

Lecture 17

Bayes Rule &

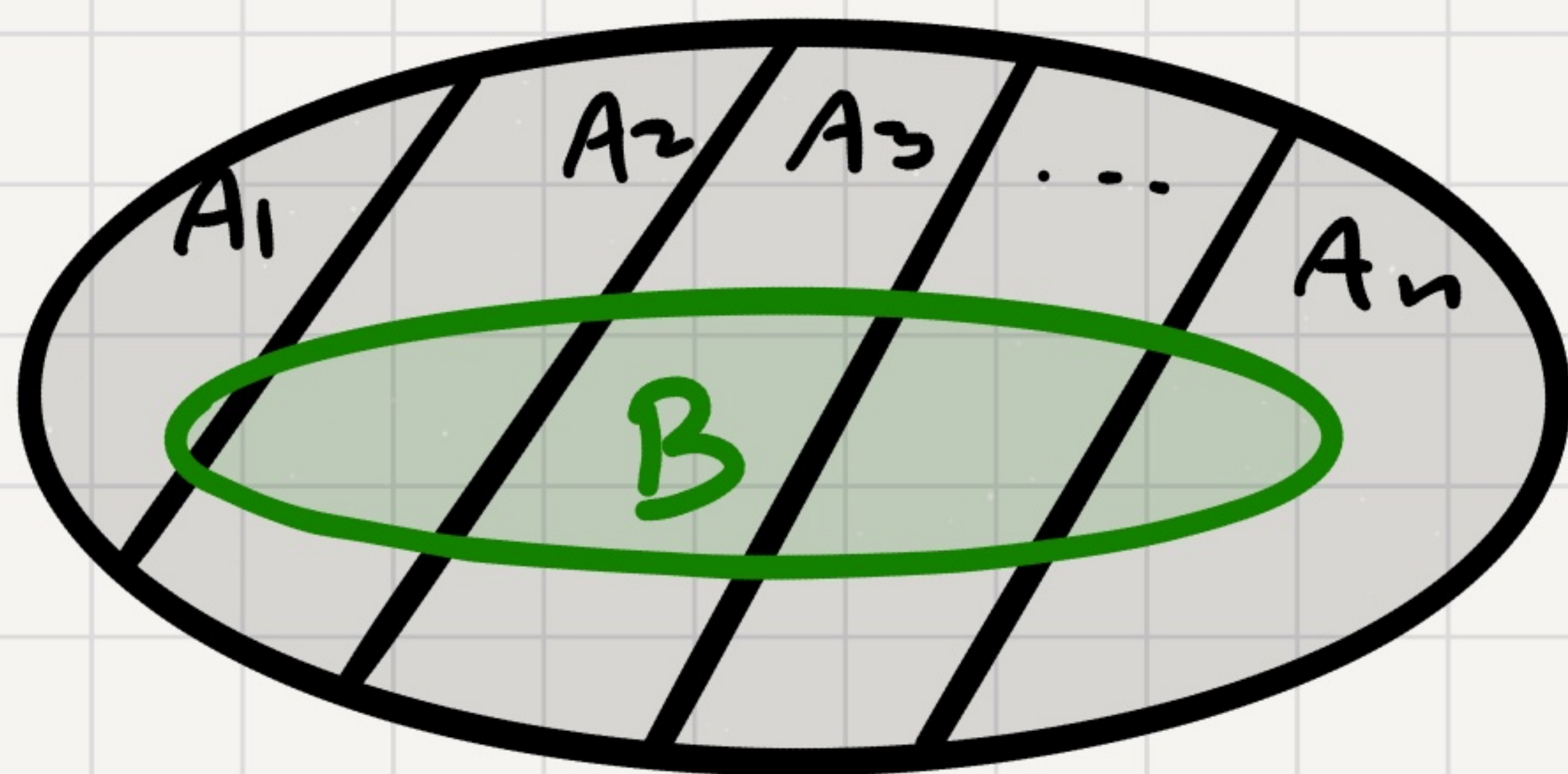
Applications in CS.

Lecture 16 Summary

- Probability is additive
- Union Bound $\Pr[A_1 \cup \dots \cup A_n] \leq \Pr[A_1] + \dots + \Pr[A_n]$
- Inclusion-Exclusion $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$
- Total probability: if A_1, \dots, A_n partition Ω
then $\Pr[B] = \Pr[A_1 \cap B] + \dots + \Pr[A_n \cap B]$
- Conditional probability: $\Pr[B|A] = \frac{\Pr[A \cap B]}{\Pr[A]}$
- Independence: $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$
- Correlation $\Pr[A \cap B] > \Pr[A] \cdot \Pr[B]$
- Bayes Rule.

Total Probability with Conditional Probability

Assume Ω is a union of disjoint events A_1, \dots, A_n .



