CSE 312

Foundations of Computing II

Lecture 7: Conditional Probabilities

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Stefano Tessaro

tessaro@cs.washington.edu

Reminder

Gradescope enroll code: M8YYEZ

Homework due tonight by 11:59pm.

Conditional Probabilities

Often we want to know how likely something is **conditioned on** something else having happened.

Example. *If we flip two fair coins, what is the probability that both outcomes are identical conditioned on the fact that at least one of them is heads?*

$\Omega = \{\mathrm{TT}, \mathrm{TH}, \mathrm{HT}, \mathrm{HH}\}$ $\forall \omega \in \Omega\colon\ \mathbb{P}(\omega) = \frac{1}{4}$

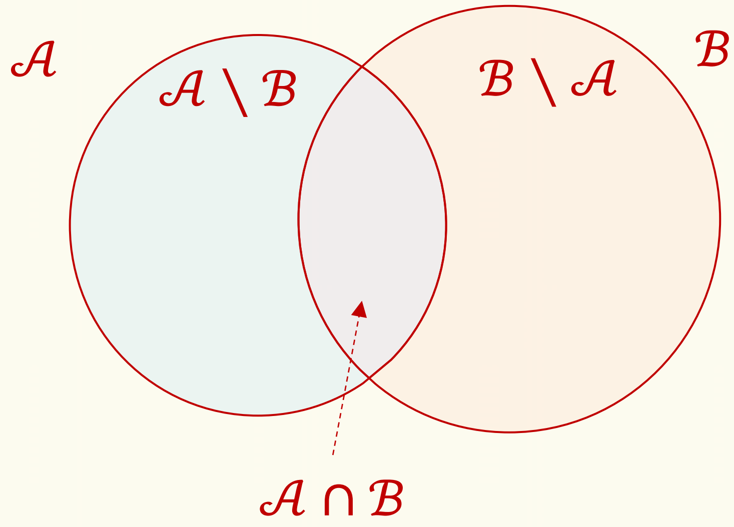
"we get heads at least once" $\mathcal{A} = \{\mathrm{TH}, \mathrm{HT}, \mathrm{HH}\}$

"same outcome" $\mathcal{B} = \{\mathrm{TT}, \mathrm{HH}\}$

If we <u>know</u> $\mathcal{A}$ happened: (1) only three outcomes are possible, and (2) only one of them leads to $\mathcal{B}$.

We expect: $\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{1}{3}$ [Verbalized: Probability of $\mathcal{B}$ conditioned on $\mathcal{A}$.]

Conditional Probability – Formal Definition



Definition. The **conditional probability** of $\mathcal{B}$ given $\mathcal{A}$ is

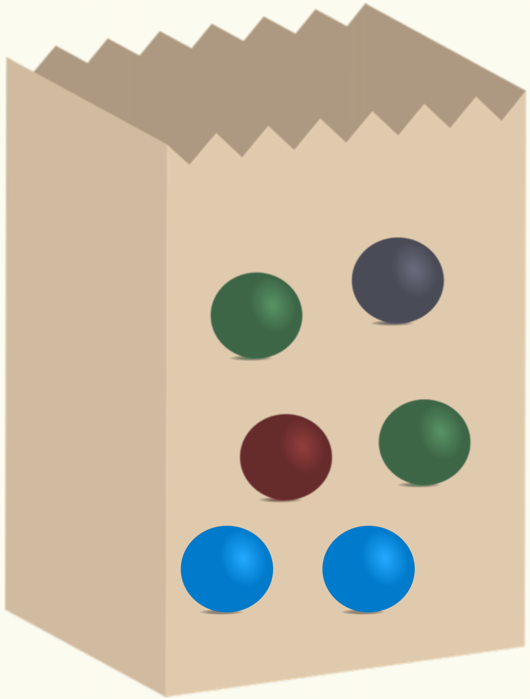
$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})}.$$

Note: This is only defined if $\mathbb{P}(\mathcal{A}) \neq 0$.

If $\mathbb{P}(\mathcal{A}) = 0$, then $\mathbb{P}(\mathcal{B}|\mathcal{A})$ is undefined.

Example – Non-uniform Case

Pick a random ball



$$\mathbb{P}(\text{green}) = \frac{1}{3}$$

$$\mathbb{P}(\text{red}) = \frac{1}{6} \qquad \mathbb{P}(\text{blue}) = \frac{1}{3} \qquad \mathbb{P}(\text{black}) = \frac{1}{6}$$

“we do not get black” $\mathcal{A} = \{\text{red}, \text{blue}, \text{green}\}$

“we get blue” $\mathcal{B} = \{\text{blue}\}$

$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\{\text{blue}\} \cap \{\text{red}, \text{blue}, \text{green}\})}{\mathbb{P}(\{\text{red}, \text{blue}, \text{green}\})} = \frac{\mathbb{P}(\{\text{blue}\})}{\mathbb{P}(\{\text{red}, \text{blue}, \text{green}\})} = \frac{1/3}{5/6} = \frac{2}{5}$$

The Effects of Conditioning

$$\mathbb{P}(\mathcal{B}) \quad \begin{matrix} < \\ = \\ > \end{matrix} \quad \mathbb{P}(\mathcal{B}|\mathcal{A})$$

"A-priori probability" / prior

All three are possible!

