6.034 FALL 2020

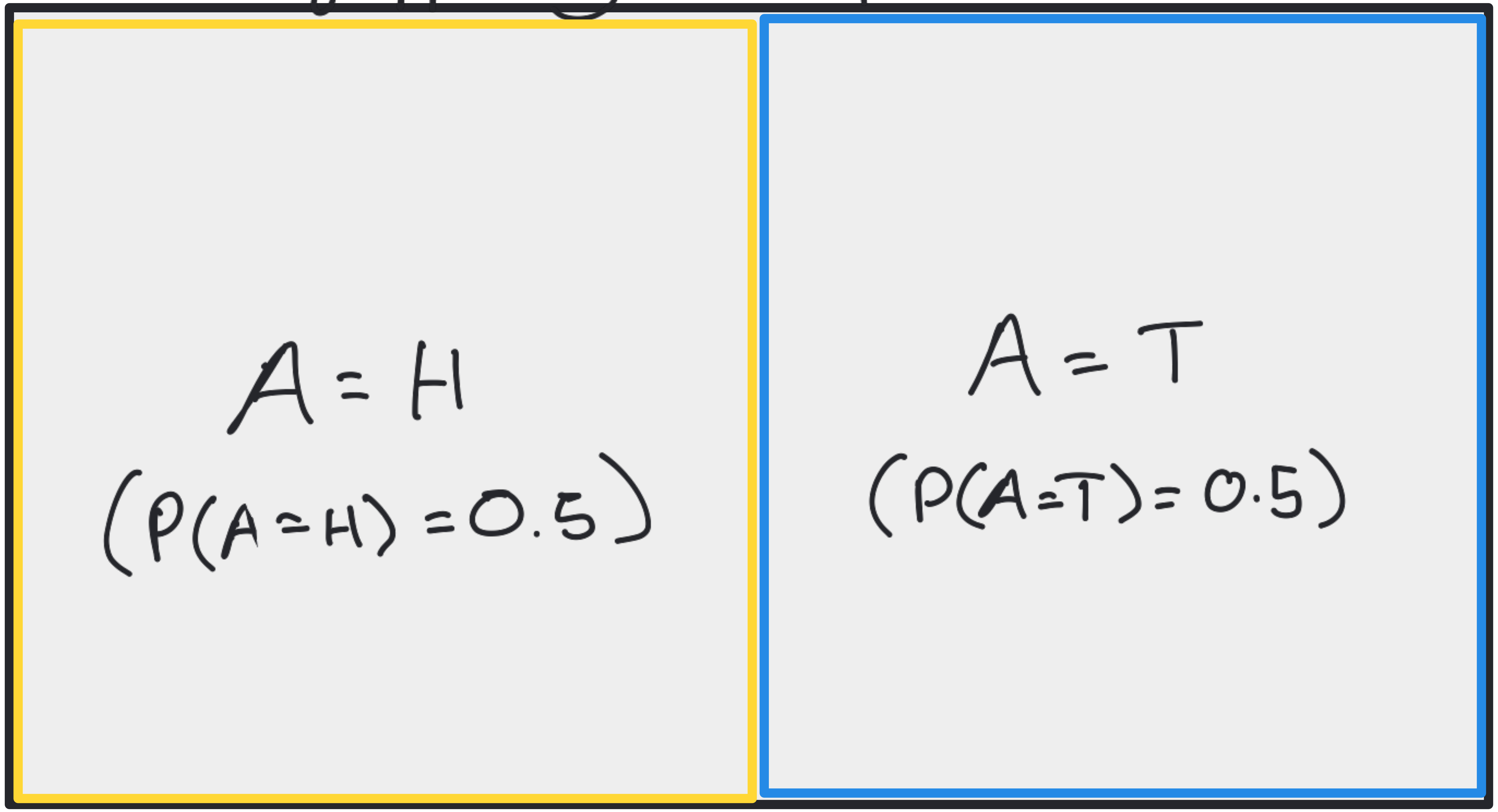
RECITATION 5: BAYES pt. 1

- ☐ PROBABILITY CRASH COURSE
- ☐ BAYES NETS

PROBABILITY: Variables & Events

A coin flipping example

Ω (1 coin flip)



$A :=$ the result of flipping a coin ($A=H$ or $A=T$)

PROBABILITY: Variables & Events

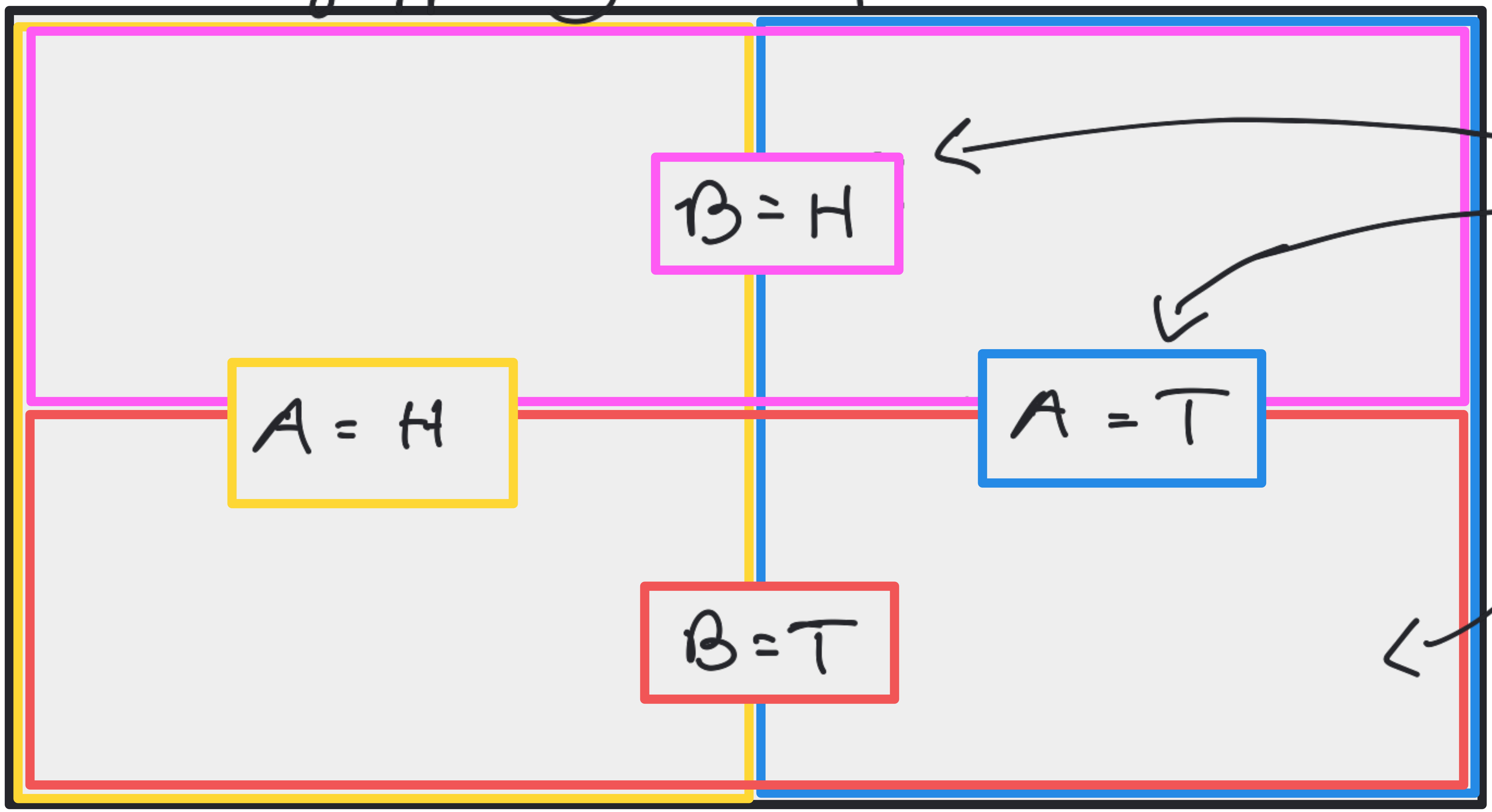
- Variables: A, B, C ...
- Events: variables take on a value
  - A = a discrete* (* simplifying assumption)
  - $A = true \rightarrow A$
  - $A = false \rightarrow \overline{A}$

  } boolean

PROBABILITY: Types of Probability

A coin flipping example

Ω (2 coin flips)