Since B is the disjoint union of $B \cap A_1, \dots, B \cap A_n$

$$Pr[B] = Pr[B \cap A_1] + \dots + Pr[B \cap A_n]$$

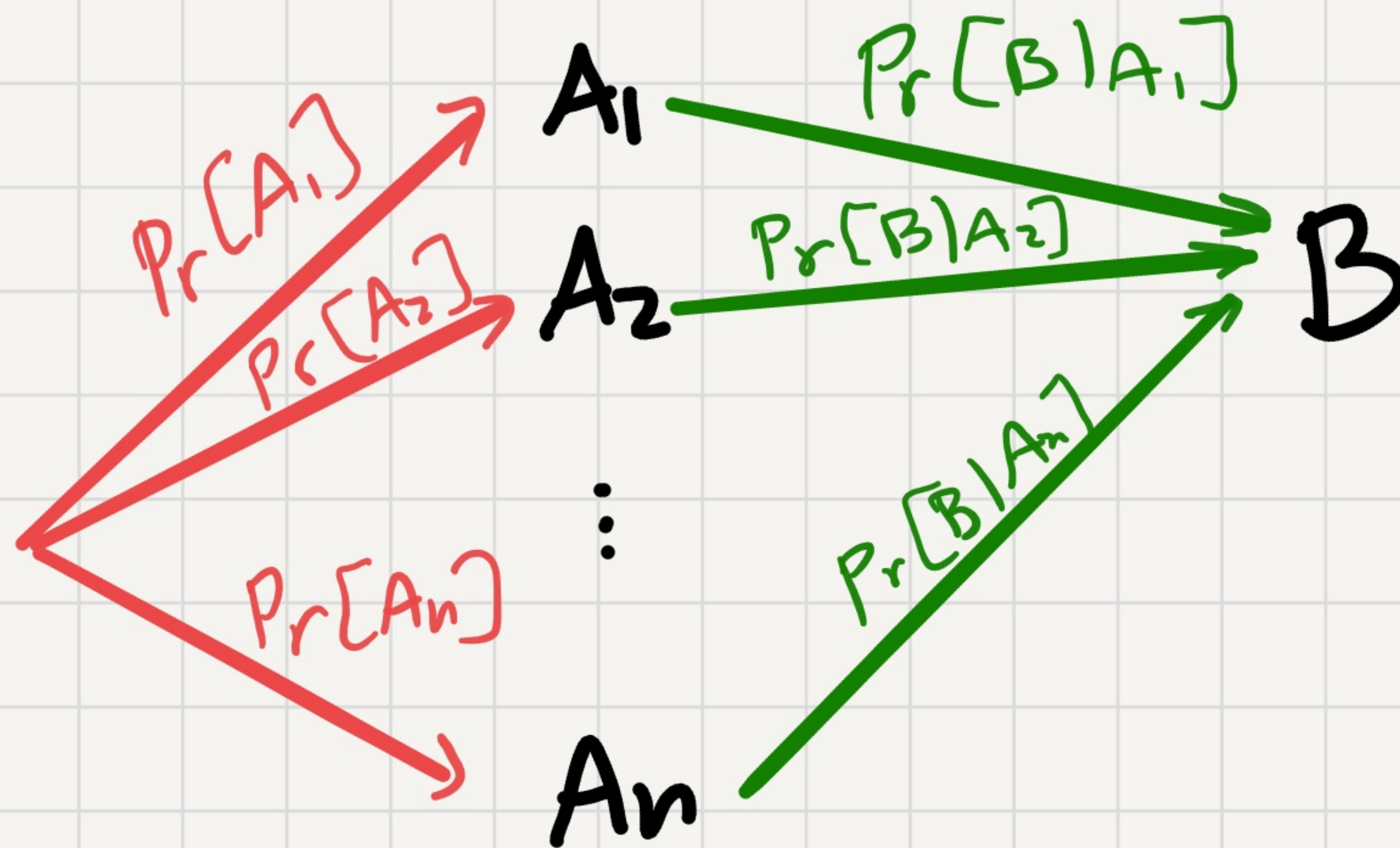
Thus, by the product rule

$$Pr[B] = Pr[A_1] \cdot Pr[B|A_1] + \dots + Pr[A_n] \cdot Pr[B|A_n]$$

Total Probability Rule with Conditional Probability

Prior Prob.

Conditional Prob.



$$Pr[B] = Pr[A_1] \cdot Pr[B|A_1] + \dots + Pr[A_n] \cdot Pr[B|A_n]$$

Bayes Rule

Suppose you know $\Pr[B|A]$, $\Pr[A]$, $\Pr[B]$

What's $\Pr[A|B]$?

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \frac{\Pr[A] \cdot \Pr[B|A]}{\Pr[B]}$$

Bayes Rule Example #1

Experiment:

1. Pick at random either a fair coin
or a biased coin with 60% chance heads.
2. Toss the coin you picked

$$\Omega = \{ (\text{fair}, H), (\text{fair}, T), (\text{biased}, H), (\text{biased}, T) \}$$

A = "coin is fair"

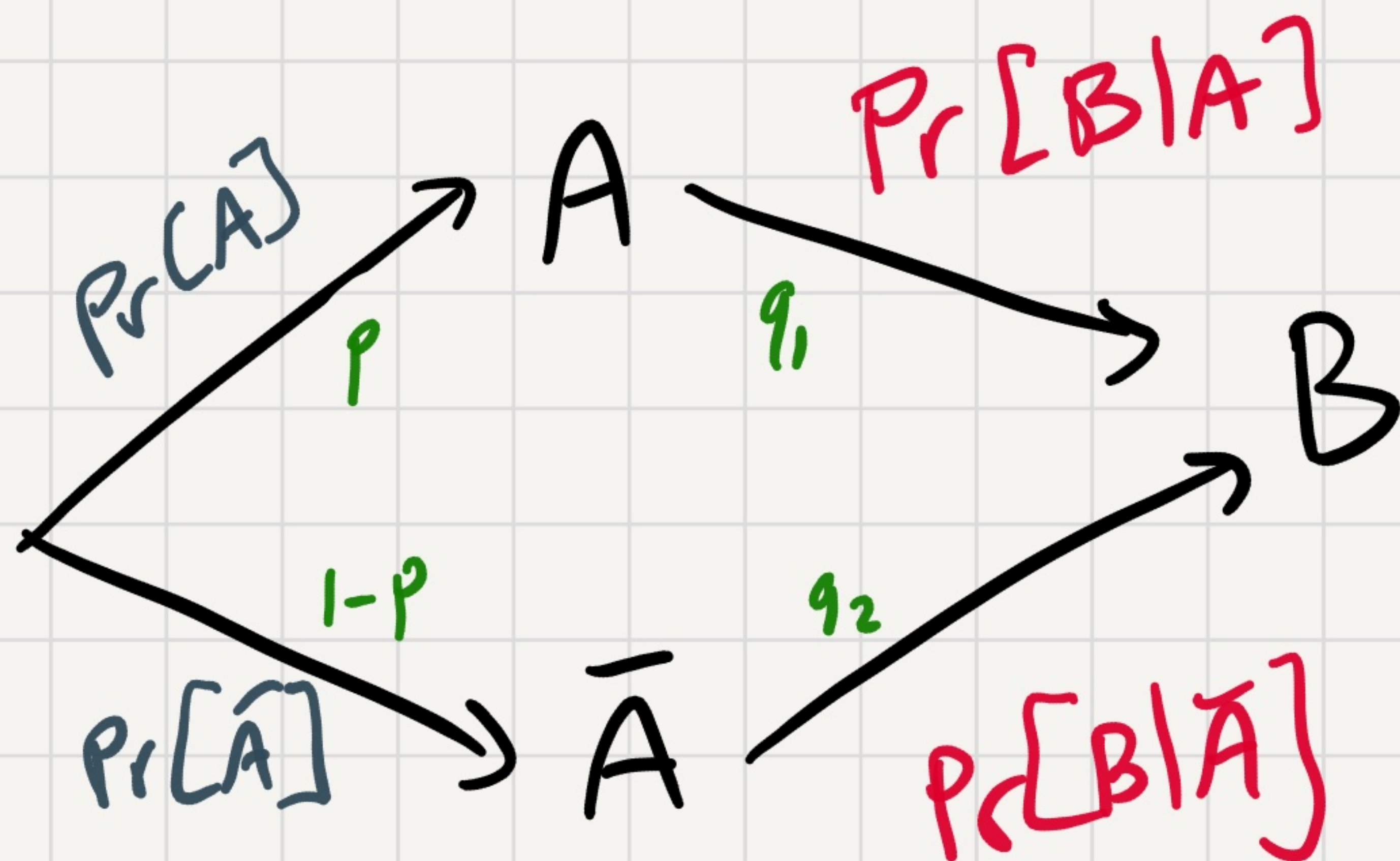
B = "got H"

$$\text{WTK: } \Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \frac{\Pr[A] \cdot \Pr[B|A]}{\Pr[B]} = \frac{0.25}{\Pr[B]}$$

$$\begin{aligned} \Pr[B] &= \Pr[A] \Pr[B|A] + \Pr[\bar{A}] \Pr[B|\bar{A}] \\ &= 0.5 \times 0.5 + 0.5 \times 0.6 = 0.55 \end{aligned}$$

