LECTURE 2: Conditioning and Bayes' rule

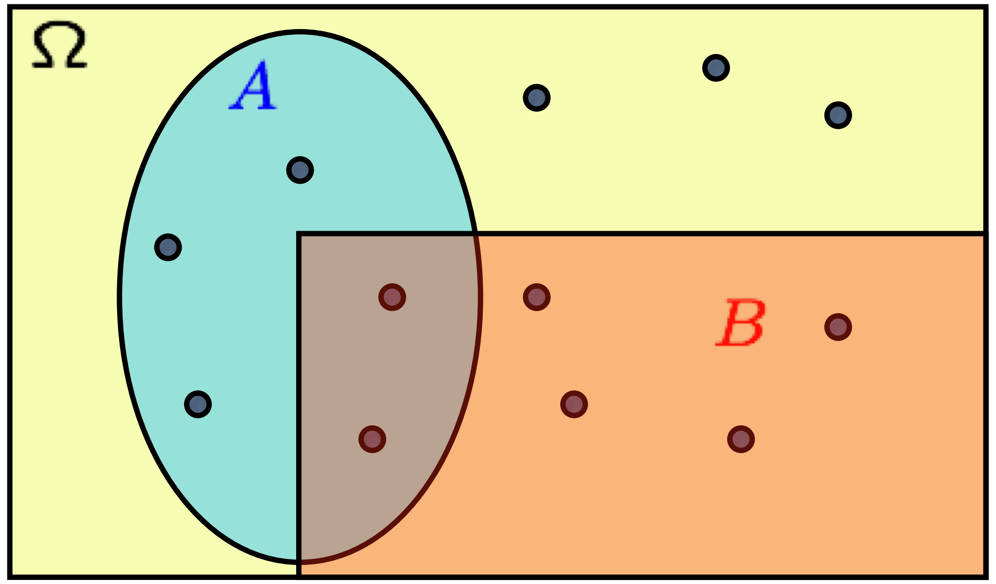
- Conditional probability

- Three **important** tools:
  - Multiplication rule
  - Total probability theorem
  - Bayes' rule ($\longrightarrow$ inference)