Joint: $P(A=x, B=y)$

Marginal: $P(B=x)$ or $P(A=x)$ alone

$A :=$ the result of the 1st coin flip

$B :=$ the result of the 2nd coin flip

PROBABILITY: Types of Probability

$\Omega = \{A, B, C\}$

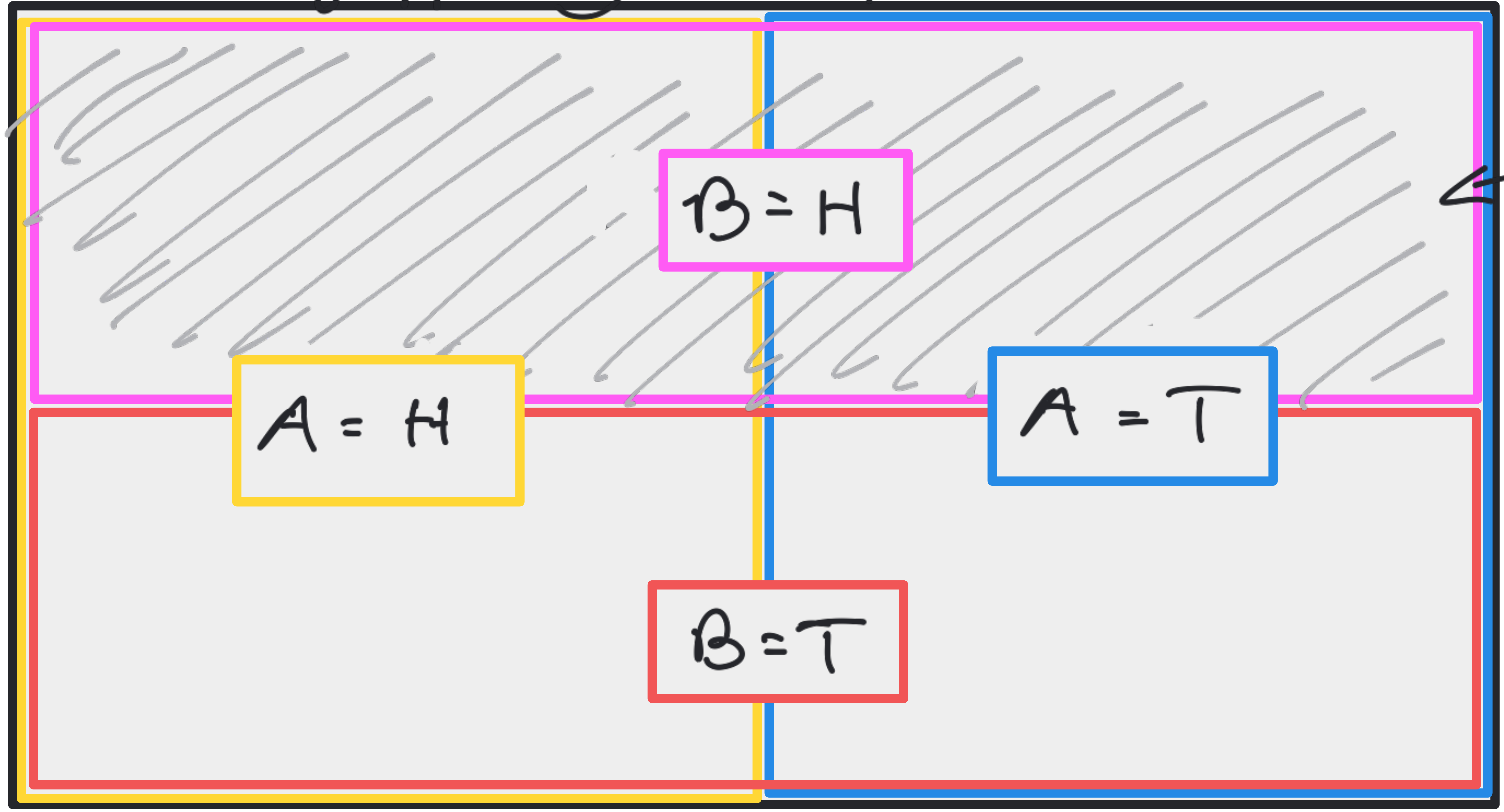
- Joint
  - Probability of an event that encompasses all the variables in Ω
  - $P(A=a, B=b, C=c)$
- Marginal
  - Probability of an event that defines an incomplete set of variables in Ω
  - $P(A=a) = \sum_b \sum_c P(A=a, B=b, C=c)$

the sum of joint probabilities

PROBABILITY: Types of Probability

A coin flipping example

Ω (2 coin flips)



(don't care)

Conditional

$P(A=x \mid \underline{B=y})$

new area of interest

A := the result of the 1st coin flip

B := the result of the 2nd coin flip

PROBABILITY: Types of Probability

$\Omega = \{A, B, C\}$

- ☐ Joint
  - ☐ $P(A=a, B=b, C=c)$
- ☐ Marginal
  - ☐ $P(A=a) = \sum_b \sum_c P(A=a, B=b, C=c)$
- ☐ Conditional
  - ☐ Likelihood of an event given another
  - ☐ $P(A=a \mid B=b, C=c) = \dfrac{P(A=a, B=b, C=c)}{P(B=b, C=c)}$ $\left(\dfrac{\text{joint}}{\text{marginal}}\right)$

PROBABILITY: Chain & Bayes Rules

$\Omega = \{A, B, C\}$

- [ ] Conditional
  - [ ] $P(A=a \mid B=b, C=c) = \dfrac{P(A=a, B=b, C=c)}{P(B=b, C=c)}$
- [ ] Chain rule
  - [ ] $P(A=a, B=b, C=c) = P(A=a \mid B=b, C=c) P(B=b, C=c)$
  - [ ] $P(A=a, B=b, C=c) = P(A=a \mid B=b, C=c) \cdot P(B=b \mid C=c) \cdot P(C=c)$

(also equivalent)

- [ ] $P(A=a, B=b, C=c) = P(C=c \mid A=a, B=b) \cdot P(A=a \mid B=b) \cdot P(B=b)$

PROBABILITY: Chain & Bayes Rules

$\Omega = \{A, B\}$

- ☐ Chain rule
  - ☐ $P(A=a, B=b) = P(A=a \mid B=b)\,P(B=b)$
    $= P(B=b \mid A=a)\,P(A=a)$
  - ☐ $P(A=a \mid B=b)\,P(B=b) = P(B=b \mid A=a)\,P(A=a)$
- ☐ Bayes Rule
  - ☐ $P(A=a \mid B=b) = \dfrac{P(B=b \mid A=a)\,P(A=a)}{P(B=b)}$

PROBABILITY: Exhaustion

$\Omega = \{A\}$

- [ ] The sum of the probabilities of all events in Ω is equal to 1.0
  - [ ] $\sum(\text{probs}) = 1$ (2nd axiom of probability)
  - [ ] boolean: $P(A=\text{True}) + P(A=\text{False}) = 1.0$
  - [ ] discrete: $P(A=a_1) + P(A=a_2) + \ldots + P(A=a_n) = 1.0$
  - [ ] conditional: $P(A=a_1 \mid \underline{B=b}) + \ldots + P(A=a_n \mid \underline{B=b}) = 1.0$