Bayes

$$\Pr[A|B] = \frac{\Pr[A] \Pr[B|A]}{\Pr[B]}$$



$$\Pr[B] = \Pr[A] \cdot \Pr[B|A] + \Pr[\bar{A}] \cdot \Pr[B|\bar{A}]$$

Bayes Rule (updated)

$$\Pr[A|B] = \frac{\Pr[A] \cdot \Pr[B|A]}{\Pr[A] \cdot \Pr[B|A] + \Pr[\bar{A}] \cdot \Pr[B|\bar{A}]} = \frac{p q_1}{p q_1 + (1-p) q_2}$$

Bayes Rule Example #2

Suppose there's a disease that occur in 0.001 of the population.

There's a test for the disease.

For a random person: $\Pr[\text{test positive} \mid \text{sick}] = 0.99$

$\Pr[\text{test positive} \mid \text{not sick}] = 0.01$

A random person arrives and tests positive

Q: what's the likelihood that he has the disease?

$$\Pr[\text{sick}] = 0.001$$

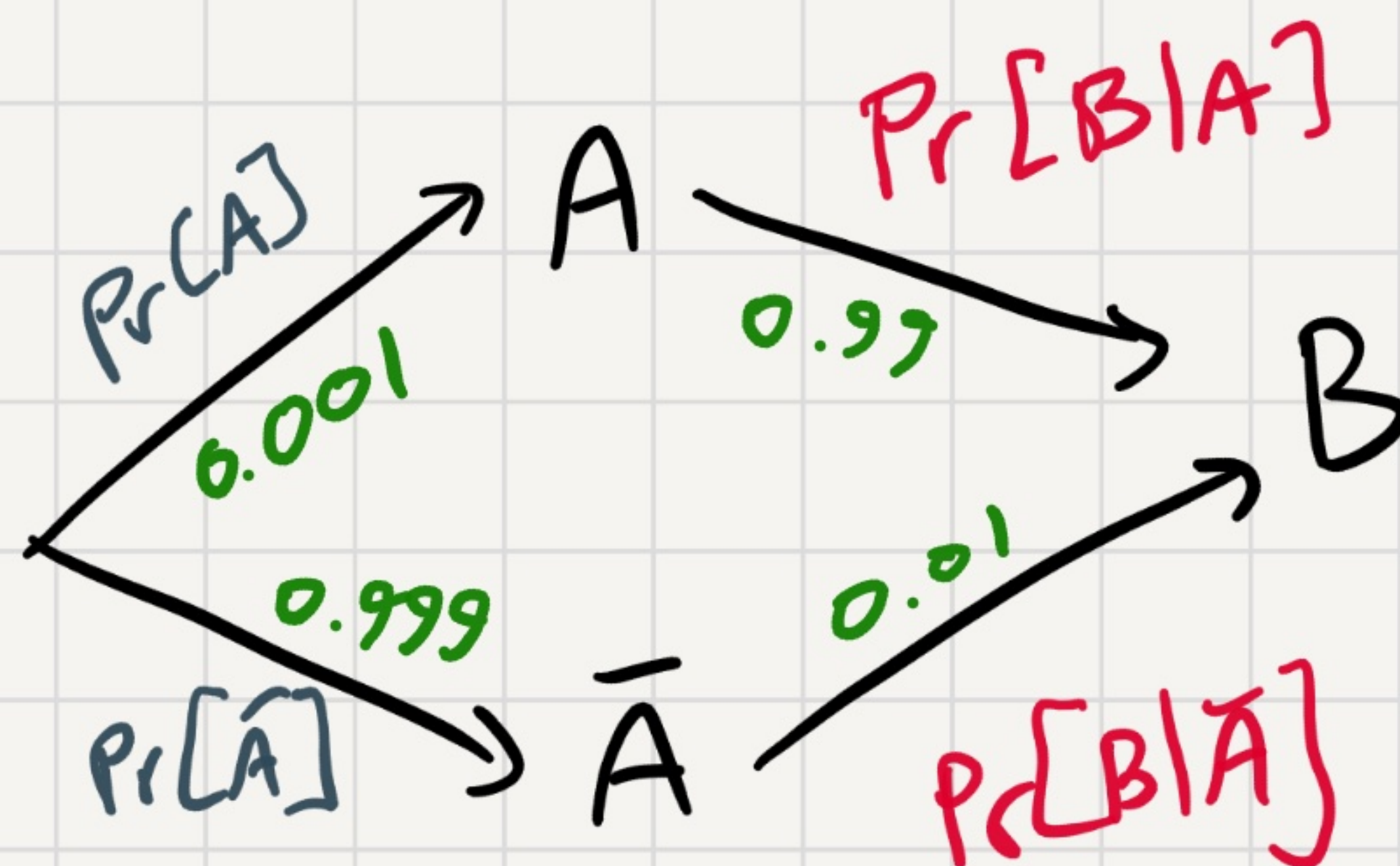
$$\Pr[\text{positive} | \text{sick}] = 0.99$$

$$\Pr[\text{positive} | \text{not sick}] = 0.01$$

A = "sick"

B = "tested positive"

