=

$P(D_1 = 4 \mid D_1 + D_2 \leq 5) = \frac{1}{10}$

$P(D_1 = 4) = 6/36$

P(A) given that B has occured

$P(A \mid B)$

$P(D_1 = 2 \mid D_1 + D_2 \leq 5) = \frac{3}{10}$

$$\frac{P(D_1 = 2 \cap D_1 + D_2 \leq 5) = \frac{3}{\cancel{36}}}{P(D_1 + D_2 \leq 5) = 10/\cancel{36}} = \frac{3}{10}$$

$$P(D_1 = 2 \mid D_1 + D_2 \leq 5) = \frac{P(D_1 = 2 \cap D_1 + D_2 \leq 5)}{P(D_1 + D_2 \leq 5)}$$

$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}}$$

Multiplication Rule

conditional

$$P[A/B] \times P[B] = \frac{P[A \cap B]}{\cancel{P[B]}} \times \cancel{P}[B]$$

$$P[A \cap B] = P[A/B] \times P[B]$$ (Multiplication Rule)

$$\updownarrow$$

$$P[A \cap B] = P[B/A] \times P[A]$$

important

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

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

Questions Is $P(A|B)$ same as $P(B|A)$?

Marginal vs Joint vs Conditional probability

Marginal

$$P(W) = \frac{184}{360}$$

$$P(C) = \frac{46}{360}$$

Joint : Two events occuring together

$$P(W \cap C) \Rightarrow \frac{30}{360}$$

$$P(W^c \cap C^c)$$

$$P(W \cap C^c)$$

Conditional

$$P(C \mid W) = 30/184$$

$$P(W \mid C) \quad 30/46$$

Tree Diagram Approach

Question: Email Spam System

- Let's say 30% of all Emails are Spam
- 70% are Non-Spam
- 80% of all spam Emails contain word 'purchase'
- 10% of Non-spam contains purchase
- Overall what % of Email will have word 'purchase'

0.31

$\boxed{0.8 * 0.3 + 0.1 * 0.7}$

Answer:

P(Spam) = 30%

P(Non-spam) = 70%

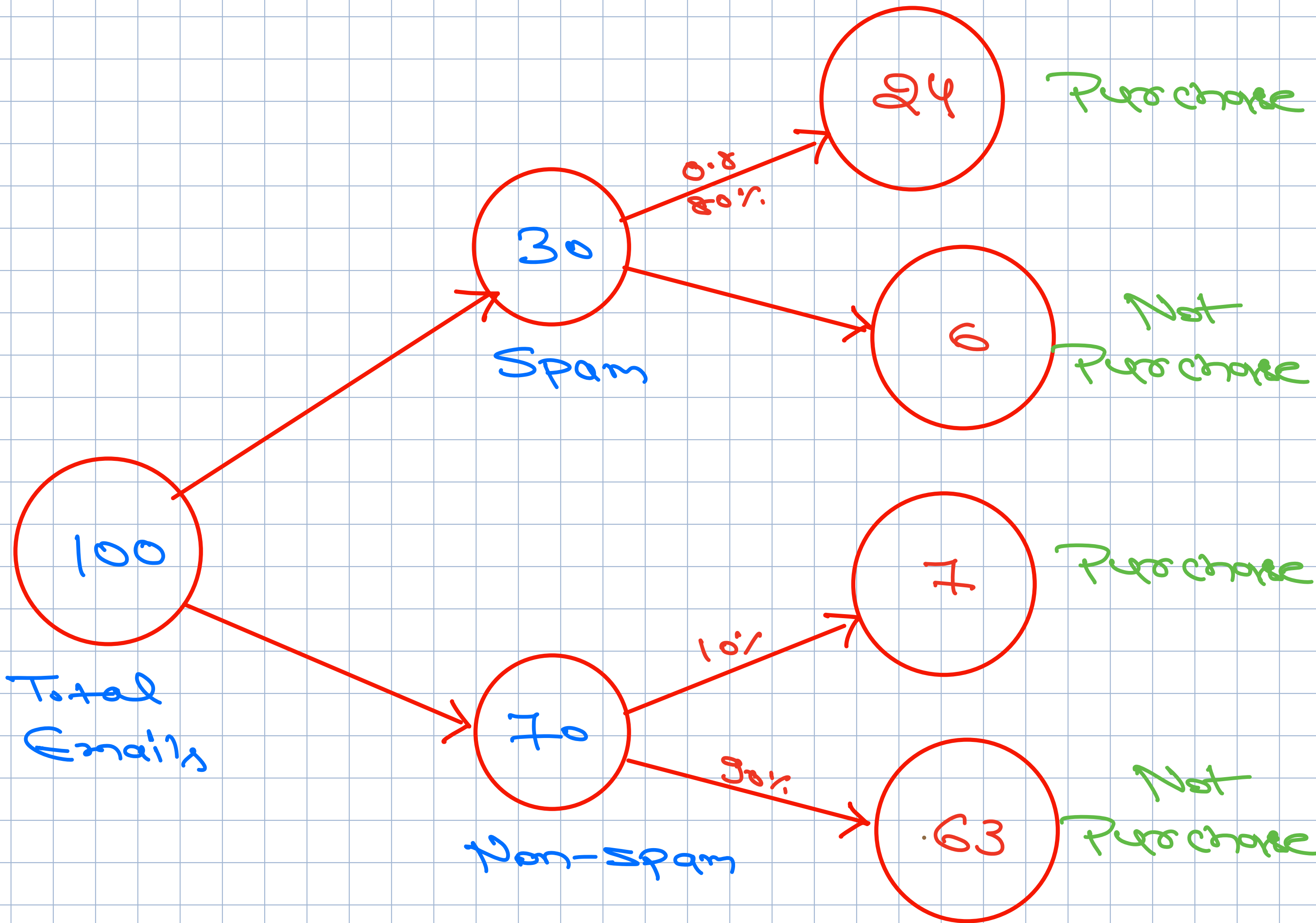
P(Purchase | spam) = 80%

P(Purchase | Non spam) = 10%

P(Purchase) = ?

P(Pur) = Pspam * P(purch | spam) + Ppur(Non-spam AN

= P(Purchase | Spam) = P(Spam | Purchase)



Purchase = 24 + 7

$$P(\text{Purchase}) = \frac{31}{100} = 0.31$$

$$\boxed{P(\text{Purchase}) = \begin{array}{l} P(\text{Purchase} \cap \text{Spam}) = \frac{24}{100} \\ \quad + \\ P(\text{Purchase} \cap \text{Non Spam}) \end{array} \begin{array}{c} \\ + \frac{7}{100} \end{array} = \frac{31}{100}}$$

Law of total Probability

$\underline{\underline{Q}}$: $P(\text{Spam} \mid \text{purchase}) = \frac{24}{31}$