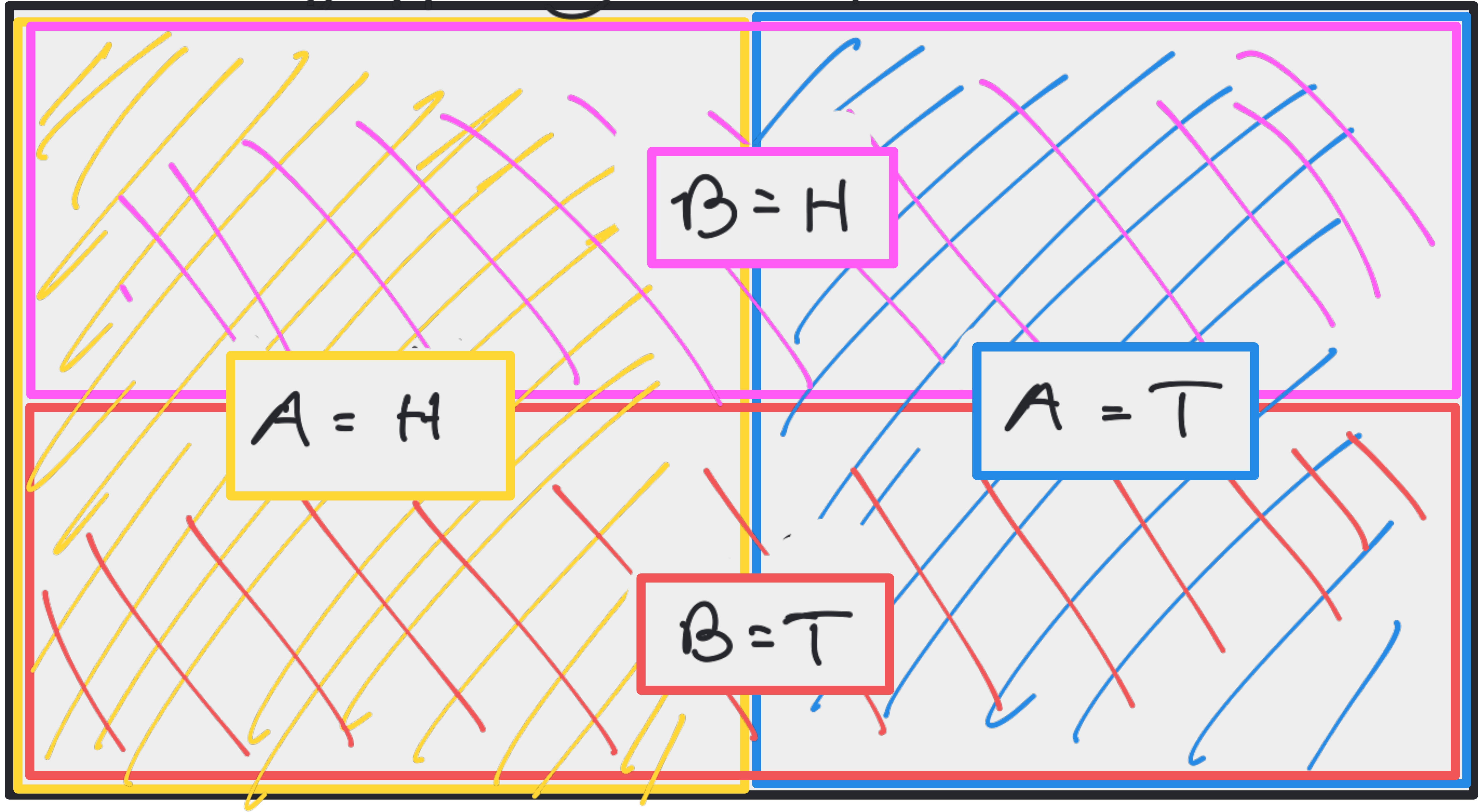
$\Omega = \{A, B\}$

$\Omega \rightarrow B$

PROBABILITY: Independence

A coin flipping example

Ω (2 coin flips)



$P(A=x \mid B=y) = P(A=x)$

$A :=$ the result of the 1st coin flip

$B :=$ the result of the 2nd coin flip

PROBABILITY: Independence

Marginal Independence

- ☐ A & B are independent ($A \perp\!\!\!\perp B$) if and only if (iff)

  $P(A|B) = P(A)$

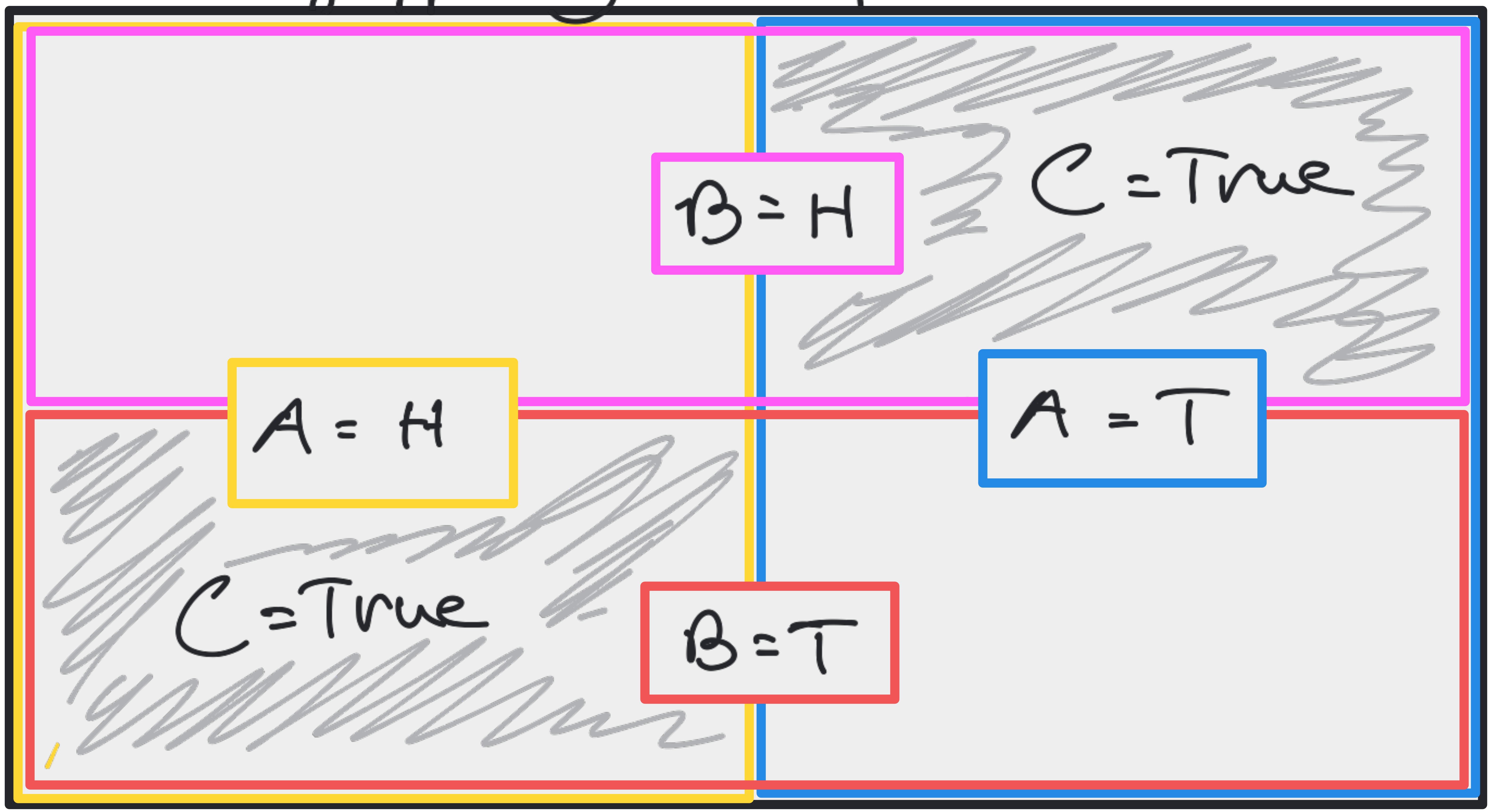
- ☐ Knowledge of one variable does not inform you of the other variable

$\Rightarrow$ ☐ $P(A,B) = P(A|B)P(B) = P(A)P(B)$

PROBABILITY: Independence

A coin flipping example

Ω (2 coin flips)



Think:

$P(A|C)$?

$P(A|B,C)$?

Are they equal?

A := the result of the 1st coin flip

B := the result of the 2nd coin flip

C := the event that both Heads and Tails occurred

PROBABILITY: Independence

Conditional Independence

- ☐ A & B are conditionally independent given C ($A \perp\!\!\!\perp B \mid C$) iff $P(A \mid B, C) = P(A \mid C)$

- ☐ Knowledge of C makes it such that knowledge of B does not inform you of A (& vice versa)

$\Rightarrow$ ☐ $P(A, B \mid C) = P(A \mid C)\,P(B \mid C)$

PROBABILITY: # of parameters

$\Omega = \{A, B\}$
(Boolean)

if $A \not\perp\!\!\!\perp B$:

A	B	$P(A,B)$
T	T	p_1
T	F	p_2
F	T	p_3
F	F	$1-(p_1+p_2+p_3)$

need 3 parameters

if $A \perp\!\!\!\perp B$

A	B	$P(A)$	$P(B)$	$P(A,B)$
T	T	p_1	p_2	$p_1 p_2$
T	F	p_1	$1-p_2$	$p_1(1-p_2)$
F	T	$1-p_1$	p_2	$(1-p_1)p_2$
F	F	$1-p_1$	$1-p_2$	$(1-p_1)(1-p_2)$

2 parameters

PROBABILITY: # of parameters

$\Omega = \{A, B\}$
(Boolean)

if $A \not\perp\!\!\!\perp B$:

params $= \left[\prod_i^n \#\text{vals of } i\right] - 1$

if A & B Boolean: $2^n - 1$

if $A \perp\!\!\!\perp B$

params $= \sum^n \left[(\#\text{vals of } i) - 1\right]$

if A & B Boolean: n

BAYES NETS: Bayes Net assumption

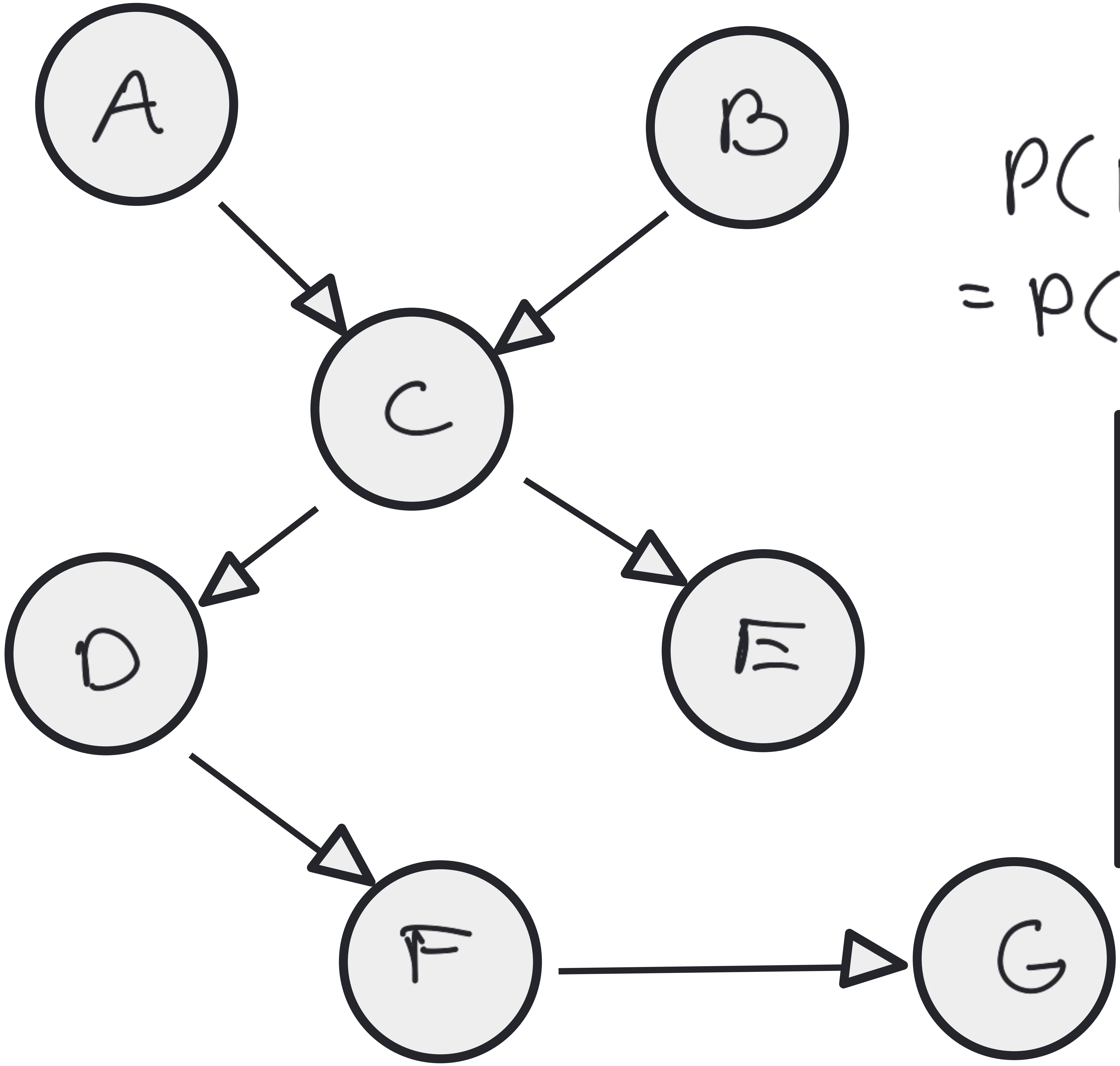
<u>Def:</u> Bayes Net

- ☐ directed, acyclic graph
- ☐ represents the independence assumption

<u>Assumption</u>

★ Each variable in the Bayes net is conditionally independent of its nondescendants given its parents

BAYES NETS: Bayes Net assumption

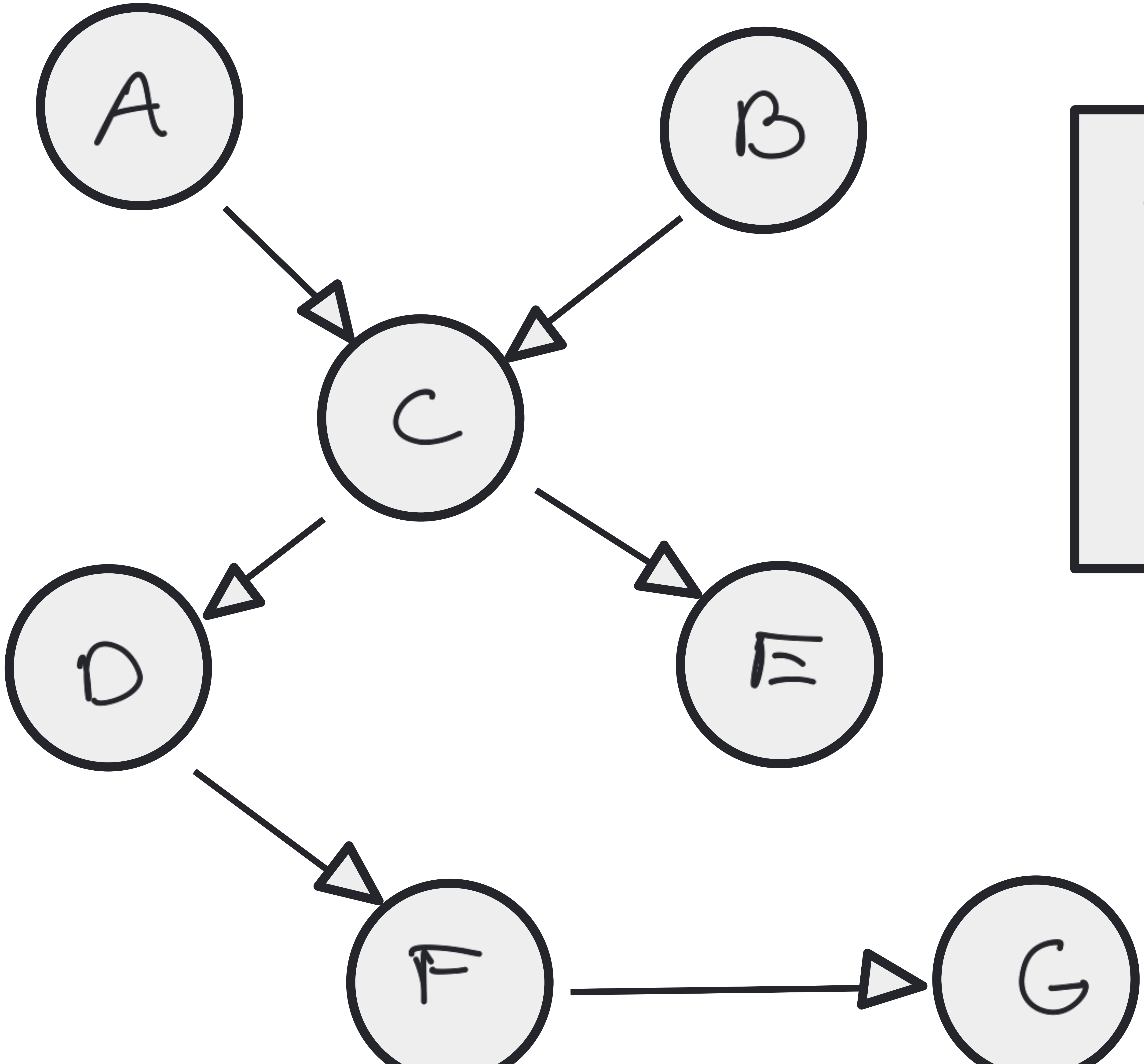


$$P(F \mid \text{nondesc}(F), \text{parents}(F))$$
$$= P(F \mid \text{parents}(F))$$

$\text{nondesc}(F) = A, B, C, D, E$
$\text{parents}(F) = D$

$$\Rightarrow P(F \mid \cancel{A}, \cancel{B}, \cancel{C}, D, \cancel{E})$$
$$= P(F \mid D)$$