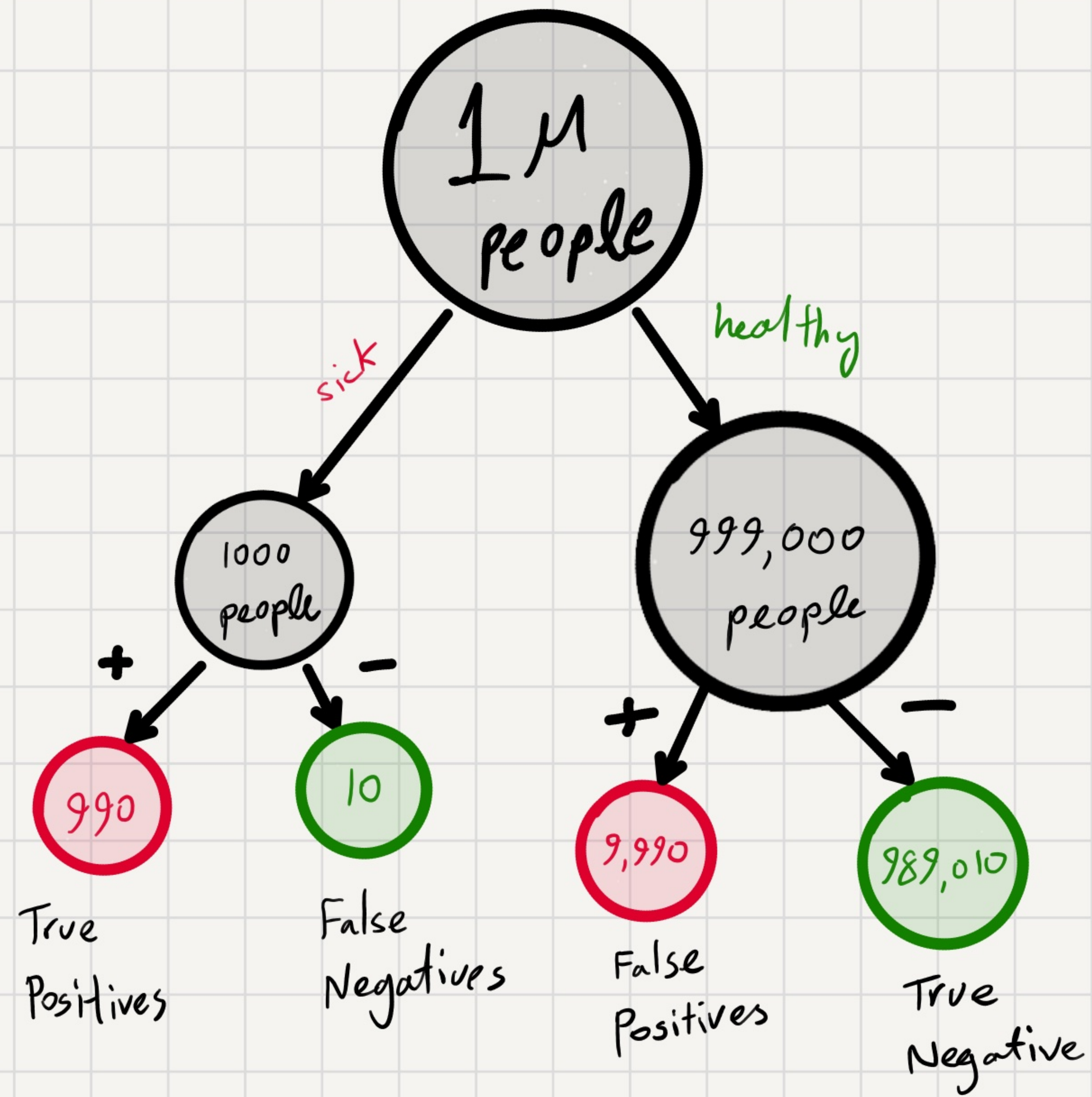
WTK: $\Pr[\text{sick} | \text{positive}]$



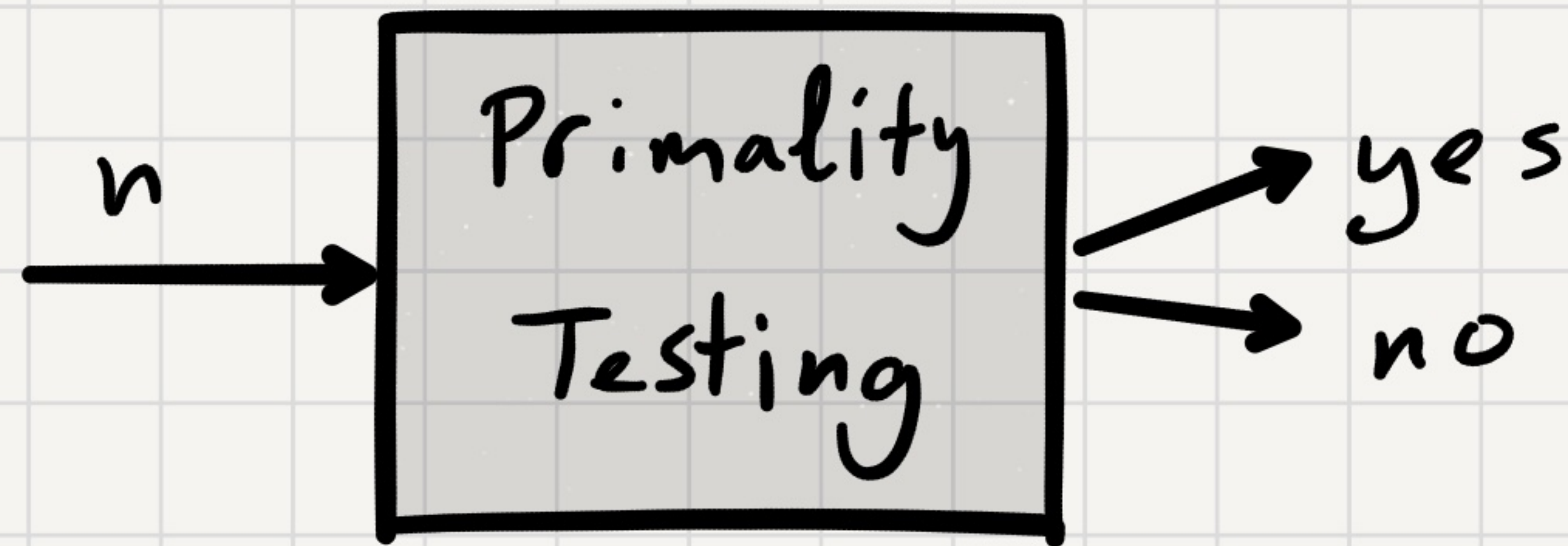
$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

$$= \frac{\Pr[A] \cdot \Pr[B|A]}{\Pr[A] \cdot \Pr[B|A] + \Pr[\bar{A}] \cdot \Pr[B|\bar{A}]} = \frac{0.001 \times 0.99}{0.001 \times 0.99 + 0.999 \times 0.01}$$

$$\approx 0.09$$



Primality Testing



Properties:

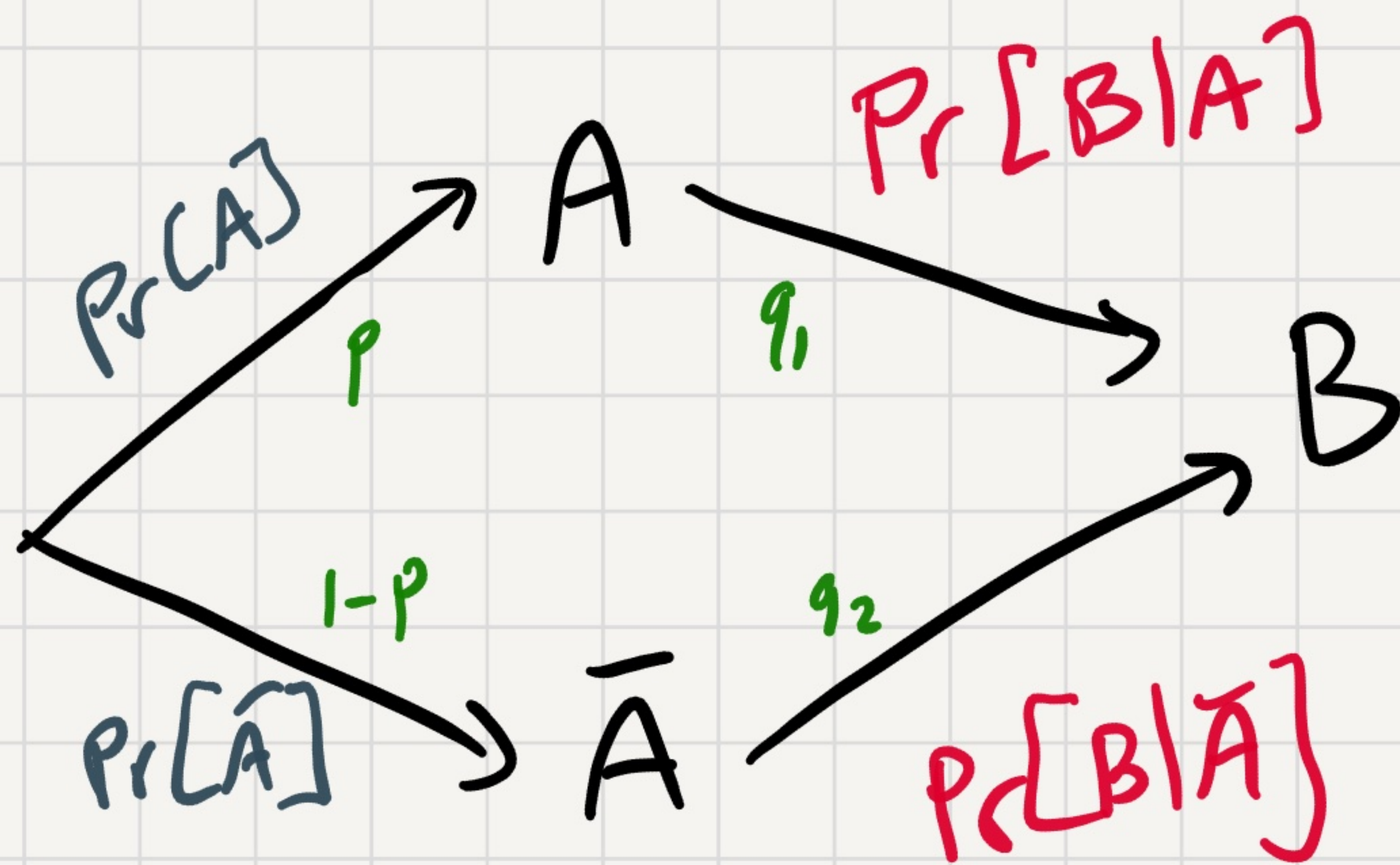
- If n is prime, always output "yes"
- If n is composite, outputs "no" w.p. $1 - \frac{1}{1000}$

what's $\Pr[n \text{ is prime} \mid \text{output is "yes"}]$?

For that, n should be random as well.

- Pick $n \in [2^{511}, 2^{512})$ uniformly at random.

Primality Testing



$$\Pr[A|B] = \frac{p q_1}{p q_1 + (1-p) q_2}$$

$A = "n \text{ is prime}"$

$B = "answer \text{ is yes}"$

$$p = \frac{1}{350}$$

$$q_1 = 1$$

$$q_2 = \frac{1}{1000}$$

$$\Pr[A|B] = \frac{p \cdot q_1}{p \cdot q_1 + (1-p) \cdot q_2} = \frac{\frac{1}{350}}{\frac{1}{350} + \frac{349}{350} \cdot \frac{1}{1000}} \approx 0.74$$

Birthday Paradox

Poll: How many students do you need to have in a classroom so that you'll definitely have a pair of students with the same birthday?

A: 23

B: 365

C: 366

D: 365×364

Birthday Paradox

How many students do you need to have in a classroom so that you'll have a pair of students with the same birthday w.p. $\geq \frac{1}{2}$.

Birthday Paradox

Assumption: Each student has a uniformly random birthday out of the 365 options.

k students.

$$\Omega = \{ (x_1, \dots, x_k) : \forall i \quad 1 \leq x_i \leq 365 \} \quad |\Omega| = 365^k$$

$$E = \text{"no collision"} = \{ (x_1, \dots, x_k) : \text{For all } i < j, x_i \neq x_j \}$$

$$|E| = 365 \cdot 364 \cdots (365 - k + 1)$$

$$\Pr[E] = \frac{|E|}{|\Omega|} = \frac{365 \cdot \cdots \cdot (365 - k + 1)}{365^k}$$

By tedious calculation when $k=23$, $\Pr[E] \leq 0.5$

and hence $\Pr[\text{exists a collision}] \geq 0.5$

Birthday Paradox

Assumption: Each student has a uniformly random birthday out of the 365 options.

k students.

$$\Omega = \{ (x_1, \dots, x_k) : \forall i \ 1 \leq x_i \leq 365 \}$$

$$E = \text{"no collision"} = \{ (x_1, \dots, x_k) : \text{For all } i < j, x_i \neq x_j \}$$

$$\bar{E} = \text{"}\exists \text{ collision"} = \{ (x_1, \dots, x_k) : \exists i < j \text{ s.t. } x_i = x_j \}$$