$P(\text{Purchase}) = P(\text{Purchase} \cap \text{Spam}) +$

$P(\text{Purchase} \cap \text{Non Spam})$

$P(A) = P(A \cap B_1) + P(A \cap B_2)$

$$\boxed{P(A \cap B) = P(A/B) \times P(B)}$$

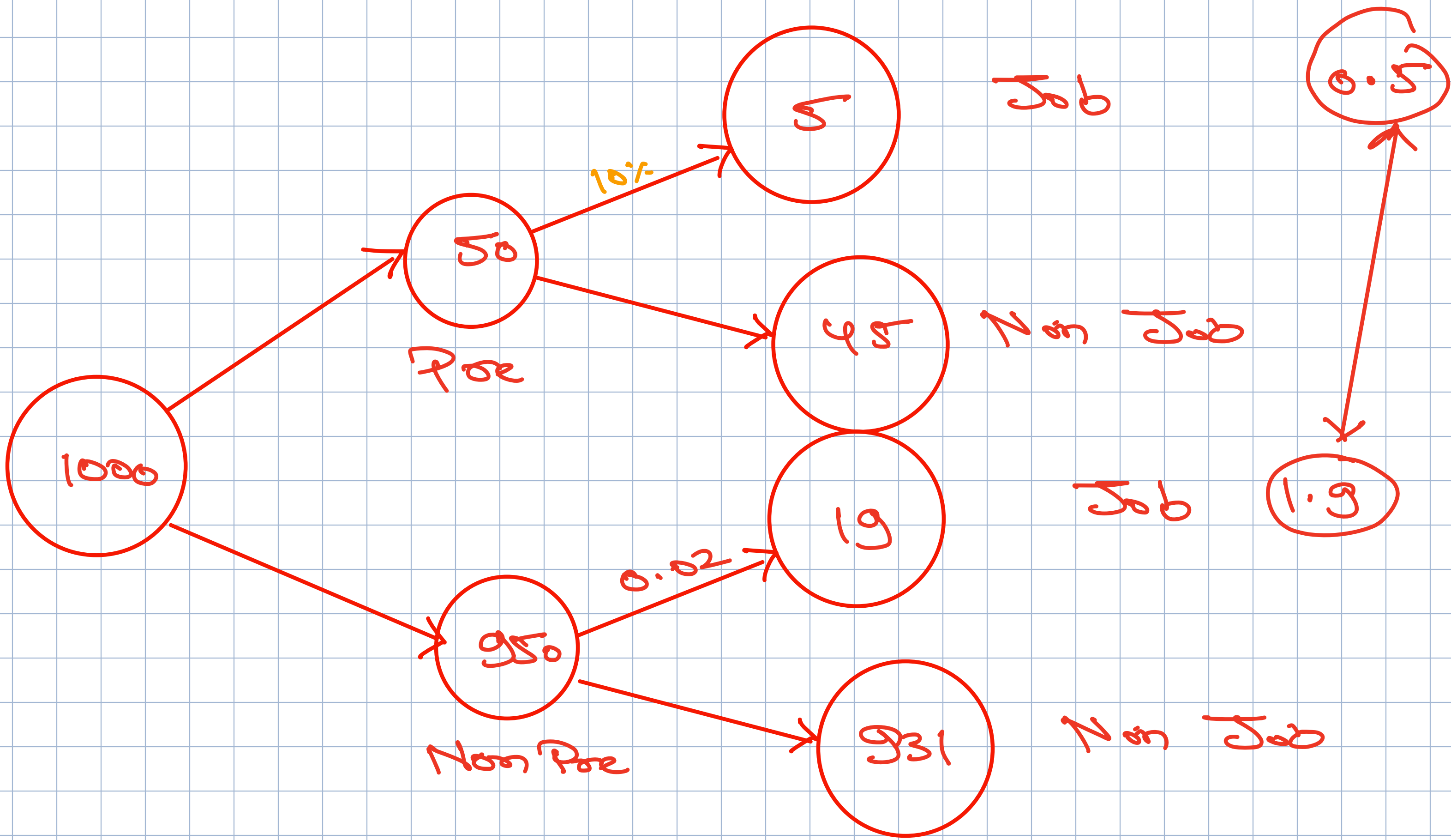
$P(A) = P(A \mid B_1) \times P B_1 + P(A \mid B_2 + P(B_2)$

$$\boxed{\text{Law of Total probability}}$$

$$\boxed{P(A) = \sum_{i=1}^{n} P(A \mid B_i) * P(B_i)}$$

Questions

It is known that -
5% of all LinkedIn users are premium users
10% of premium users are actively seeking new job opportunities.
Only 2% of non-premium users are actively seeking new job opportunities.
Overall, what percentage of people are actively seeking new job opportunities



$19 + 5 = 24$

$P(\text{job}) = \frac{24}{1000}$

Questions

: Solve above Questions with using Law of Total probability

Summary

1. Conditional Probability:

- $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

2. Multiplication Rule:

- $P(A \cap B) = P(A \mid B) \cdot P(B)$

3) Law of Total Probability:

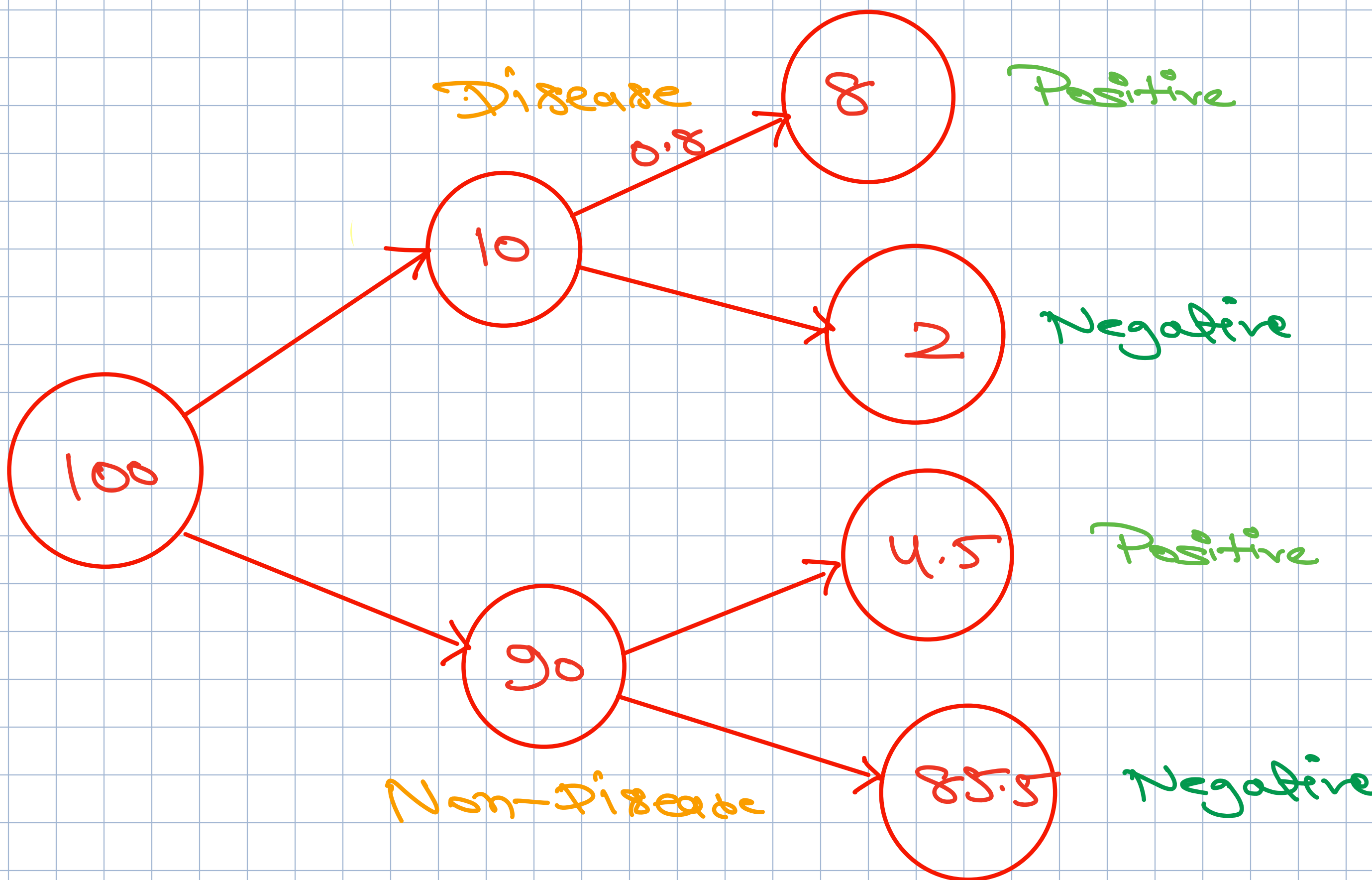
- $P(A) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i)$