"A-posteriori probability" / posterior

Prior Examples – A-posteriori vs a-priori

“heads at least once” $\mathcal{A} = \{\mathrm{TH}, \mathrm{HT}, \mathrm{HH}\}$

“same outcome” $\mathcal{B} = \{\mathrm{TT}, \mathrm{HH}\}$

$$\mathbb{P}(\mathcal{B}) = \frac{1}{2} > \mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{1}{3}$$

“we do not get black” $\mathcal{A} = \{\mathrm{red}, \mathrm{blue}, \mathrm{green}\}$

“we get blue” $\mathcal{B} = \{\mathrm{blue}\}$

$$\mathbb{P}(\mathcal{B}) = \frac{1}{3} < \mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{2}{5}$$

Independence

Definition. Two events $\mathcal{A}$ and $\mathcal{B}$ are (statistically) **independent** if

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}).$$

Note: If $\mathcal{A}$, $\mathcal{B}$ independent, and $\mathbb{P}(\mathcal{A}) \neq 0$, then:

$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})} = \frac{\mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})}{\mathbb{P}(\mathcal{A})} = \mathbb{P}(B)$$

Reads as "The probability that $\mathcal{B}$ occurs is independent of $\mathcal{A}$."

Independence - Example

Assume we toss two fair coins

"first coin is heads" $\mathcal{A} = \{\mathrm{HH}, \mathrm{HT}\}$

"second coin is heads" $\mathcal{B} = \{\mathrm{HH}, \mathrm{TH}\}$

$$\mathbb{P}(\mathcal{A}) = 2 \times \frac{1}{4} = \frac{1}{2}$$

$$\mathbb{P}(\mathcal{B}) = 2 \times \frac{1}{4} = \frac{1}{2}$$

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\{HH\}) = \frac{1}{4} = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B})$$

Note here we have defined the probability space **assuming** independence, so quite unsurprising – but this makes it all precise.

Gambler's fallacy

Assume we toss **51** fair coins.
Assume we have seen **50** coins, and they are all "heads".
What are the odds the **51**$^{\text{st}}$ coin is also "heads"?

$\mathcal{A}$ = first 50 coins are heads

B = 51$^{\text{st}}$ coin is "heads"

51$^{\text{st}}$ coin is independent of outcomes of first 50 tosses!

$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})} = \frac{2^{-51}}{2^{-50}} = \frac{1}{2}$$

Gambler's fallacy = Feels like it's time for "tails"!?

Conditional Probability Define a Probability Space

The probability conditioned on $\mathcal{A}$ follows the same properties as (unconditional) probability.

Example. $\mathbb{P}(\mathcal{B}^c|\mathcal{A}) = 1 - \mathbb{P}(\mathcal{B}|\mathcal{A})$

Formally. $(\Omega, \mathbb{P})$ is a probability space + $\mathbb{P}(\mathcal{A}) > 0$

$\Rightarrow$ $(\Omega, \mathbb{P}(\cdot\,|\mathcal{A}))$ is a probability space

Recap

- $\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})}$.
- Independence: $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B})$.

Chain Rule

$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})} \quad \Rightarrow \quad \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{A} \cap \mathcal{B})$$

Theorem. (Chain Rule) For events $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$,

$$\mathbb{P}(\mathcal{A}_1 \cap \cdots \cap \mathcal{A}_n) = \mathbb{P}(\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_2|\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_3|\mathcal{A}_1 \cap \mathcal{A}_2) \cdots \mathbb{P}(\mathcal{A}_n|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \cdots \cap \mathcal{A}_{n-1})$$

(Proof: Apply above iteratively / formal proof requires induction)

Chain Rule – Applications

Often probability space $(\Omega, \mathbb{P})$ is given **implicitly**.

- Convenient: definition via a **sequential process.**
  - Use chain rule (implicitly) to define probability of outcomes in sample space.
- Allows for easy definition of experiments where $|\Omega| = \infty$

Sequential Process – Example

Setting: A fair die is thrown, and each time it is thrown, regardless of the history, it is equally likely to show any of the six numbers.

Rules: In each round

- If outcome = 1,2 → Alice wins
- If outcome = 3 → Bob wins
- Else, play another round

Sequential Process – Example

Rules: At each step:

- If outcome = 1,2 → Alice wins
- If outcome = 3 → Bob wins
- Else, play another round

$$\mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{A} \cap \mathcal{B})$$

Events:

- $\mathcal{A}_i$ = Alice wins in round i
- $\mathcal{N}_i$ = nobody wins in round i

$$\mathbb{P}(\mathcal{A}_1) = \frac{1}{3}$$

$$\mathbb{P}(\mathcal{A}_2) = \mathbb{P}(\mathcal{A}_2 \cap \mathcal{N}_1)$$

$$= \mathbb{P}(\mathcal{N}_1) \times \mathbb{P}(\mathcal{A}_2|\mathcal{N}_1)$$

$$= \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$$

[The event $\mathcal{A}_2$ implies $\mathcal{N}_1$, and this means that $\mathcal{A}_2 \cap \mathcal{N}_2 = \mathcal{A}_2$]

Sequential Process – Example

Rules: At each step:
- If outcome = 1,2 → Alice wins
- If outcome = 3 → Bob wins
- Else, play another round