BAYES NETS: d-separation



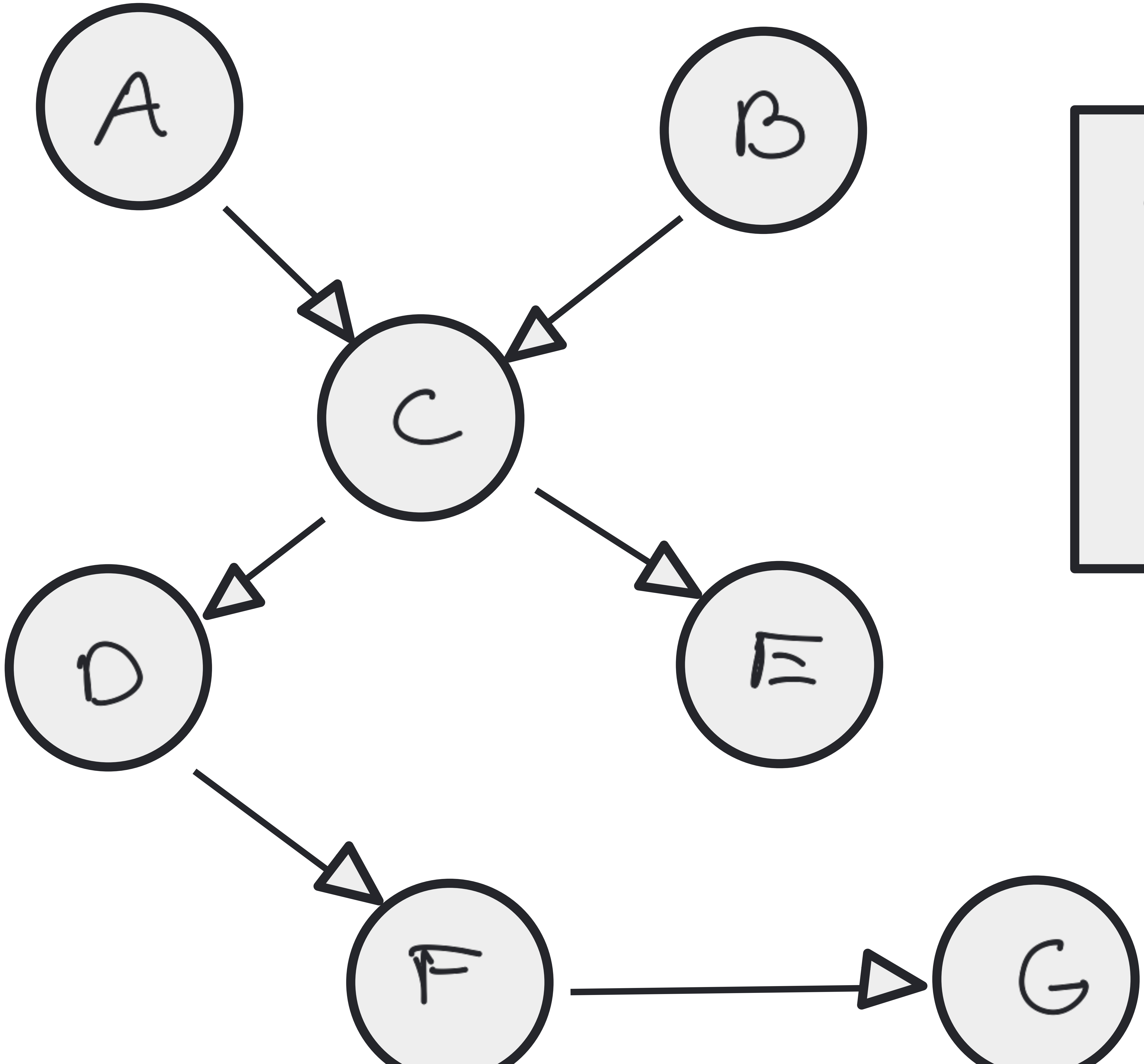
Want to know:

- ☐ $A \perp\!\!\!\perp B \mid D, F$?
- ☐ $P(A \mid B, D, F) \stackrel{?}{=} P(A \mid D, F)$?

BAYES NETS: d-separation

1) Draw an ancestral graph (NO non-ancestors)
2) Connect each set of parents in the graph
   - □ find variables that share a child and draw an edge between them
3) Disorientation
   - □ Make the graph undirected
4) Delete givens and their edges
5) Read the graph
   - □ if there is a path between 2 variables => $\not\perp\!\!\!\perp$
   - □ if 2 variables are completely disconnected => $\perp\!\!\!\perp^*$

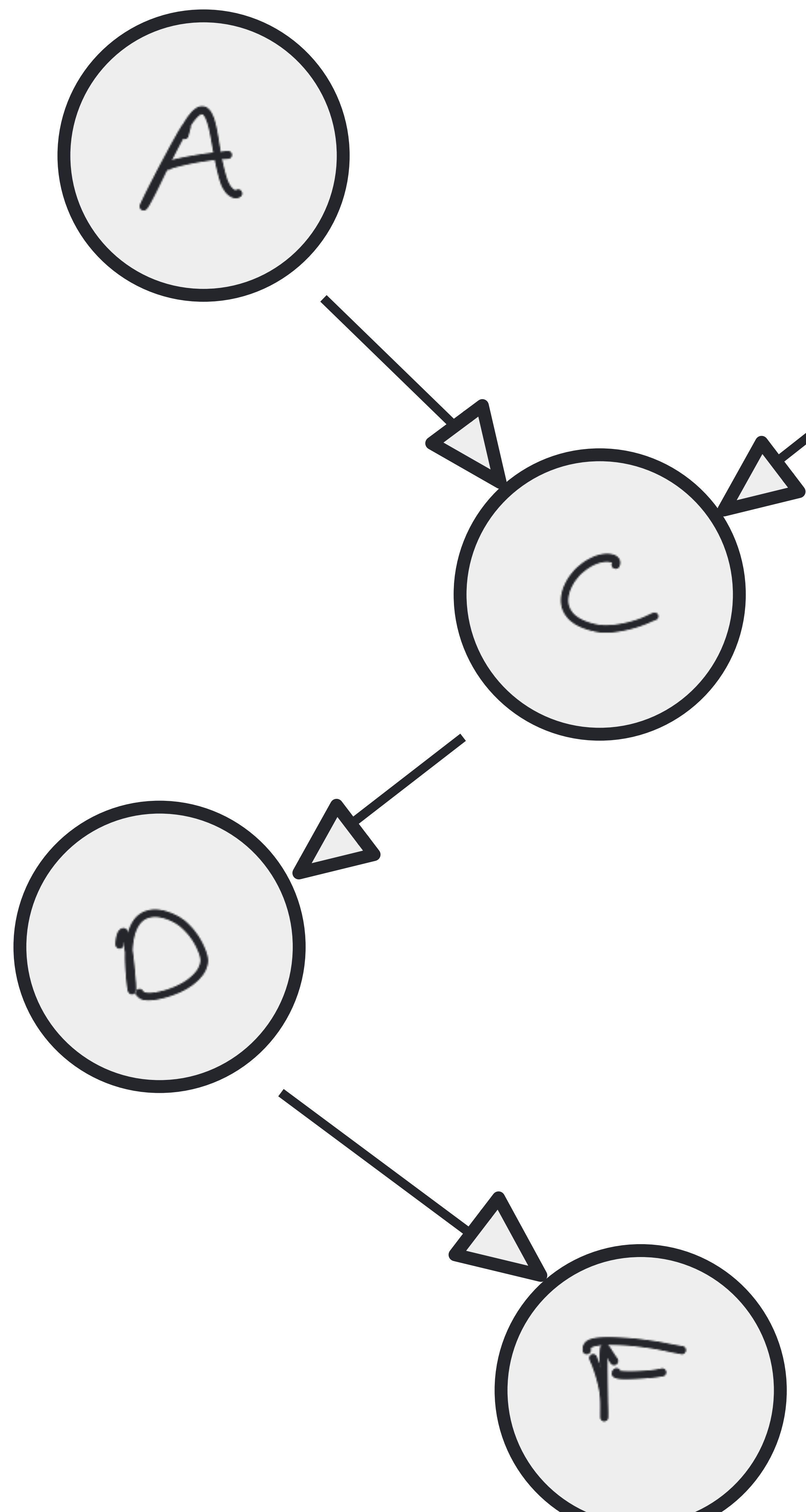
BAYES NETS: d-separation



Want to know:

- ☐ $A \perp\!\!\!\perp B \mid D, F$?
- ☐ $P(A \mid B, D, F) \stackrel{?}{=} P(A \mid D, F)$?