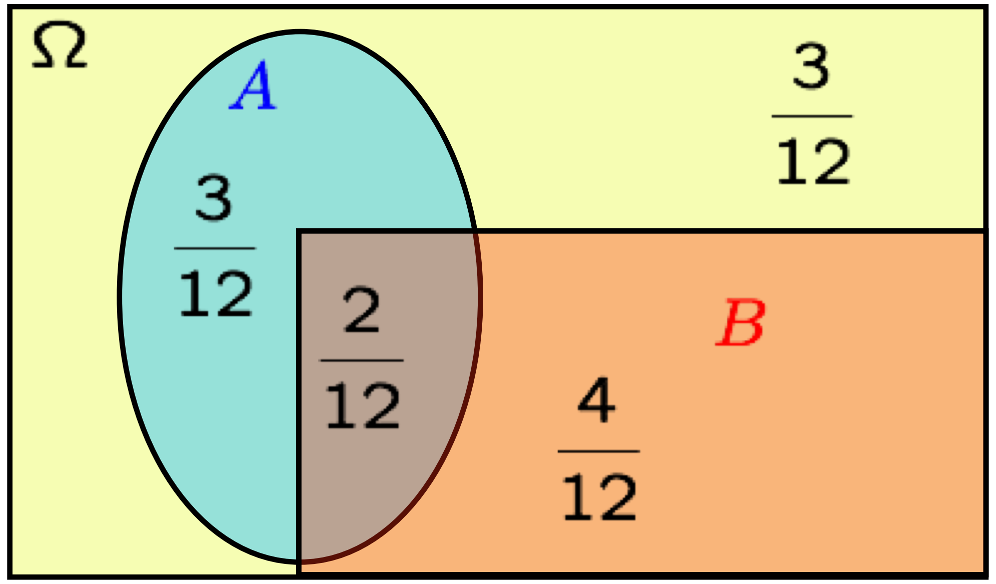
The idea of conditioning

Use new information to revise a model

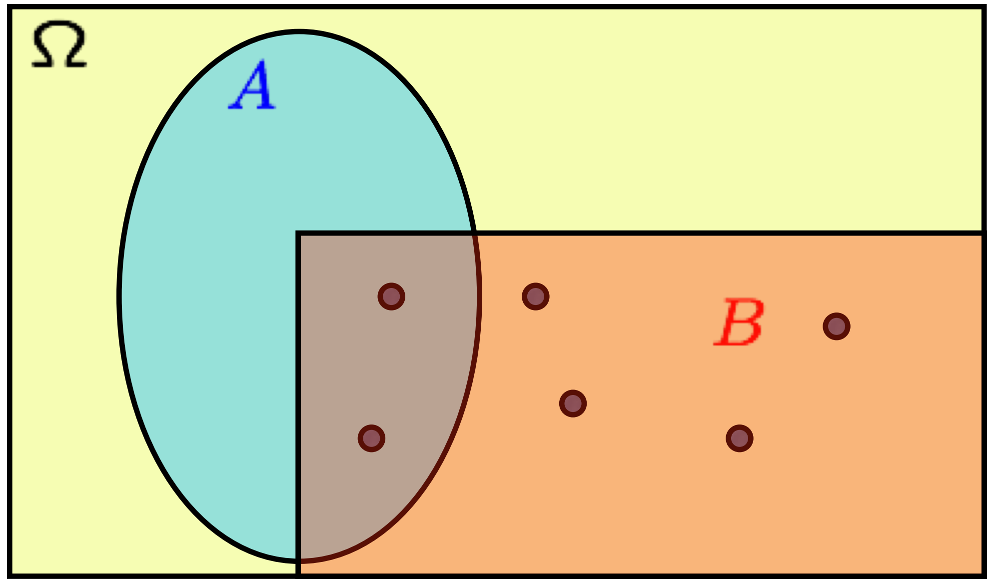
Assume 12 equally likely outcomes



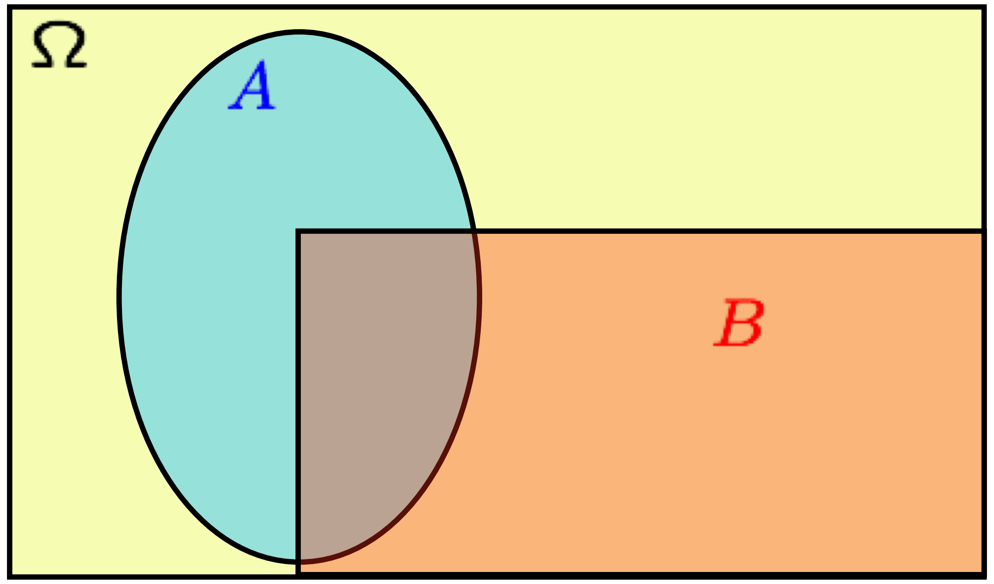
$$\mathbf{P}(A) = \frac{5}{12} \qquad \mathbf{P}(B) = \frac{6}{12}$$