Events:
- $\mathcal{A}_i$ = Alice wins in round i
- $\mathcal{N}_i$ = nobody wins in round i

$$\mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{A} \cap \mathcal{B})$$

$$\mathbb{P}(\mathcal{A}_i) = \mathbb{P}(\mathcal{A}_i \cap \mathcal{N}_1 \cap \mathcal{N}_2 \cap \cdots \cap \mathcal{N}_{i-1})$$

$$= \mathbb{P}(\mathcal{N}_1) \times \mathbb{P}(\mathcal{N}_2|\mathcal{N}_1) \times \mathbb{P}(\mathcal{N}_3|\mathcal{N}_1 \cap \mathcal{N}_2)$$

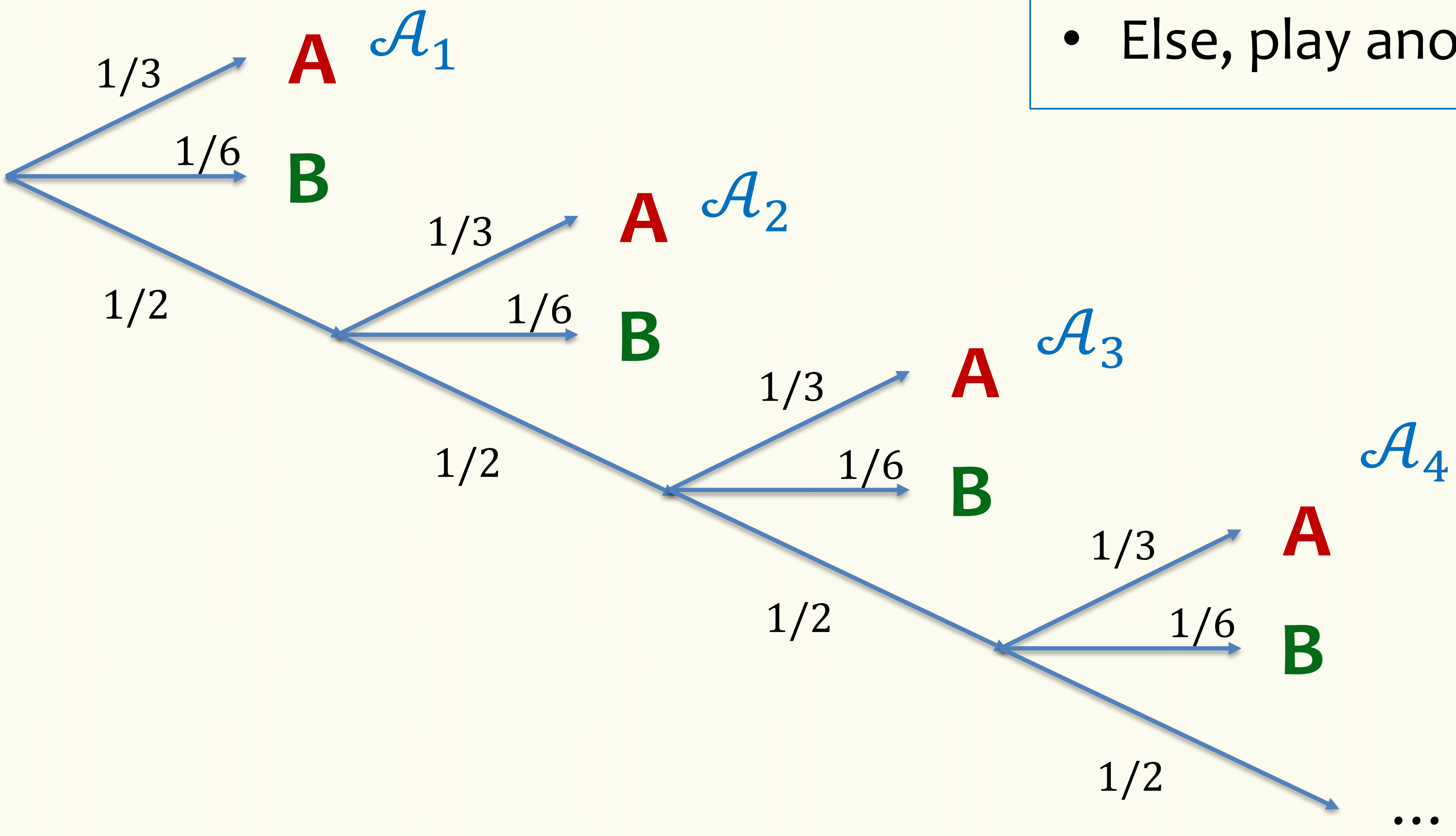
$$\cdots \times \mathbb{P}(\mathcal{N}_{i-1}|\mathcal{N}_1 \cap \mathcal{N}_2 \cap \cdots \cap \mathcal{N}_{i-2}) \times \mathbb{P}(\mathcal{A}_i|\mathcal{N}_1 \cap \mathcal{N}_2 \cap \cdots \cap \mathcal{N}_{i-1})$$

$$= \left(\frac{1}{2}\right)^{i-1} \times \frac{1}{3}$$

Sequential Process – Example

Rules: At each step:

- If outcome = 1,2 → **Alice** wins
- If outcome = 3 → **Bob** wins
- Else, play another round



Sequential Process – Crazy Math?

$\mathcal{A}_i$ = Alice wins in round i $\rightarrow \mathbb{P}(\mathcal{A}_i) = \left(\frac{1}{2}\right)^{i-1} \times \frac{1}{3}$

What is the probability that Alice wins?

$$\mathbb{P}(\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots) = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i \times \frac{1}{3} = \frac{1}{3} \times \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i = \frac{1}{3} \times 2 = \frac{2}{3}$$

Fact. If $|x| < 1$, then $\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$.

