Probability in Computing

LECTURE 8

CS 237

Reminders

- HW4 due Thursday

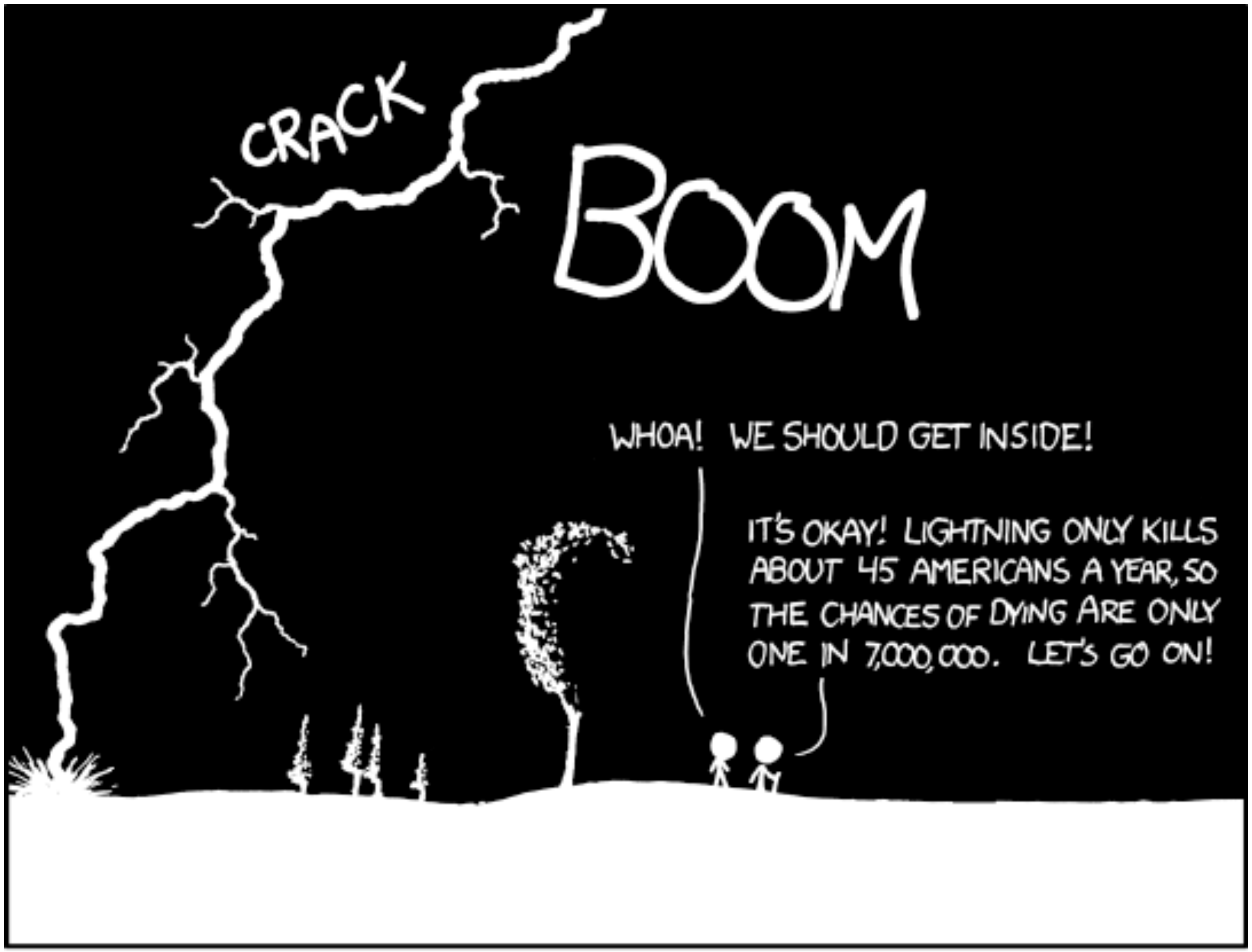
Last time

- Probability Mass Functions (PMFs)
- Probability Density Functions (PDFs)
- Cumulative Distribution Functions (CDFs)