BAYES NETS: d-separation

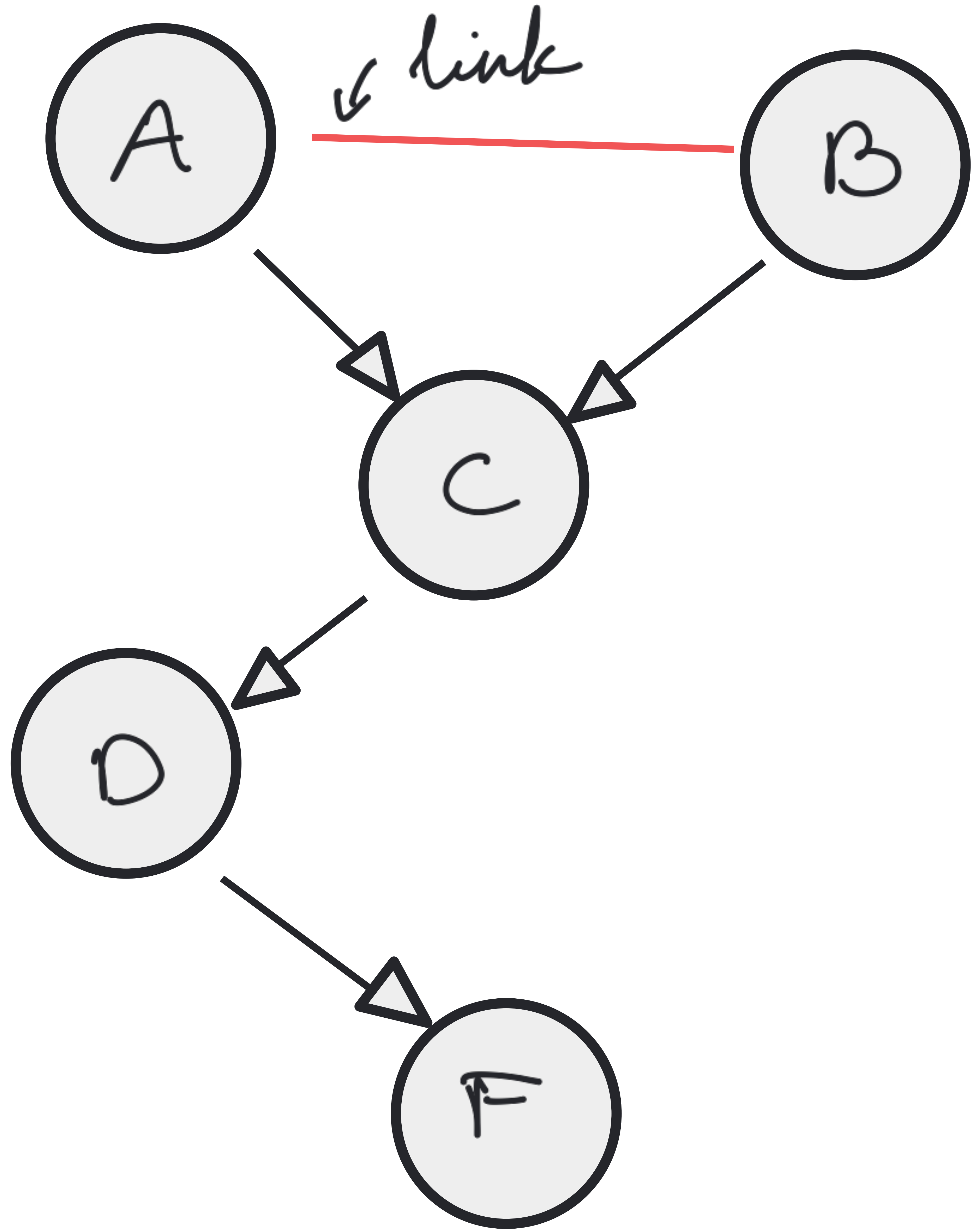


Want to know:

- ☐ $A \perp\!\!\!\perp B \mid D, F$?
- ☐ $P(A \mid B, D, F) \stackrel{?}{=} P(A \mid D, F)$?

① Draw an ancestral graph

BAYES NETS: d-separation



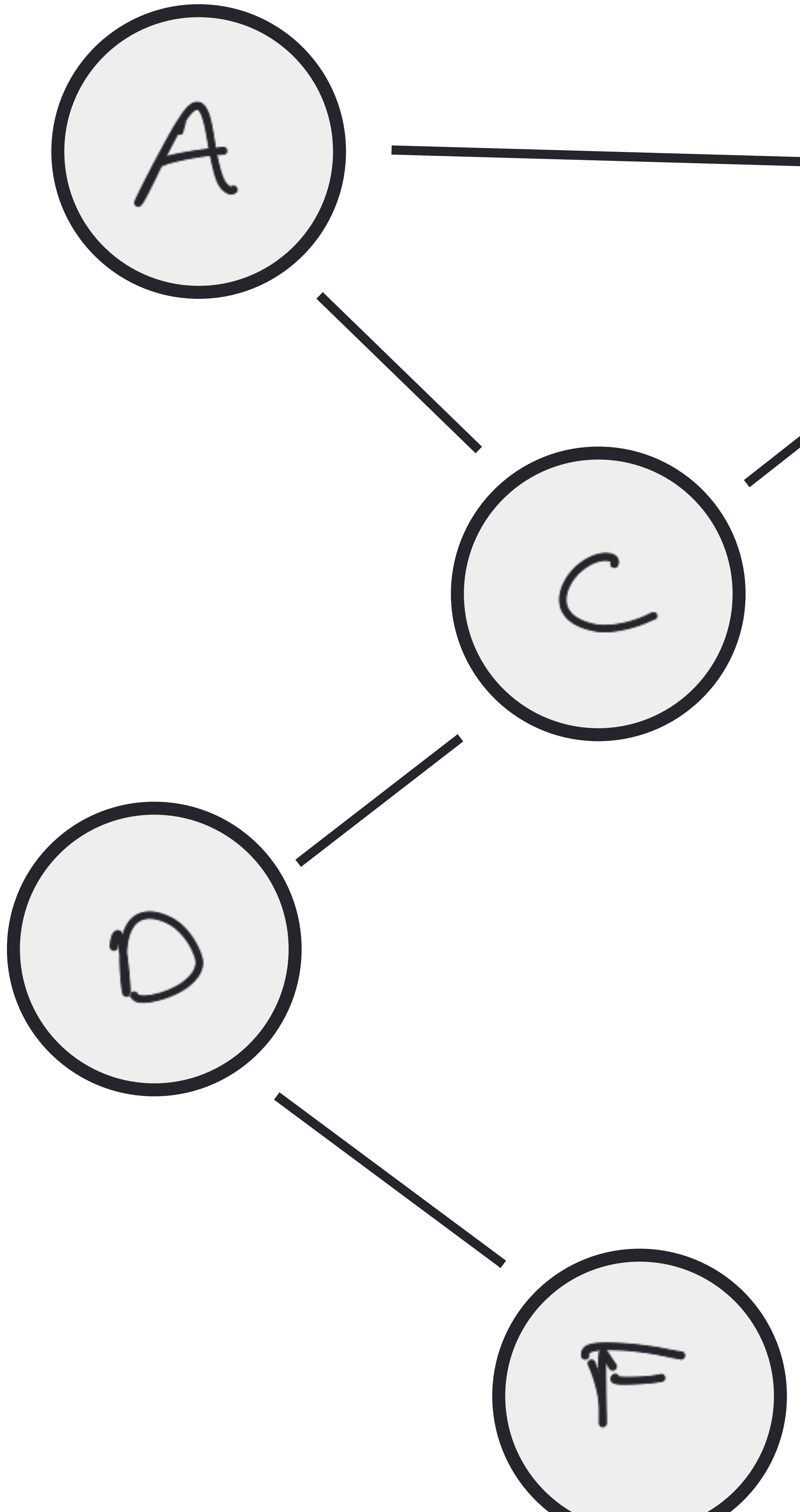
Want to know:

☐ $A \perp\!\!\!\perp B \mid D, F$?

☐ $P(A \mid B, D, F) \stackrel{?}{=} P(A \mid D, F)$?

② link the parents

BAYES NETS: d-separation

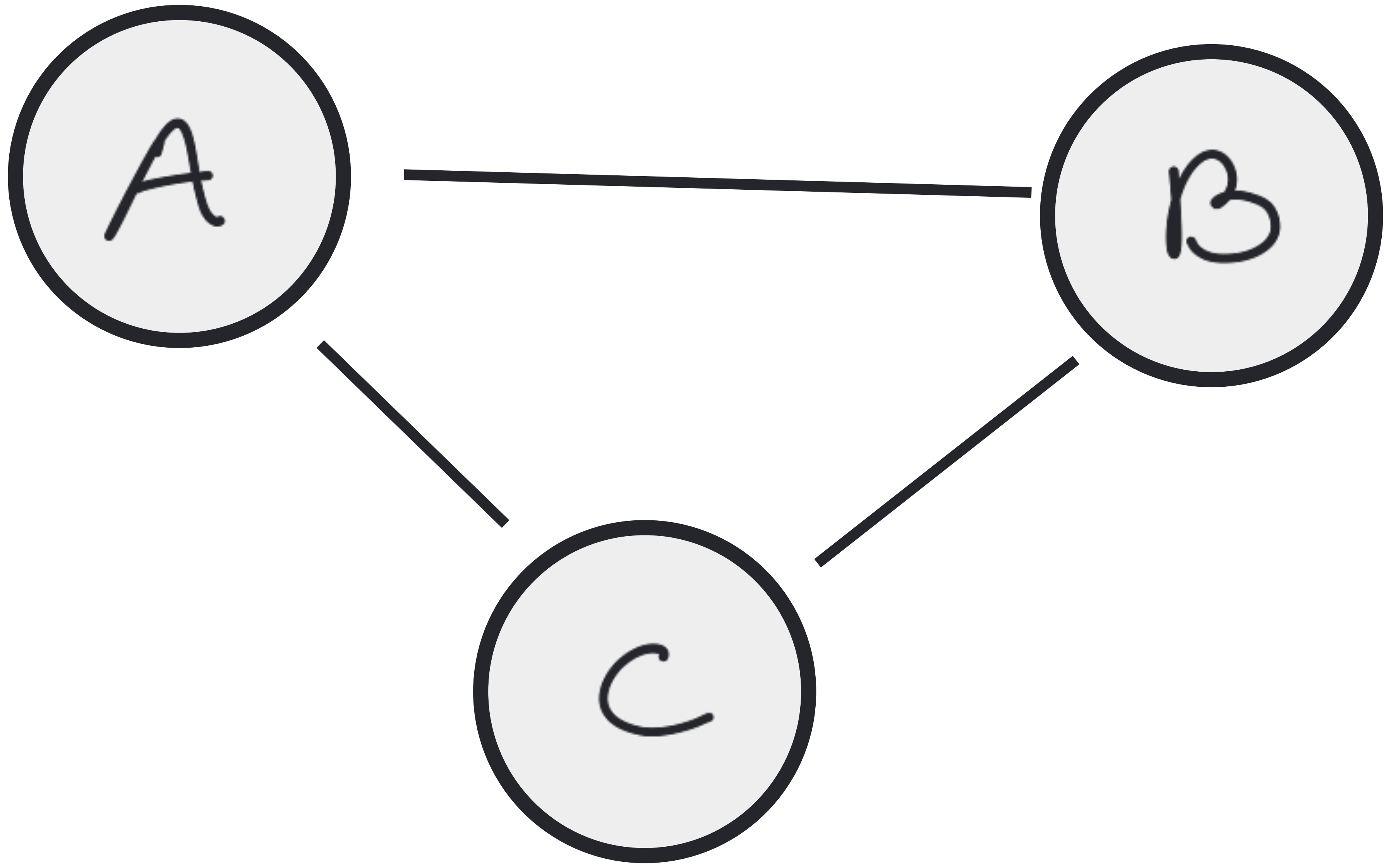


Want to know:

- ☐ $A \perp\!\!\!\perp B \mid D, F$?
- ☐ $P(A \mid B, D, F) \overset{?}{=} P(A \mid D, F)$?