Sequential Process – Another Example

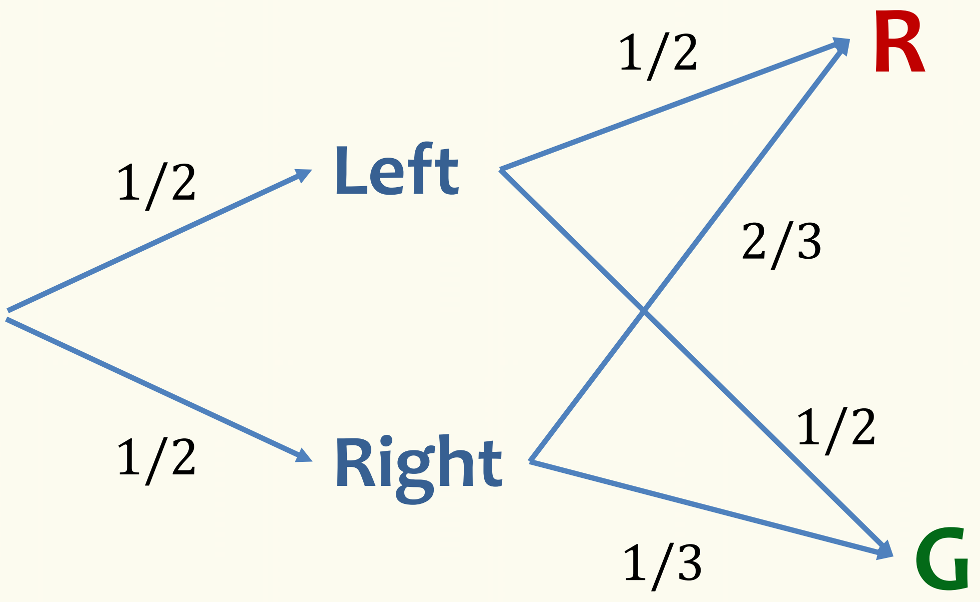
Alice has two pockets:

- **Left pocket:** Two red balls, two green balls
- **Right pocket:** One red ball, two green balls.

Alice picks a random ball from a random pocket.

[Both pockets equally likely, each ball equally likely.]

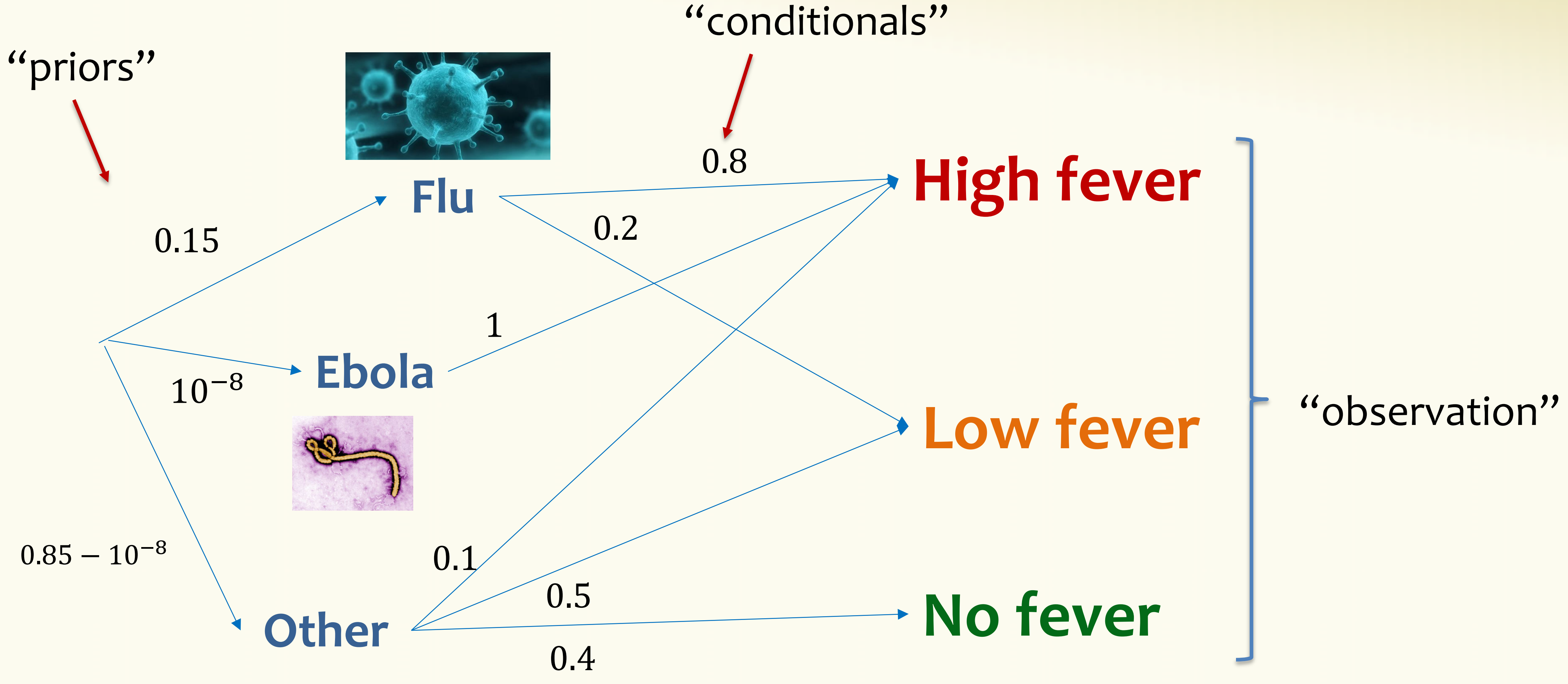
Sequential Process – Another Example



$$\mathbb{P}(\mathbf{R}) = \mathbb{P}(\mathbf{R} \cap \text{Left}) + \mathbb{P}(\mathbf{R} \cap \text{Right})$$ **(Law of total probability)**