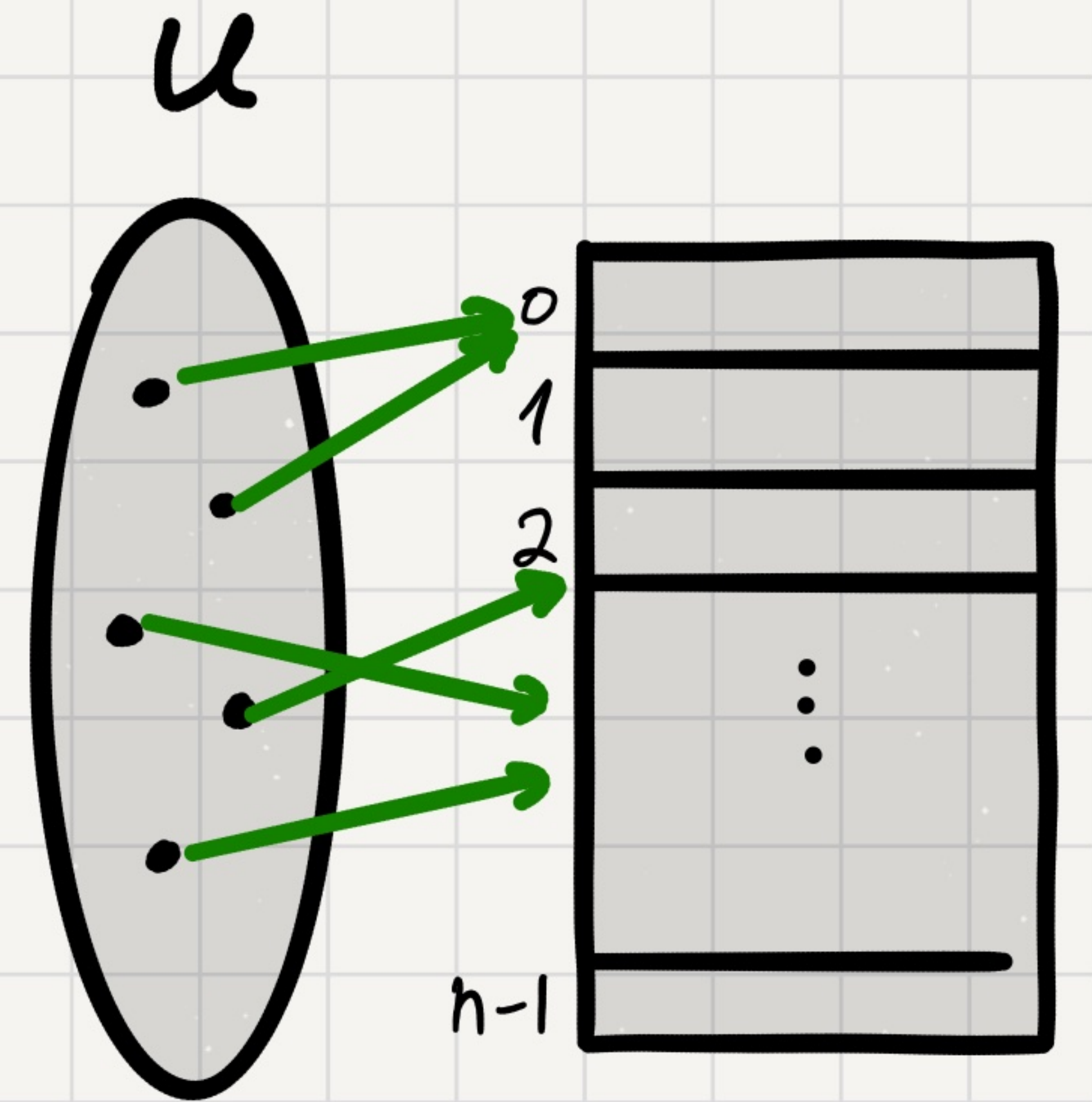
How to get a simple upper bound on $\Pr[\bar{E}]$?

For $1 \leq i < j \leq k$ let $A_{i,j}$ be the event indicating " $x_i = x_j$ "

$$\Pr[\bar{E}] = \Pr\left[\bigcup_{1 \leq i < j \leq k} A_{i,j}\right] \leq \sum_{1 \leq i < j \leq k} \Pr[A_{i,j}] = \binom{k}{2} \cdot \frac{1}{365}.$$

Birthday Paradox in Computing

You create a hash table of size n and add k elements to it using a random hash function.



Q: What's the prob. for a collision?

$$\Pr[\text{no collision}] = \frac{n \cdot \dots \cdot (n-k+1)}{n^k}$$

$$\Pr[\exists \text{ collision}] \leq \sum_{1 \leq i < j \leq k} \Pr[\text{elements } i \& j \text{ collide}] = \binom{k}{2} \cdot \frac{1}{n}.$$

n balls in n bins

What's the maximum capacity?

Let j be a parameter.

Denote by $A_j =$ "there exists a bin with $\geq j$ balls"

$\Pr[A_j] = ?$

Let $A_{j,1}, \dots, A_{j,n}$ be events $A_{j,i} =$ "bin # i has $\geq j$ balls"

$A_j = A_{j,1} \vee \dots \vee A_{j,n}$ $\Pr[A_j] \leq \sum_{i=1}^n \Pr[A_{j,i}]$

n balls in n bins

What's the maximum capacity?

$A_{j,i}$ = "bin # i has $\geq j$ balls" A_j = " \exists bin with $\geq j$ balls"

$$\Pr[A_{j,i}] \leq \frac{\binom{n}{j} \cdot n^{n-j}}{n^n} = \binom{n}{j} \cdot \frac{1}{n^j} = \frac{n \cdot \dots \cdot (n-j+1)}{j! n^j} \leq \frac{1}{j!}$$

$$\Pr[A_j] \leq \Pr[A_{j,1}] + \dots + \Pr[A_{j,n}] \leq n \cdot \frac{1}{j!}$$

For example for $n = 1,000,000$, $j = 10$, $\Pr[A_j] \leq 0.28$

Independence

We say that events A and B are independent if $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.

Example: Toss two fair coins

$A =$ "first coin is H"

$B =$ "second coin is H"