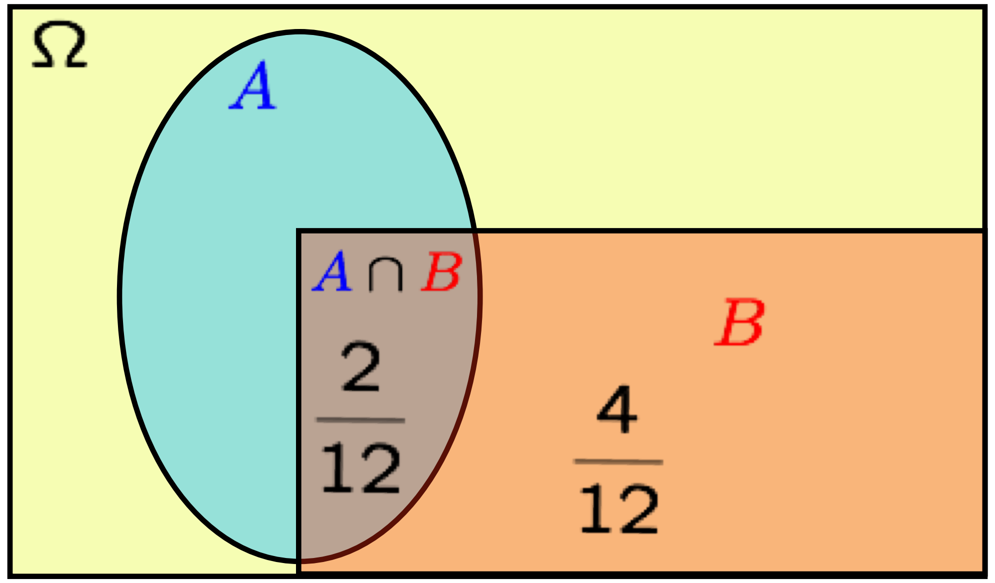
If told B occurred:



$$\mathbf{P}(A \mid B) = \qquad \mathbf{P}(B \mid B) =$$



Definition of conditional probability

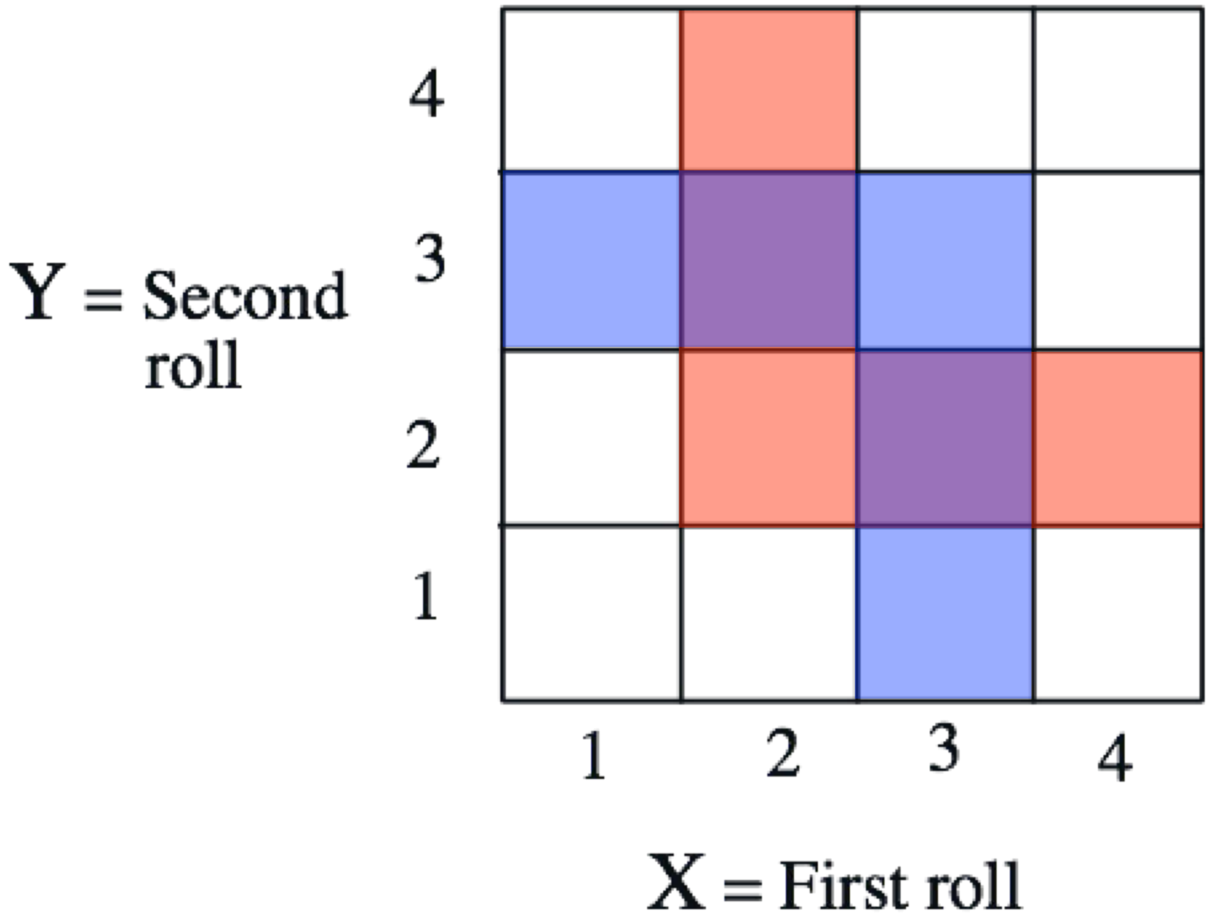


- $\mathbf{P}(A \mid B) =$ "probability of A, given that B occurred"