$$= \mathbb{P}(\text{Left}) \times \mathbb{P}(\mathbf{R} | \text{Left}) + \mathbb{P}(\text{Right}) \times \mathbb{P}(\mathbf{R} | \text{Right})$$

$$= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{2}{3} = \frac{1}{4} + \frac{1}{3} = \frac{7}{12}$$



Assume we observe high fever, what is the probability that the subject has Ebola?

Posterior: $\mathbb{P}(\textbf{Ebola}|\textbf{High fever})$

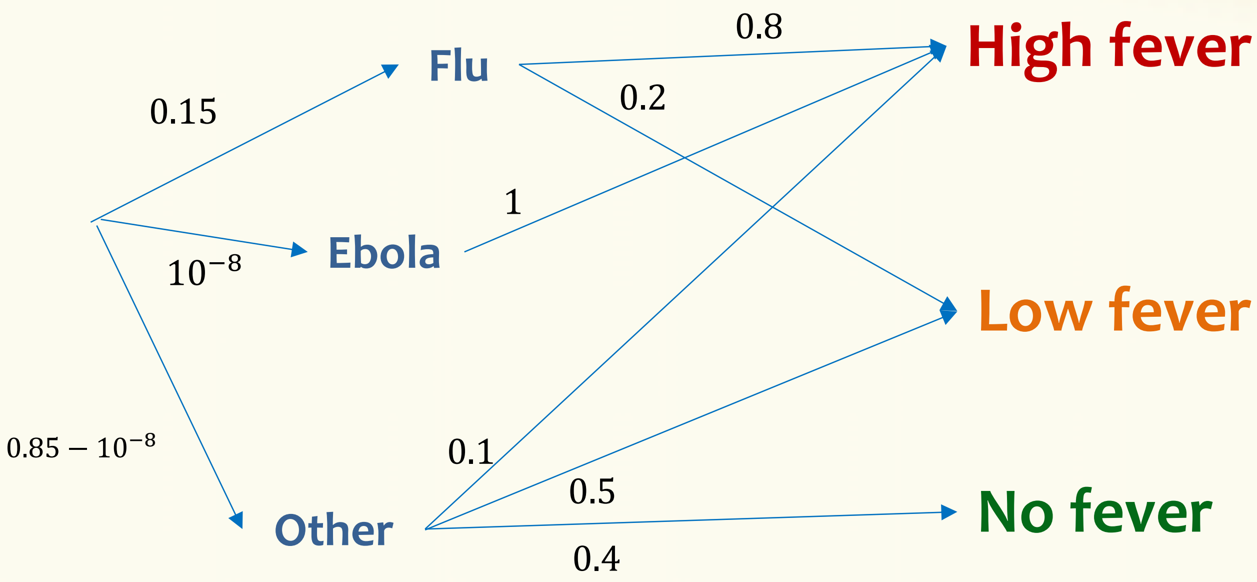
Bayes Rule

Theorem. (Bayes Rule) For events $\mathcal{A}$ and $\mathcal{B}$, where $\mathbb{P}(\mathcal{A}), \mathbb{P}(\mathcal{B}) > 0$,

$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{B}) \cdot \mathbb{P}(\mathcal{A}|\mathcal{B})}{\mathbb{P}(\mathcal{A})}$$

Rev. Thomas Bayes [1701-1761]

Proof: $\mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{A} \cap \mathcal{B})$



$$\mathbb{P}(\text{Ebola}|\text{High fever}) = \frac{\mathbb{P}(\text{Ebola}) \cdot \mathbb{P}(\text{High fever}|\text{Ebola})}{\mathbb{P}(\text{High fever})}$$

$$= \frac{10^{-8} \cdot 1}{0.15\times 0.8 + 10^{-8}\times 1 + (0.85 - 10^{-8})\times 0.1} \approx 7.4\times 10^{-8}$$

$\mathbb{P}(\text{Flu}|\text{High fever}) \approx 0.89$

Most-likely a-posteriori outcome (MLA)