③ Disorientation

BAYES NETS: d-separation

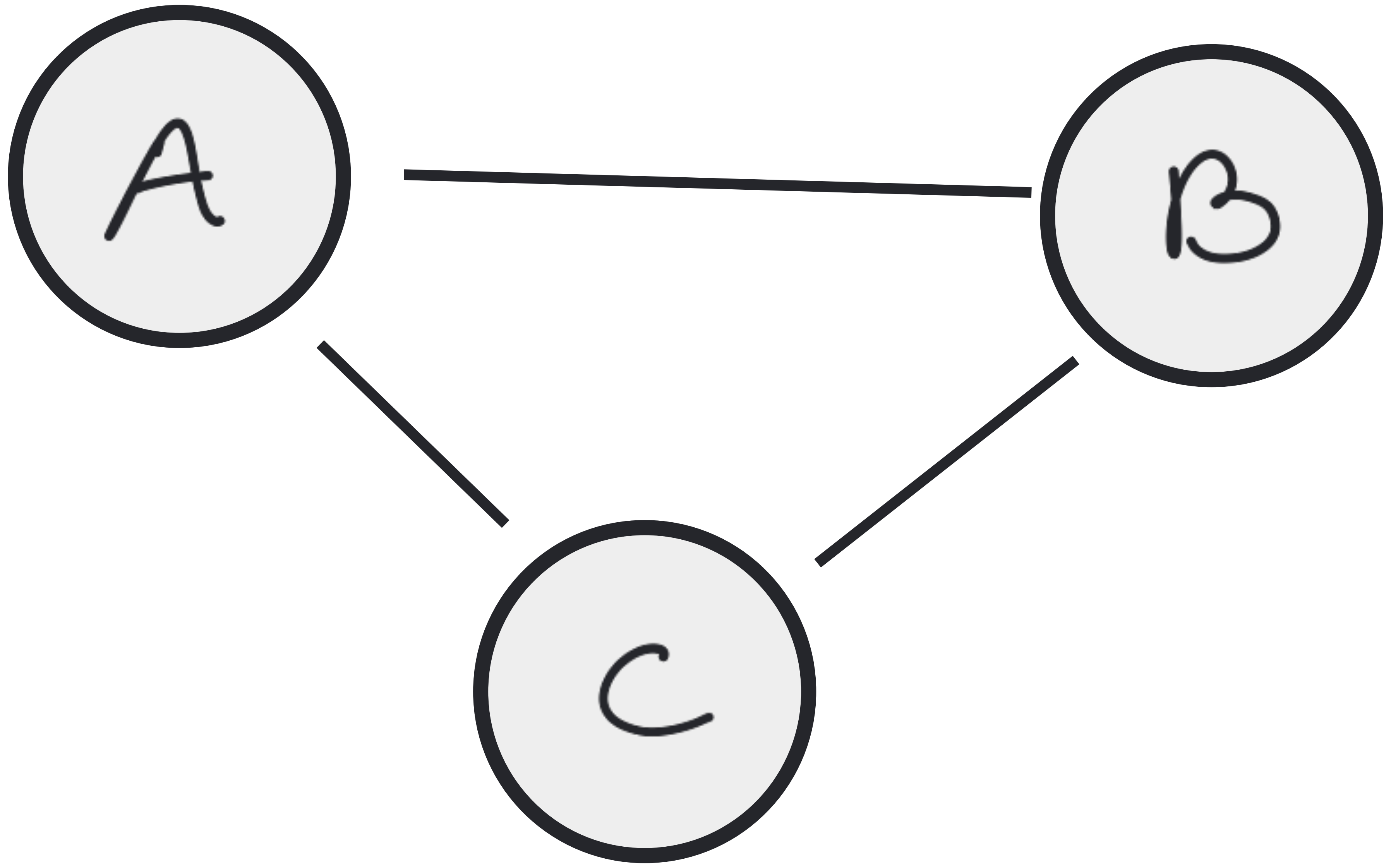


Want to know:

- ☐ $A \perp\!\!\!\perp B \mid D, F$?
- ☐ $P(A \mid B, D, F) \stackrel{?}{=} P(A \mid D, F)$?

④ Delete givens & their edges

BAYES NETS: d-separation



Want to know:

- ☐ $A \perp\!\!\!\perp B \mid D, F$?
- ☐ $P(A \mid B, D, F) \stackrel{?}{=} P(A \mid D, F)$?

⑤ Read the graph

Conclusion: $A \not\!\perp\!\!\!\perp B \mid D, F$ *

* A & B are structurally $\not\!\perp\!\!\!\perp$, but mathematically may still be $\perp\!\!\!\perp$

BAYES NETS: Explaining away

- Consider two different ailments: the flue and a broken arm.

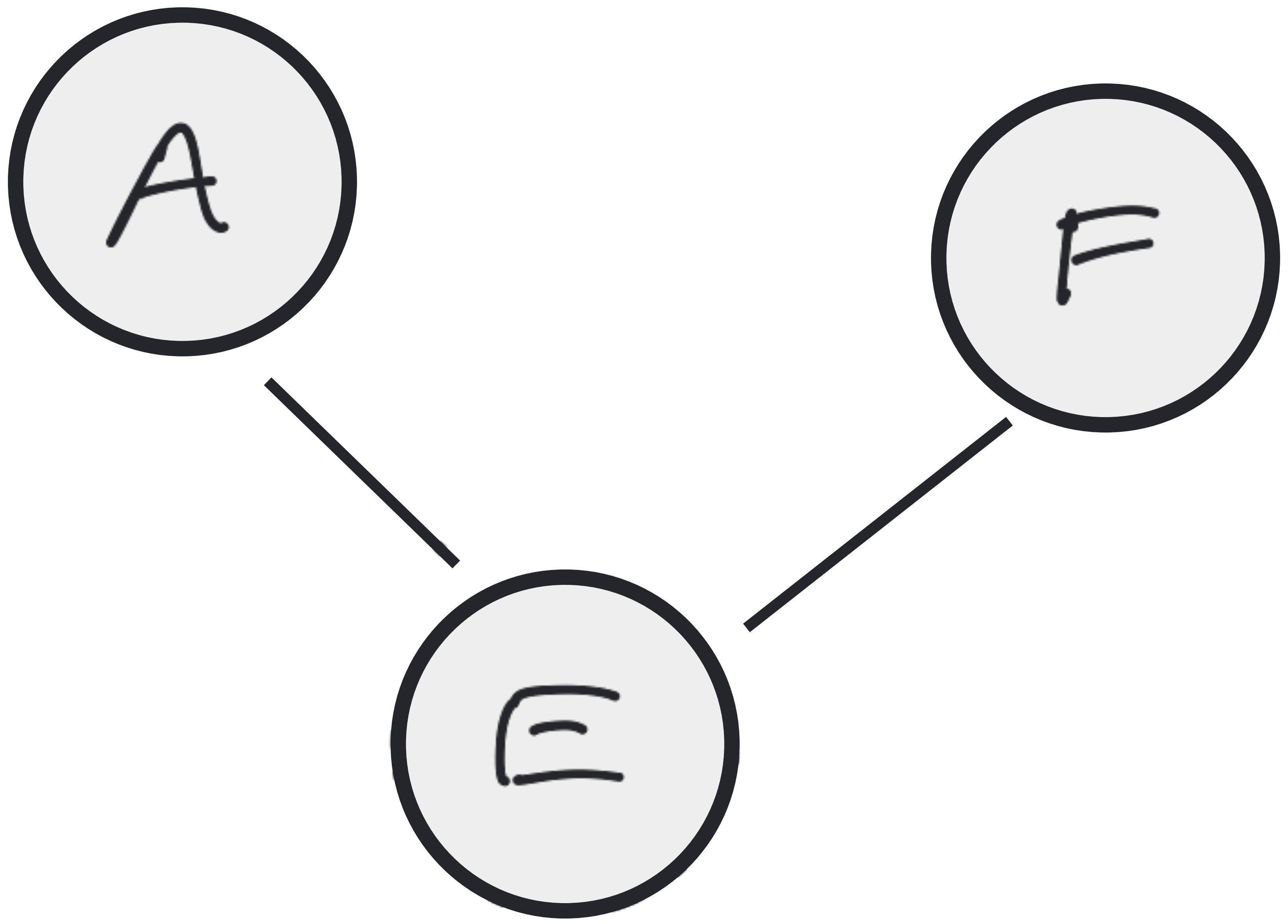
Different ailments
so $A \perp\!\!\!\perp F$

$A :=$ the event that Héctor has a broken arm

$F :=$ the event that Héctor has the flue

BAYES NETS: Explaining away

- Consider two different ailments: the flue and a broken arm.