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

defined only when $\mathbf{P}(B) > 0$

Example: two rolls of a 4-sided die



- Let B be the event: $\min(X, Y) = 2$

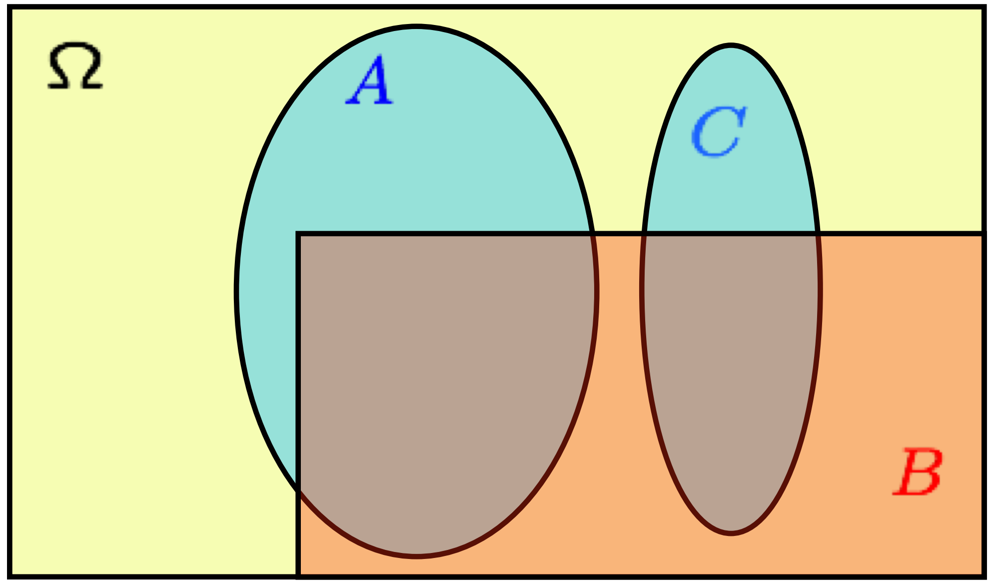
  Let $M = \max(X, Y)$

$\mathbf{P}(M = 1 \mid B) =$

$\mathbf{P}(M = 3 \mid B) =$