Questions

```
A disease affects 10% of the population.
Among those who have the disease, 80% get “positive” test result
Among those who don’t have the disease, 5% get “positive” test result.
What is P(+ve | Disease) ?

a) 0.1
b) 0.8
c) 0.05
d) 0.85
```

10



$$P(+ve \mid Disease) = \frac{8}{10} = 0.8$$

$$\downarrow$$

$$P(+ve \cap Disease) / Disease$$

A disease affects 10% of the population. Among thos who have the disease, 80% get “positive” test result Among those who don’t have the disease, 5% get “positive” test result. What is P(+ve ∩ Disease)?

4 options

$P_{+ve} \cap$ disease

$P[A \cap B] \Rightarrow P_{A/B} \times P_B$

Baye's Theorem

important

Likelihood

Marginal

$$P[A|B] \Rightarrow \frac{P[B|A] \times P[A]}{P[B]}$$

Posterior Probability

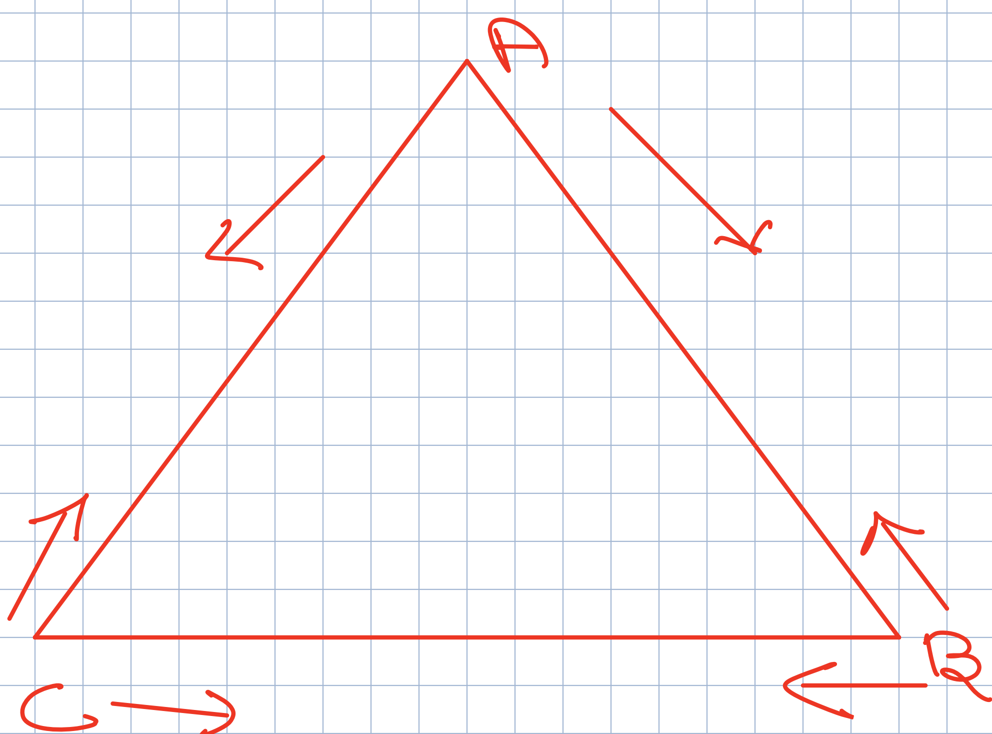
Marginal

Questions : Derive Above formula

Hint :