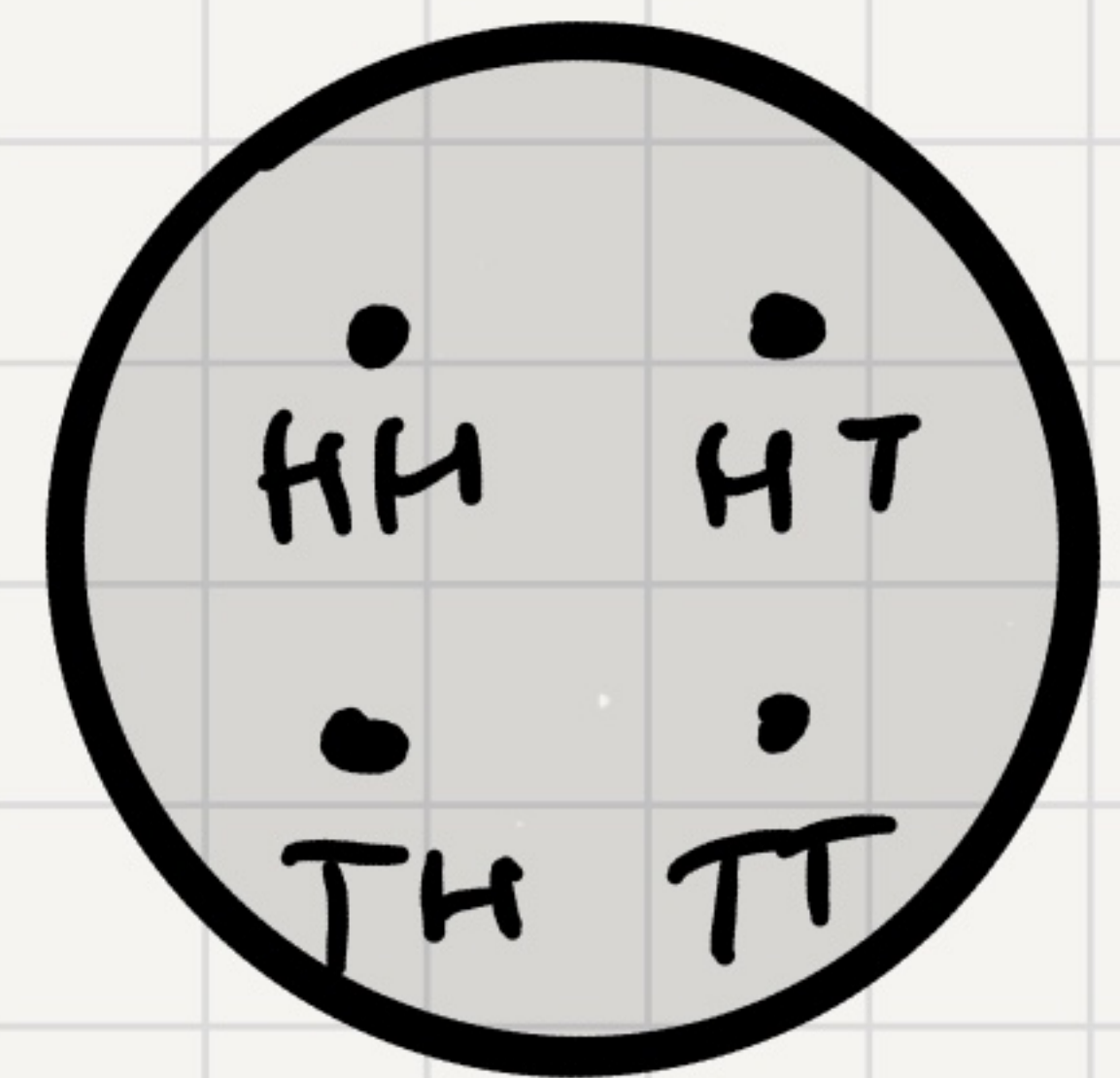
$C =$ "same result"

Poll: Which are independent

1: A, B

2: A, C

3: B, C



Independence

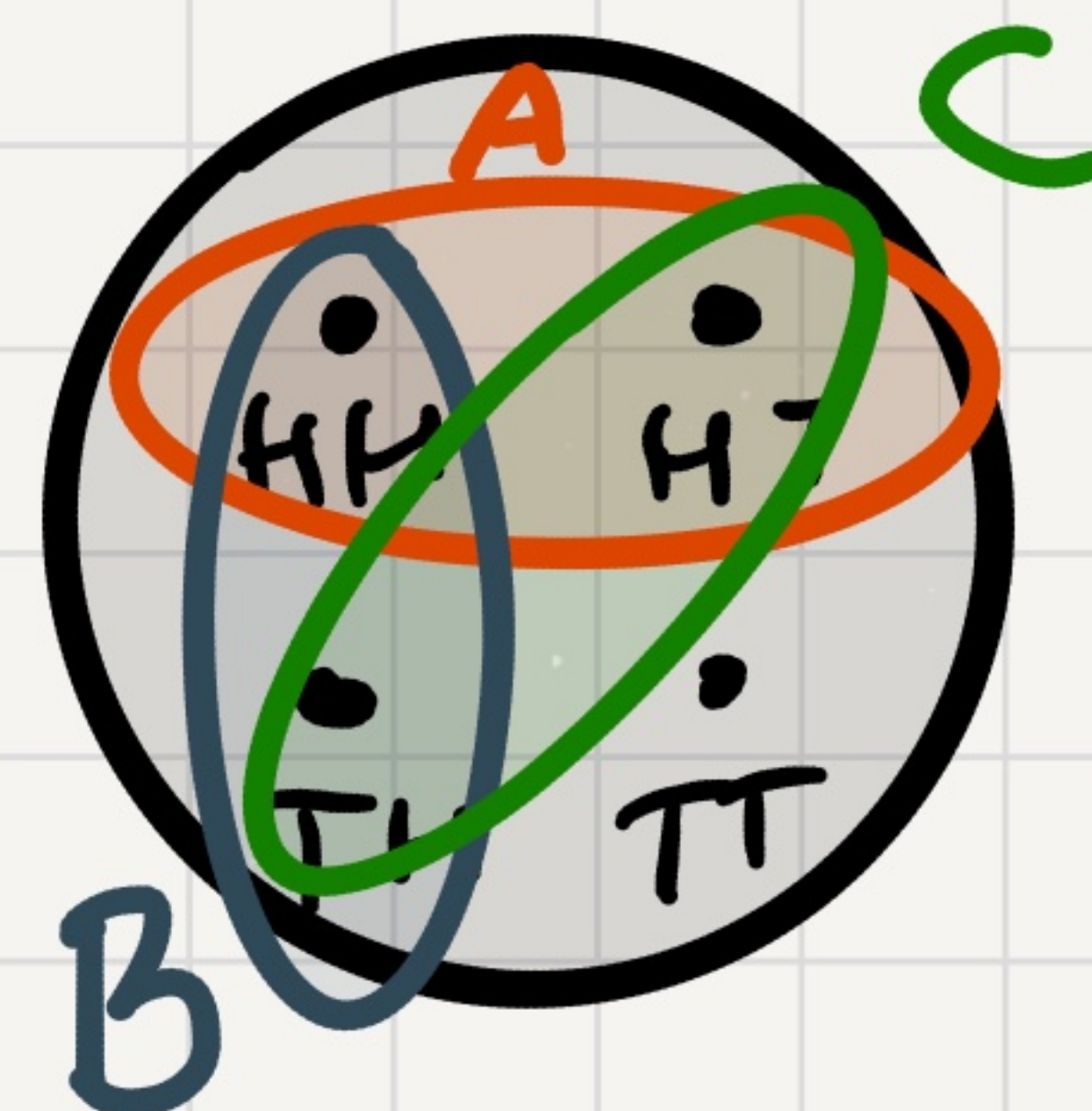
We say that events A and B are independent if $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.