Conditional probabilities share properties of ordinary probabilities

$\mathbf{P}(A \mid B) \geq 0$ assuming $\mathbf{P}(B) > 0$



$\mathbf{P}(\Omega \mid B) =$

$\mathbf{P}(B \mid B) =$

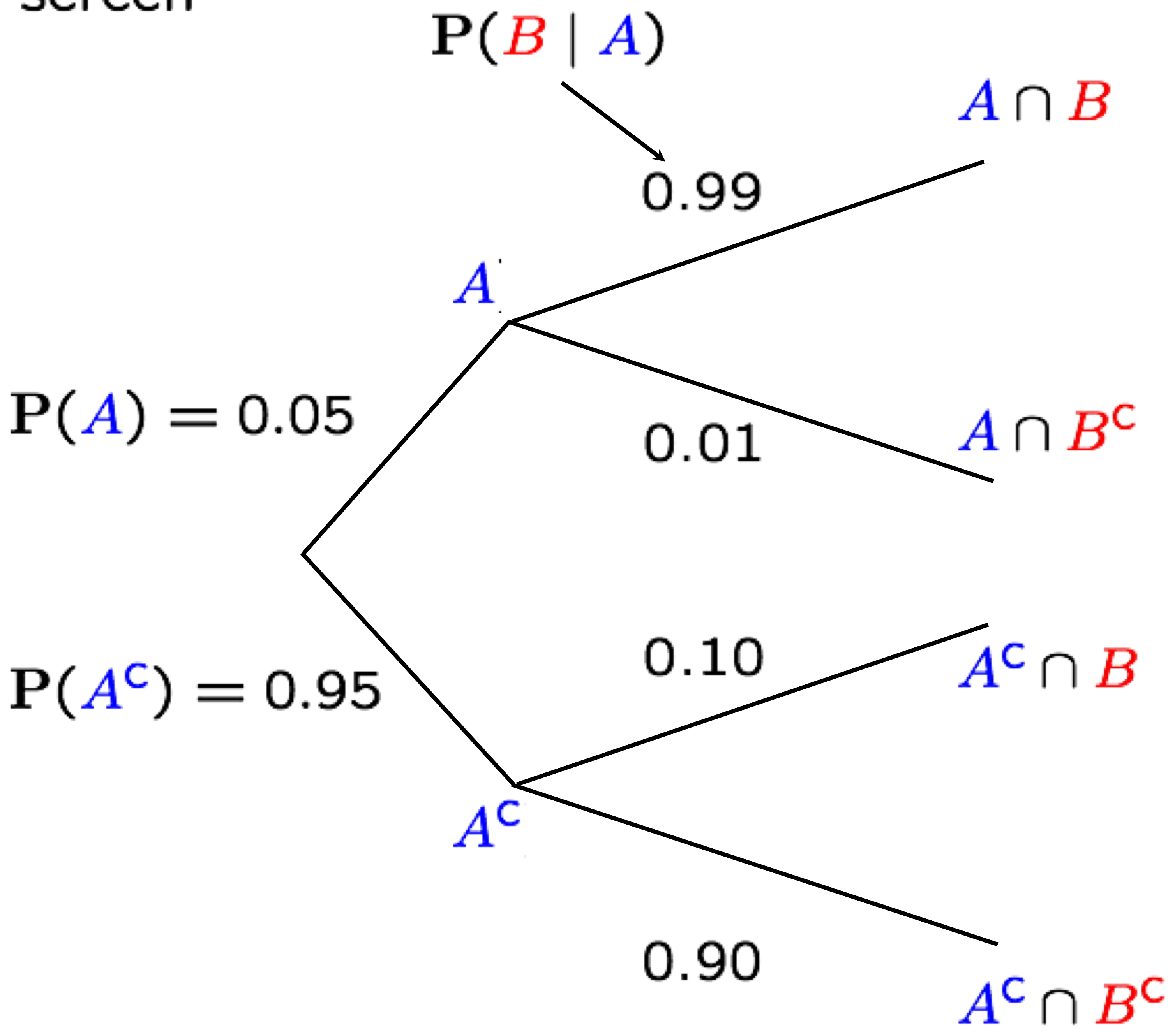
If $A \cap C = \varnothing$, then $\mathbf{P}(A \cup C \mid B) = \mathbf{P}(A \mid B) + \mathbf{P}(C \mid B)$