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

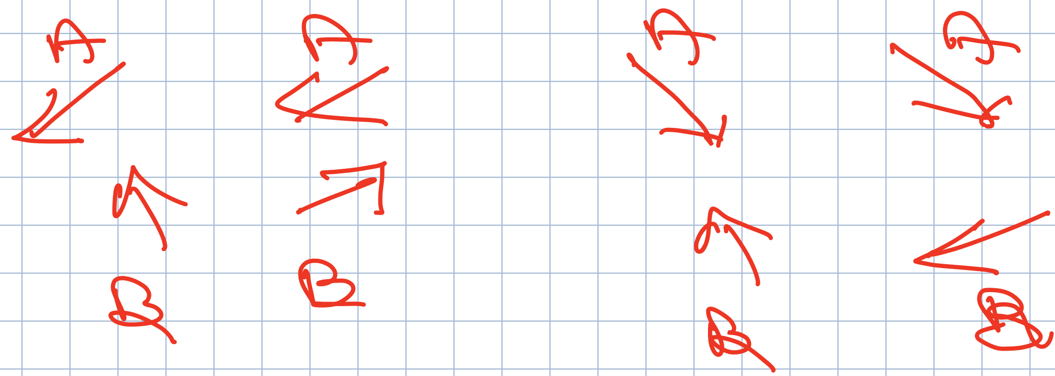
$$P(B \cap A) = P(B/A) \times P(A)$$

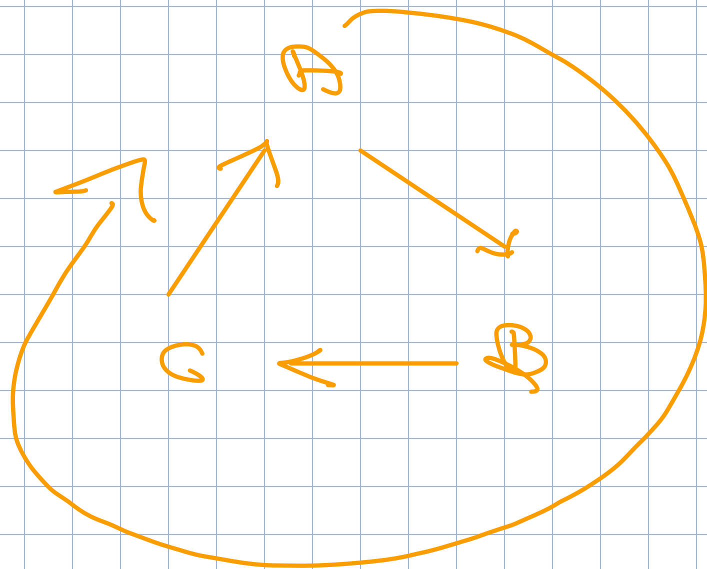


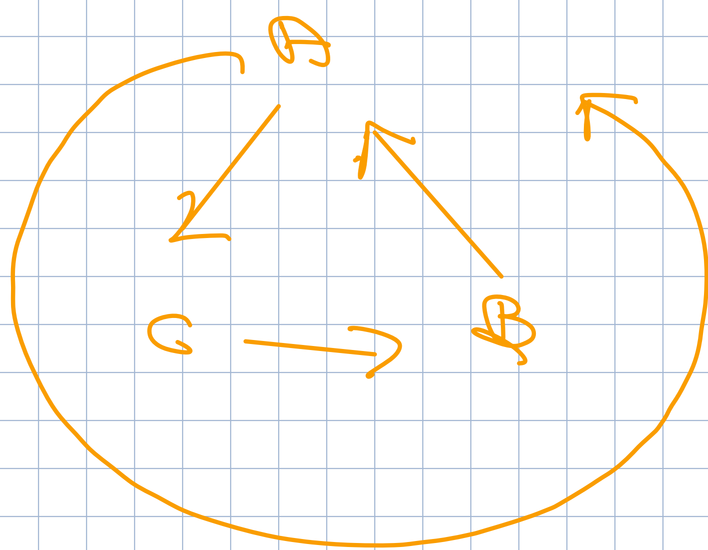
A

Total Possible $\Rightarrow 2 \times 2 \times 2$

$\Rightarrow 8$ ways







$P((0|1)) = \frac{8-2}{8} = \frac{6}{8}$