A:= the event that Hector has a broken arm

F:= the event that Héctor has the flue

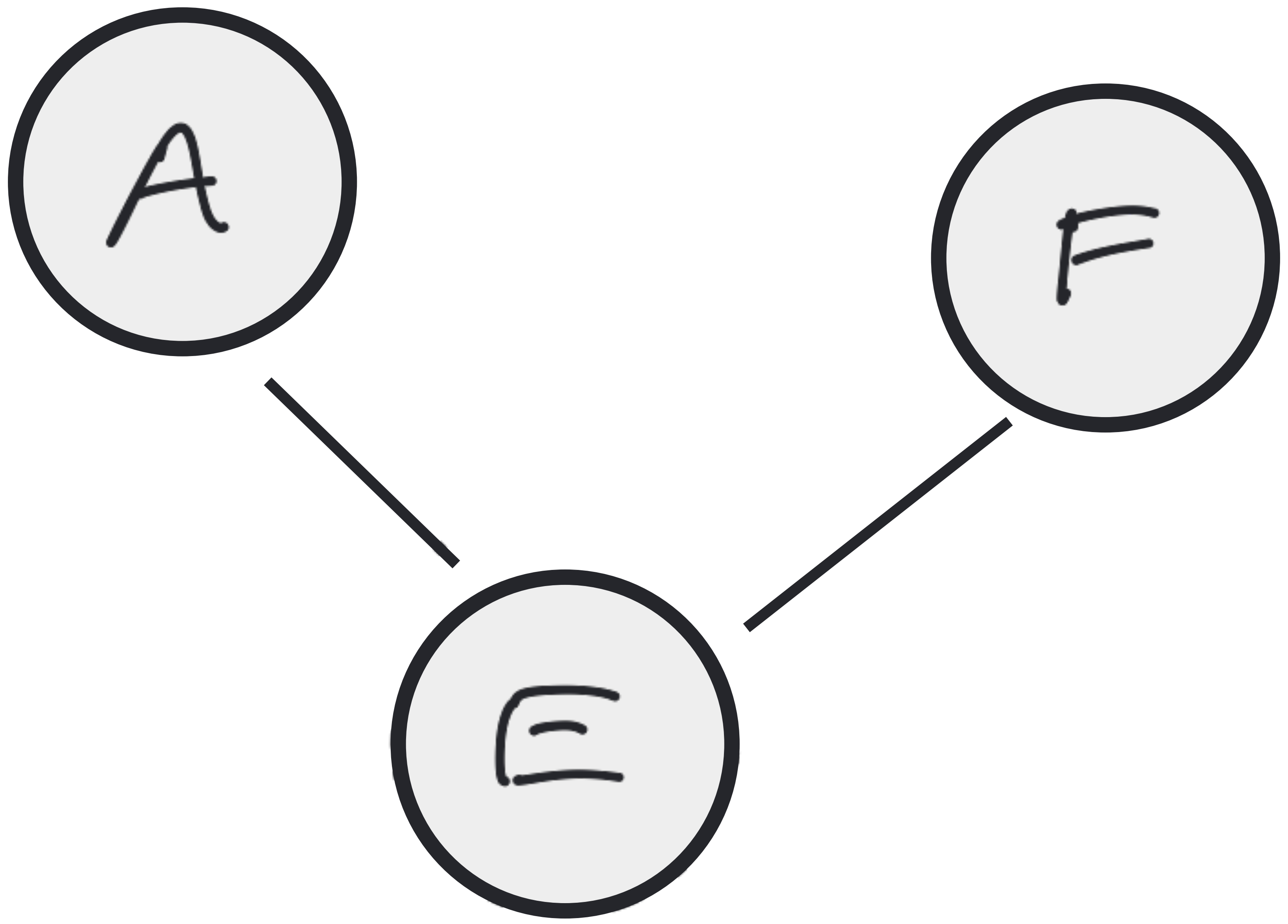
E:= the event that Héctor is in the Emergency Room

Think: what is $P(A)$ compared to $P(A|E)$?

what is $P(B)$ compared to $P(B|E$?

BAYES NETS: Explaining away

- Consider two different ailments: the flue and a broken arm.



A:= the event that Héctor has a broken arm

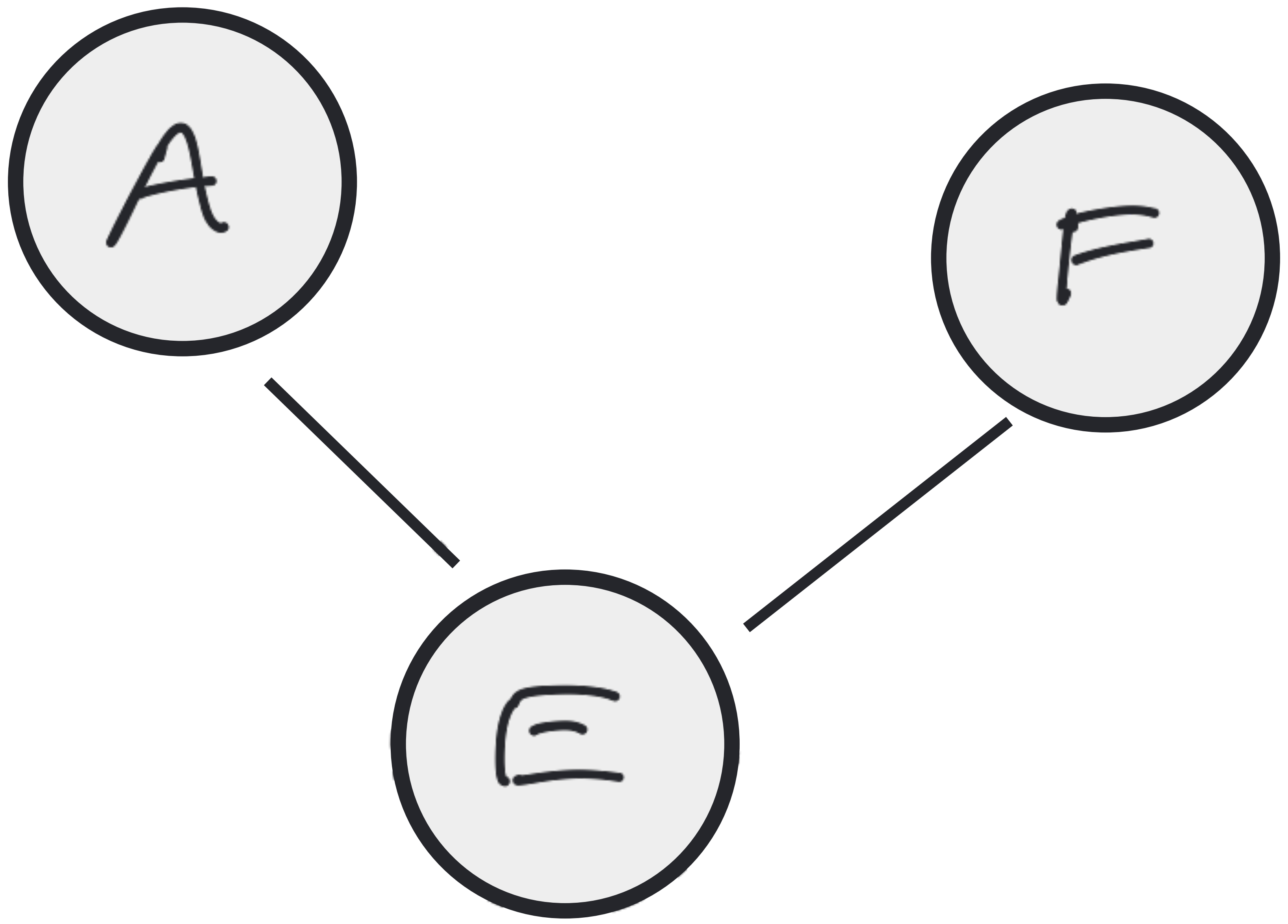
F:= the event that Héctor has the flue

E:= the event that Héctor is in the Emergency Room

$P(A) < P(A|E)$ → $P(A|B,E)$ or $P(B|A,E)$?
$P(B) < P(B|E)$ →

BAYES NETS: Explaining away

- Consider two different ailments: the flu and a broken arm.



A := the event that Héctor has a broken arm

F := the event that Héctor has the flu

E := the event that Héctor is in the Emergency Room

$$P(A \mid B, E) < P(A \mid E)$$

$$P(B \mid A, E) < P(B \mid E)$$