Models based on conditional probabilities

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

Event A: Airplane is flying above

Event B: Something registers on radar screen



- $P(A \cap B) =$

- $P(B) =$

- $P(A \mid B) =$

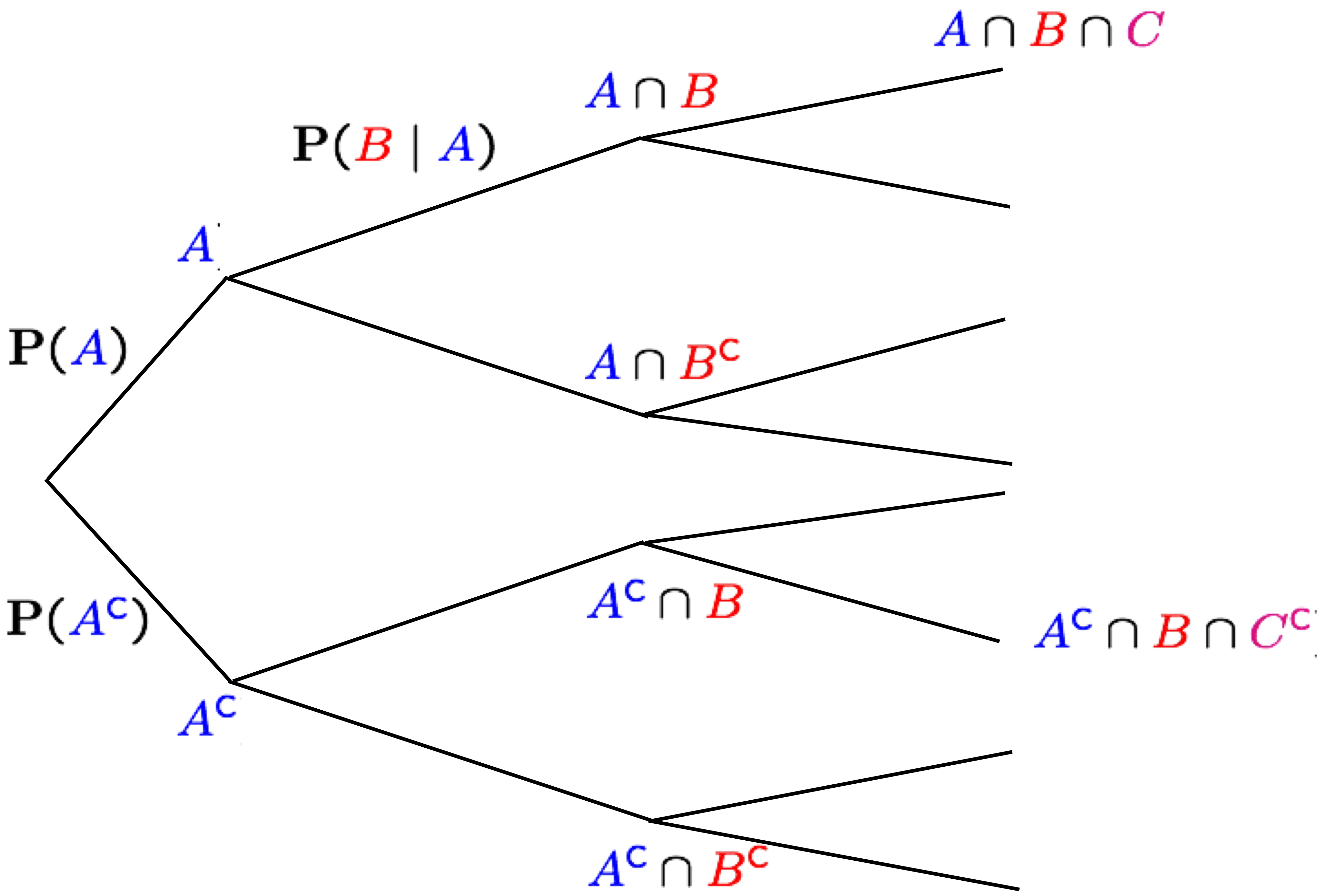
The multiplication rule

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

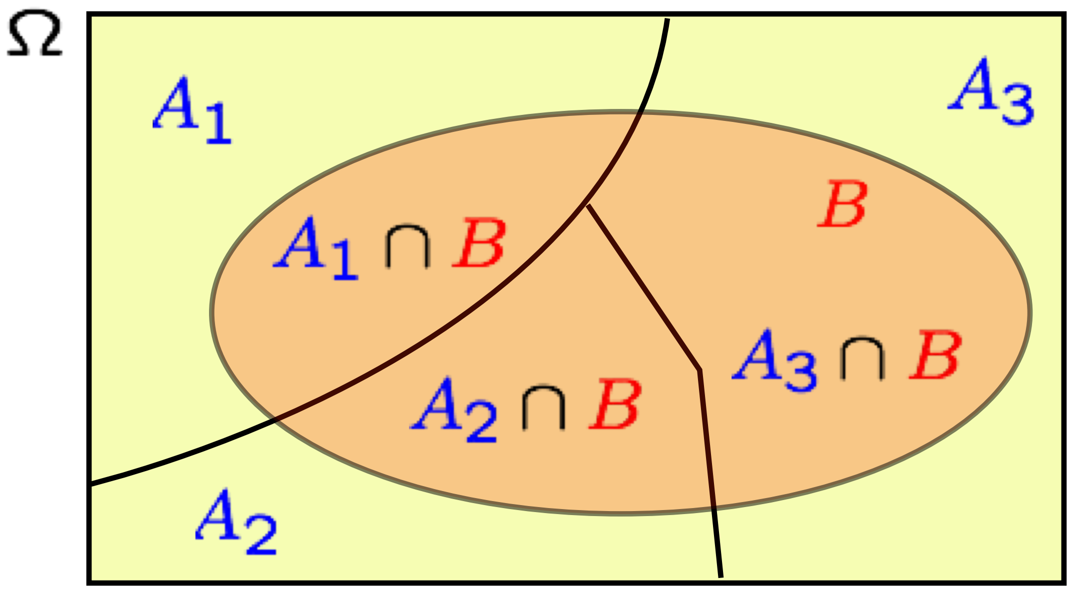
$$\mathbf{P}(A \cap B) = \mathbf{P}(B)\,\mathbf{P}(A \mid B)$$