$\mathbb{P}(\text{Other}|\text{High fever}) \approx 0.11$

Bayes Rule – Example

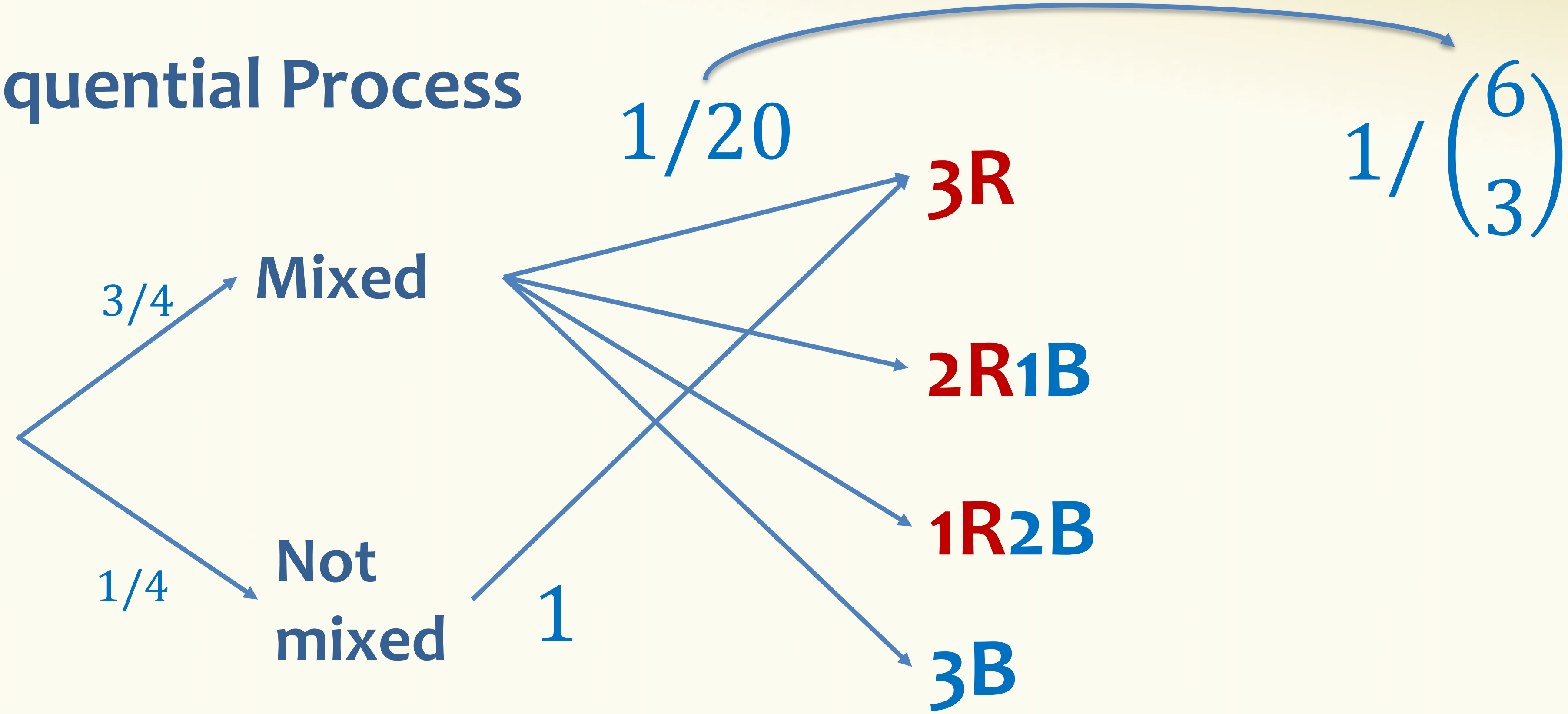
Setting: An urn contains 6 balls:

- 3 **red** and 3 **blue** balls w/ probability ¾
- 6 **red** balls w/ probability ¼

We draw three balls at random from the urn.

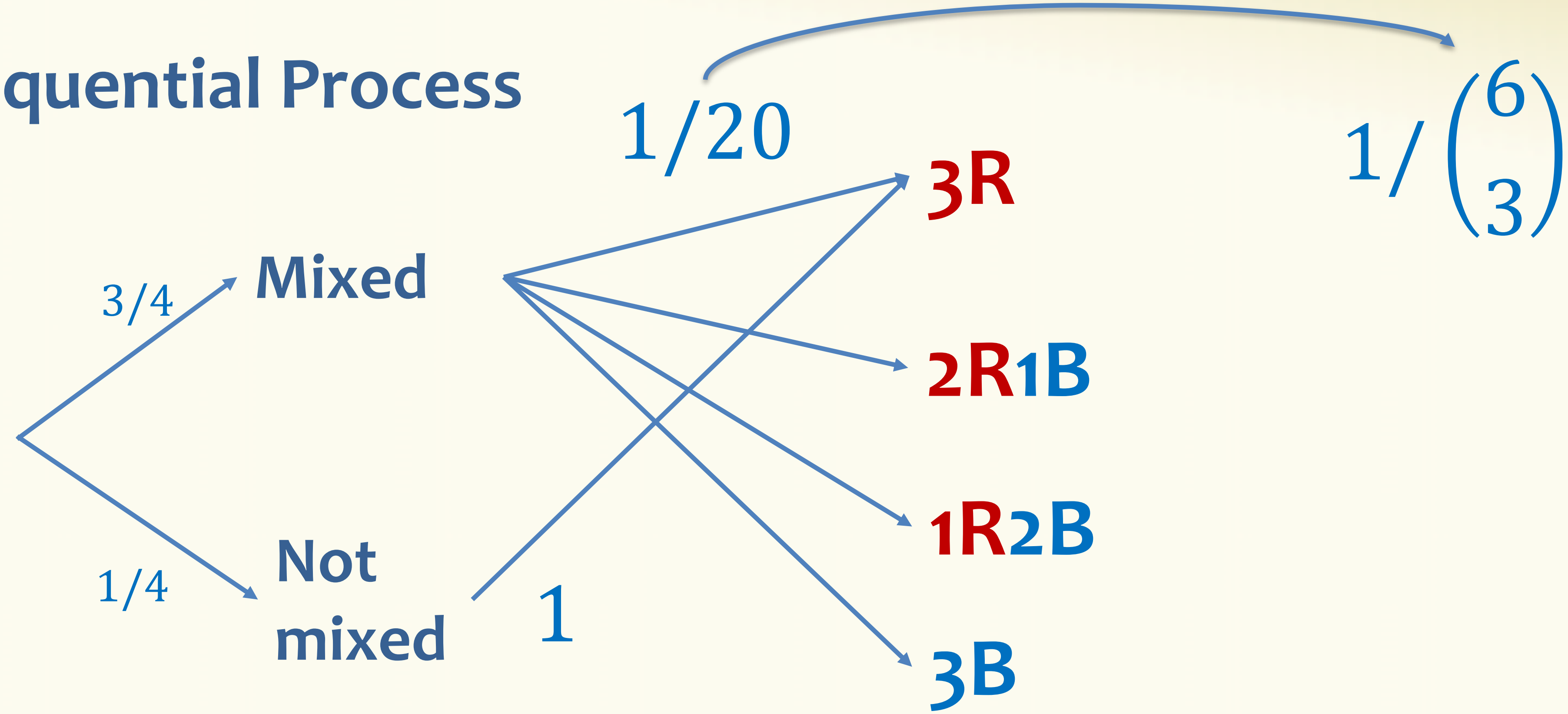
All three balls are red. What is the probability that the remaining (undrawn) balls are all blue?