Example: Toss two fair coins

$A =$ "first coin is H"

$B =$ "second coin is H"

$C =$ "same result"



A & B are "indep."

A & C are "indep."

B & C are "indep."

Independence

We say that events A and B are independent if $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.

Example: Toss two fair coins

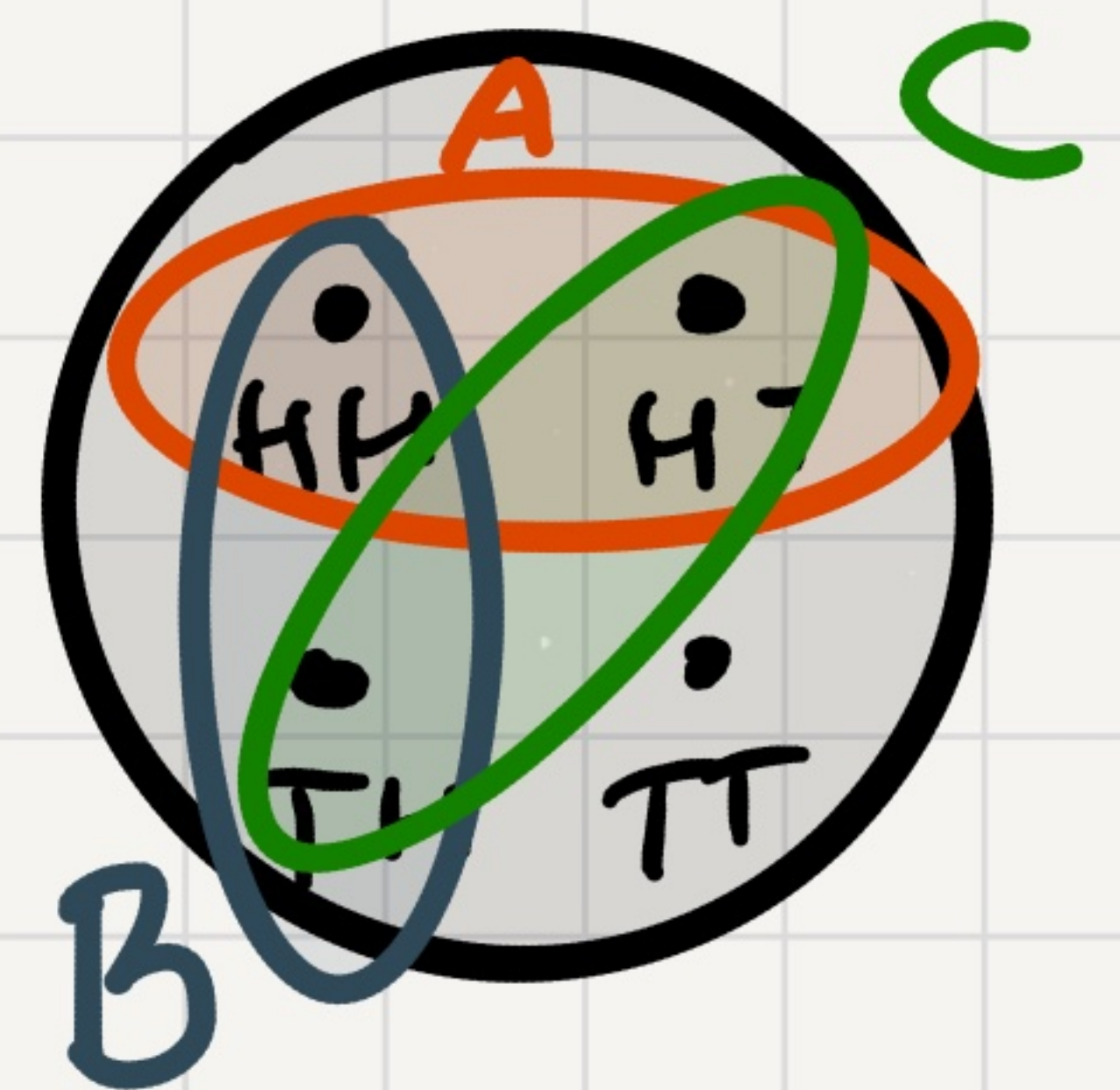
$A =$ "first coin is H"

$B =$ "second coin is H"

$C =$ "same result"

Are A, B, C independent together?

$$\Pr[C | A \cap B] = 1$$



Definition Pairwise Independence

We say that events A_1, \dots, A_n are pairwise independent if

$\forall i \neq j$ A_i and A_j are independent.

Definition Mutual Independence

We say that events A_1, A_2, A_3 are **mutually** independent if

they are pairwise indep. and $P_r[A_1 \cap A_2 \cap A_3] = P_r[A_1] \cdot P_r[A_2] \cdot P_r[A_3]$

Definition Pairwise Independence

We say that events A_1, \dots, A_n are pairwise independent if

$\forall i \neq j$ A_i and A_j are independent.

Definition Mutual Independence

We say that events A_1, \dots, A_n are **mutually** independent if for each

non-empty subset $I \subseteq \{1, \dots, n\}$ $\Pr\left[\bigcap_{i \in I} A_i\right] = \prod_{i \in I} \Pr[A_i]$

Secret Sharing

n shares threshold: 3

Degree 2 polynomial

How to pick the polynomial: let $p > n$ be a prime

1. Pick $s = a_0 \in \{0, 1, \dots, p-1\}$ uniformly at random.

2. Pick $a_1, a_2 \in \{0, 1, \dots, p-1\}$ " " "

$$f(x) = a_2 x^2 + a_1 x + a_0$$

Share i : $f(i)$

$$\Pr [s = a \mid f(x_1) = y_1, f(x_2) = y_2] =$$

$$= \frac{\Pr [s = a, f(x_1) = y_1, f(x_2) = y_2]}{\Pr [f(x_1) = y_1, f(x_2) = y_2]} = \frac{1/p^3}{p/p^3} = 1/p$$

Secret Sharing

n shares threshold: 3

Degree 2 polynomial

How to pick the polynomial: let $p > n$ be a prime

1. Pick $s = a_0 \in \{0, 1, \dots, p-1\}$ uniformly at random.

2. Pick $a_1, a_2 \in \{0, 1, \dots, p-1\}$ " " "

$$f(x) = a_2 x^2 + a_1 x + a_0$$

Share i : $f(i)$

Every three shares are independent:

$$\Pr[f(x_1) = y_1 \cap f(x_2) = y_2 \cap f(x_3) = y_3] = \frac{1}{p^3}$$

but 4 shares are not.