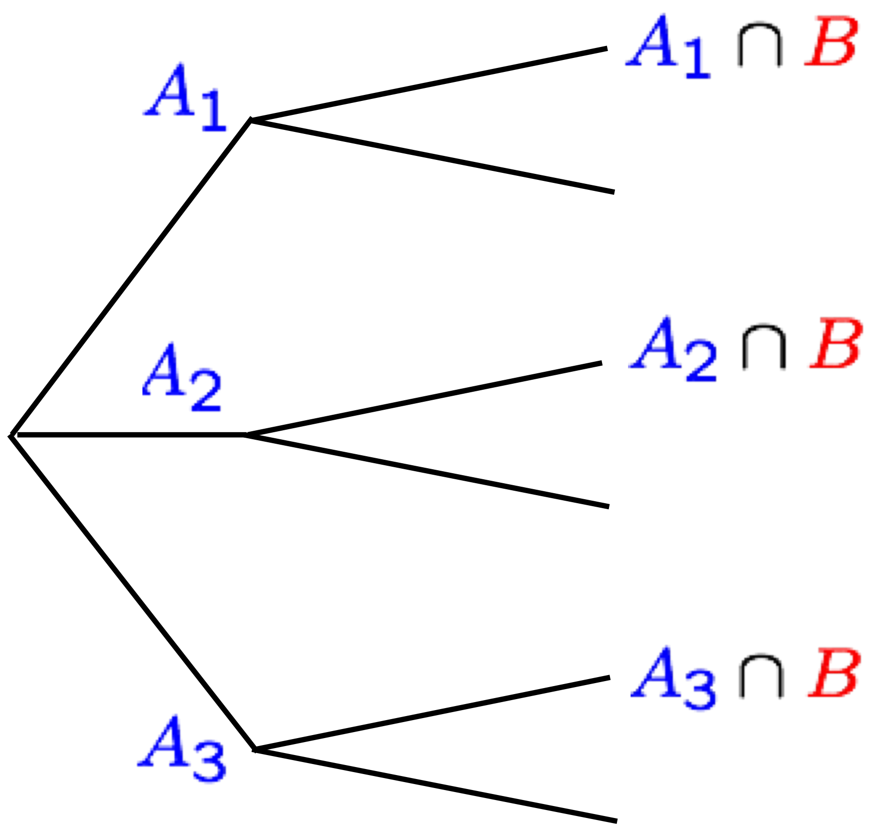
$$= \mathbf{P}(A)\,\mathbf{P}(B \mid A)$$