Sequential Process



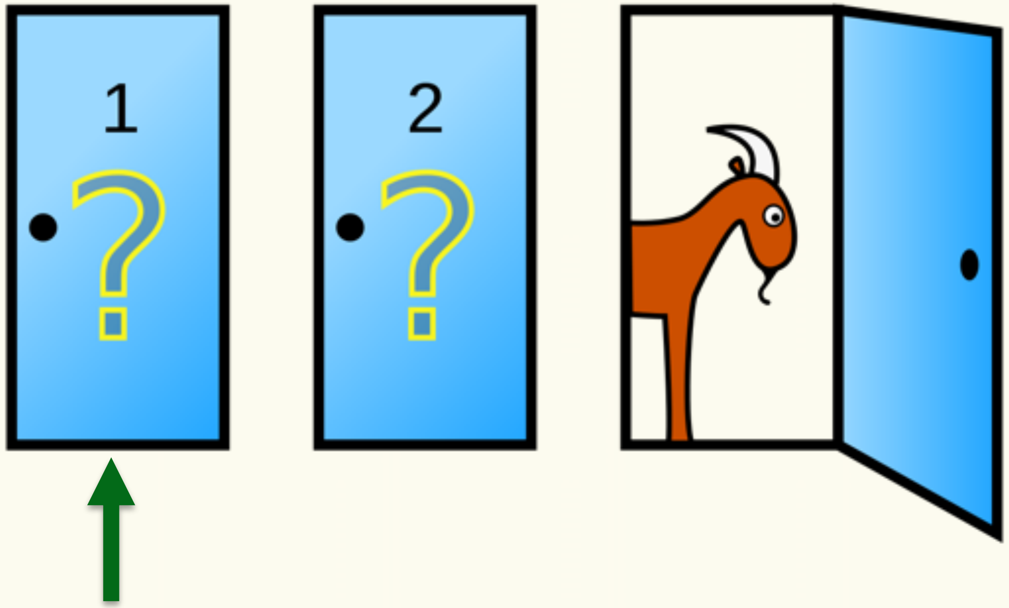
Wanted: $\mathbb{P}(\textbf{Mixed}|\textbf{3R})$

Sequential Process



$$\mathbb{P}(\text{Mixed}|\text{3R}) = \frac{\mathbb{P}(\text{Mixed})\mathbb{P}(\text{3R}|\text{Mixed})}{\mathbb{P}(\text{3R})} = \frac{\frac{3}{4}\times\frac{1}{20}}{\frac{3}{4}\times\frac{1}{20}+\frac{1}{4}\times 1} \approx 0.13$$