$$\mathbf{P}(A^c \cap B \cap C^c) =$$



Total probability theorem



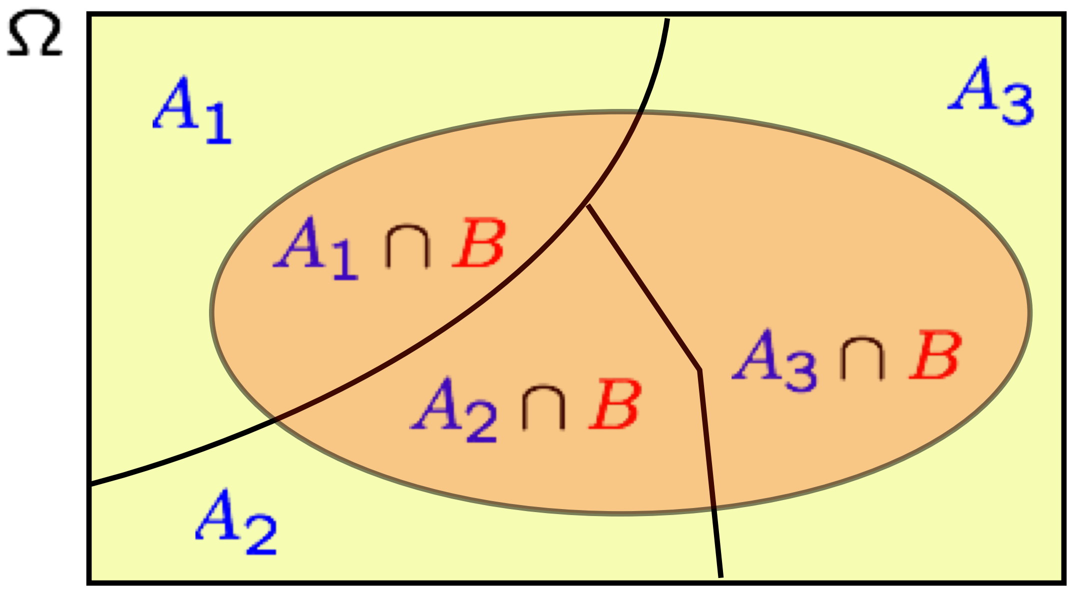


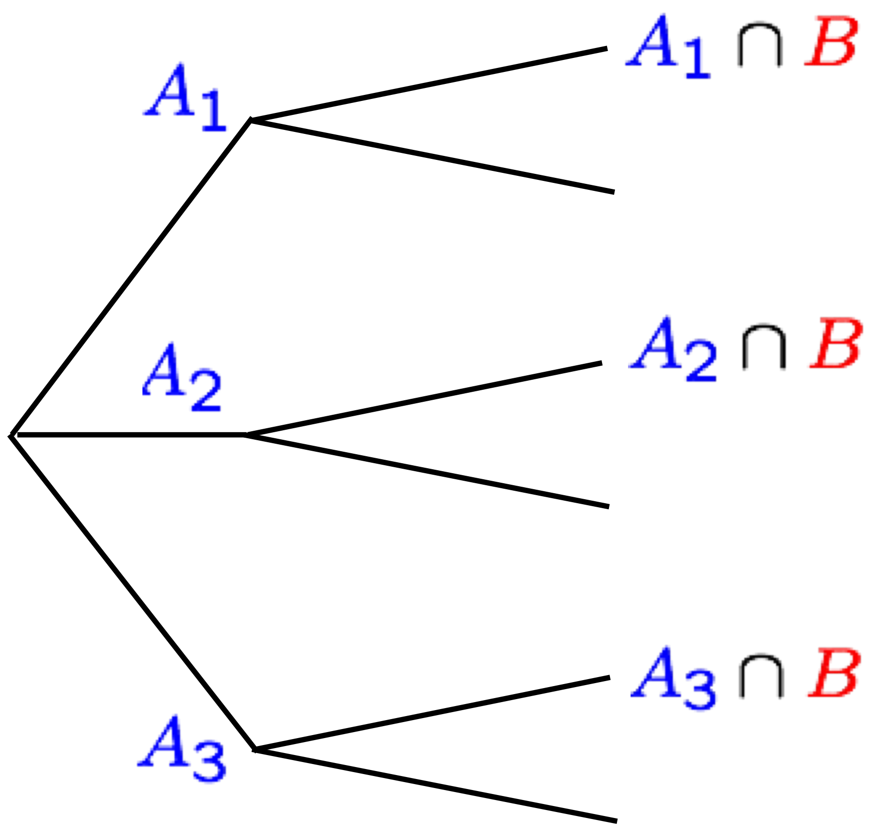
- Partition of sample space into A_1, A_2, A_3
- Have $\mathbf{P}(A_i)$, for every i
- Have $\mathbf{P}(B \mid A_i)$, for every i

$\mathbf{P}(B) =$

$$\mathbf{P}(B) = \sum_i \mathbf{P}(A_i)\,\mathbf{P}(B \mid A_i)$$

Bayes' rule





- Partition of sample space into A_1, A_2, A_3
- Have $\mathbf{P}(A_i)$, for every i — initial "beliefs"
- Have $\mathbf{P}(B \mid A_i)$, for every i

revised "beliefs," given that B occurred:

$$\mathbf{P}(A_i \mid B) =$$

$$\mathbf{P}(A_i \mid B) = \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\sum_j \mathbf{P}(A_j)\mathbf{P}(B \mid A_j)}$$

Bayes' rule and inference

- Thomas Bayes, presbyterian minister (c. 1701-1761)
- "Bayes' theorem," published posthumously
- systematic approach for incorporating new evidence

- Bayesian inference
  - initial beliefs $\mathbf{P}(A_i)$ on possible causes of an observed event B
  - model of the world under each A_i: $\mathbf{P}(B \mid A_i)$

$$A_i \xrightarrow[\mathbf{P}(B \mid A_i)]{\text{model}} B$$

  - draw conclusions about causes

$$B \xrightarrow[\mathbf{P}(A_i \mid B)]{\text{inference}} A_i$$