The Monty Hall Problem



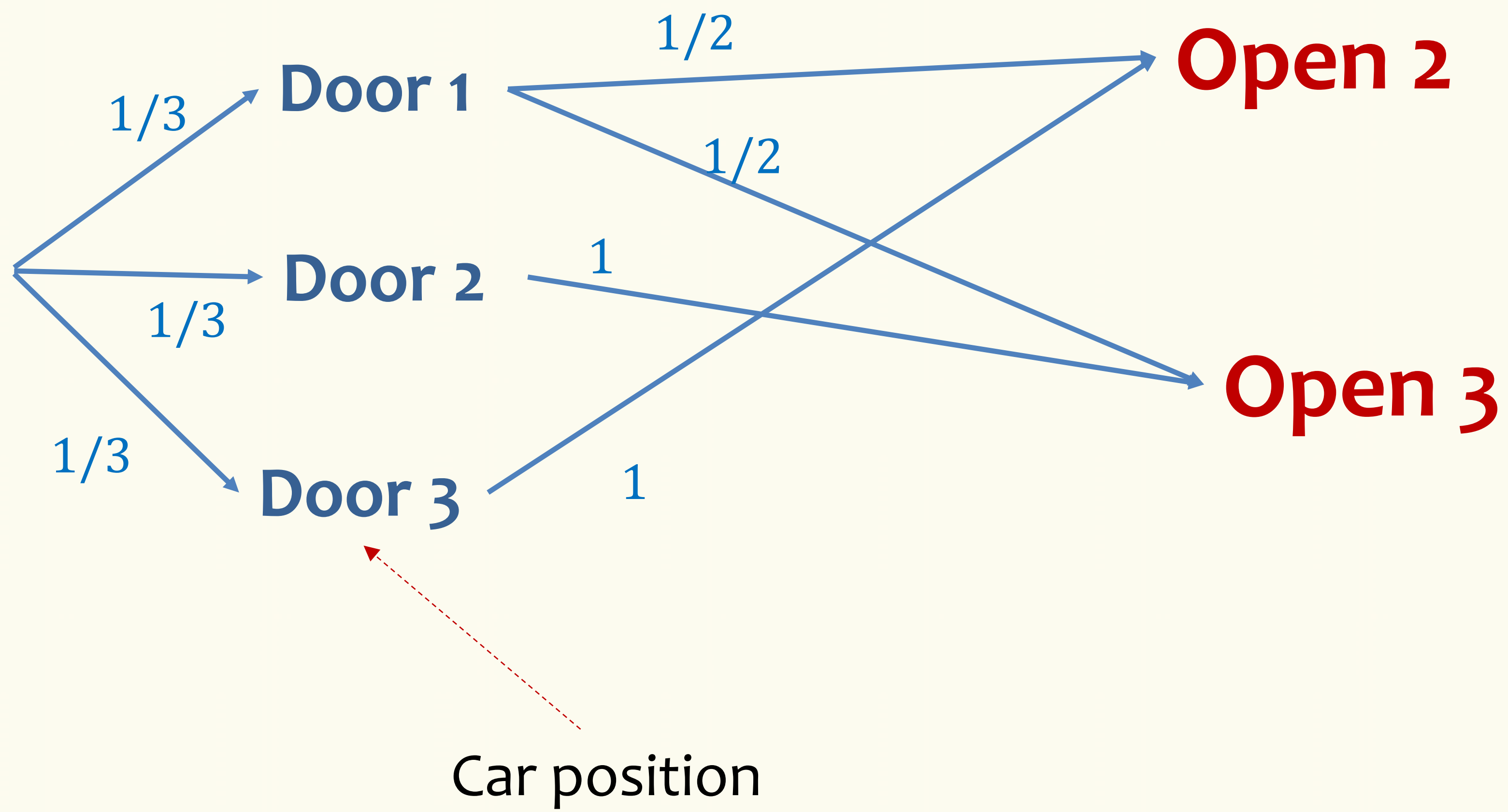
Your choice

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

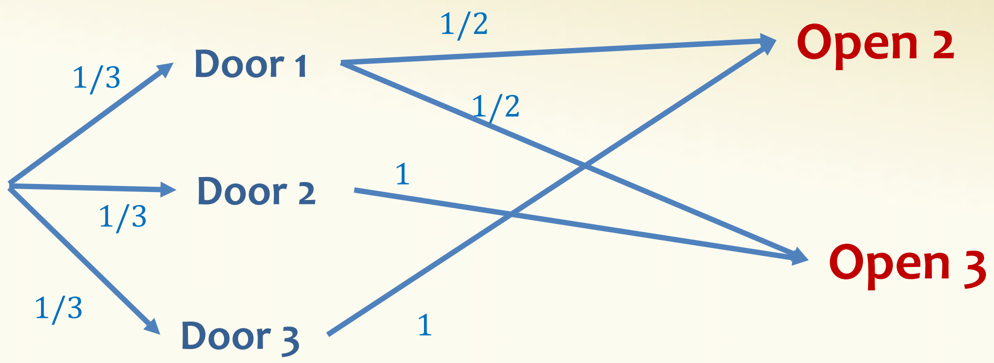
What would you do?

Monty Hall

Say you picked (without loss of generality) **Door 1**



Monty Hall



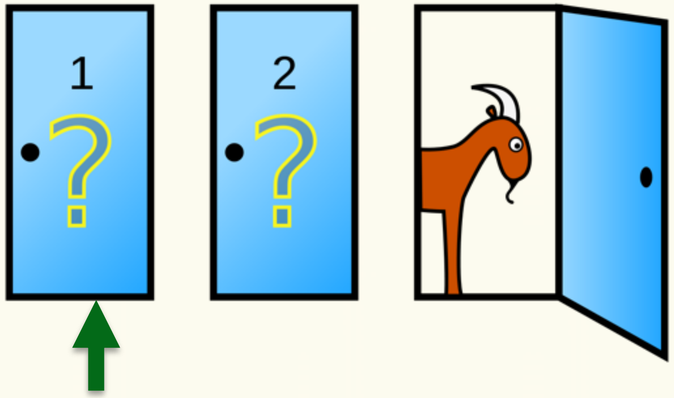
$$\mathbb{P}(\text{Door 1}|\text{Open 3}) = \frac{\mathbb{P}(\text{Door 1})\mathbb{P}(\text{Open 3}|\text{Door 1})}{\mathbb{P}(\text{Open 3})}$$

$$= \frac{\frac{1}{3}\times\frac{1}{2}}{\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times 1} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

$$\mathbb{P}(\text{Door 2}|\text{Open 3}) = 1 - \mathbb{P}(\text{Door 1}|\text{Open 3}) = 2/3$$

Monty Hall

Bottom line: Always swap!



Your choice