Today

- Conditional Probability

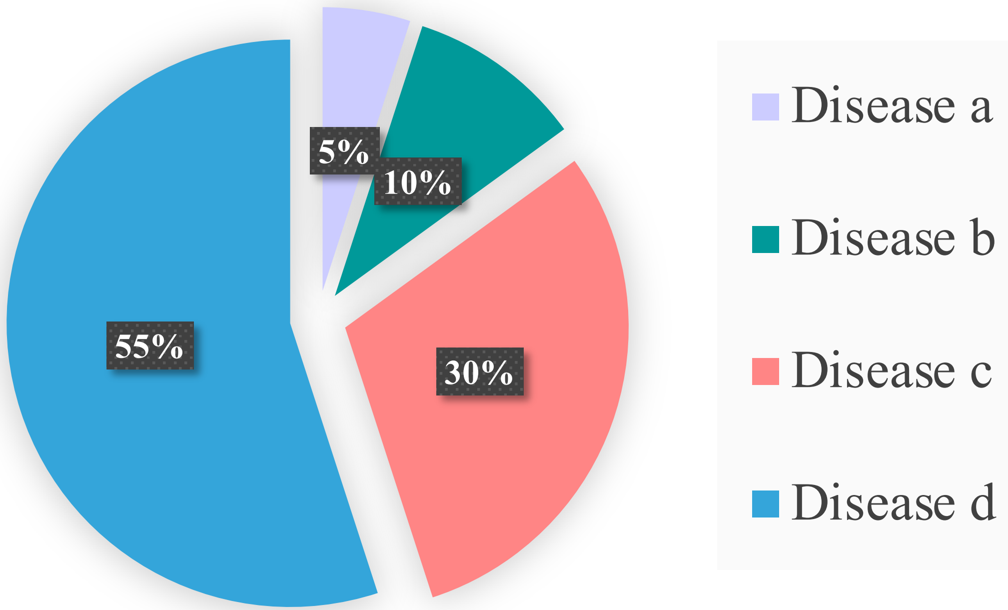
The annual death rate among people who know this statistic is one in six



https://xkcd.com/795/

Incorporating New Information

- A patient has some unknown disease
- Based on the symptoms, the doctor estimates that the patient has:



- Then a test reveals that the disease is neither b nor d
- Based on this information, what is the chance of having each disease?

Incorporating New Information

- A test eliminated diseases b and d

	Probability	
	Before the test	After the test
Disease a	**0.05**	?
Disease b	**0.1**	?
Disease c	**0.3**	?
Disease d	**0.55**	?

Incorporating New Information

- A test eliminated diseases b and d

	Probability: Before the test	Probability: After the test
Disease a	**0.05**	?
Disease b	**0.1**	→ 0
Disease c	**0.3**	?
Disease d	**0.55**	→ 0

Incorporating New Information

- A test eliminated diseases b and d

	Probability: Before the test		Probability: After the test
Disease a	$\mathbf{0.05}$	→	$\mathbf{0.05}x$
Disease b	$\mathbf{0.1}$	→	$\mathbf{0}$
Disease c	$\mathbf{0.3}$	→	$\mathbf{0.3}x$
Disease d	$\mathbf{0.55}$	→	$\mathbf{0}$

- In the absence of any other information, the new probabilities of a and b should be **proportional** to their original probabilities.

Incorporating New Information

- A test eliminated diseases b and d

	Probability: Before the test		Probability: After the test
Disease a	$\mathbf{0.05}$	→	$\frac{0.05}{0.35}$
Disease b	$\mathbf{0.1}$	→	$\mathbf{0}$
Disease c	$\mathbf{0.3}$	→	$\frac{0.3}{0.35}$
Disease d	$\mathbf{0.55}$	→	$\mathbf{0}$

- Probabilities should add up to 1: $0.05x + 0.3x = 1$

$$x = \frac{1}{0.35}$$

Incorporating New Information

- A test eliminated diseases b and d

	Probability: Before the test		Probability: After the test
Disease a	$\mathbf{0.05}$	→	$\frac{0.05}{0.35}$
Disease b	$\mathbf{0.1}$	→	$\mathbf{0}$
Disease c	$\mathbf{0.3}$	→	$\frac{0.3}{0.35}$
Disease d	$\mathbf{0.55}$	→	$\mathbf{0}$

- Note: $\Pr(\textit{"the disease is neither b nor d"}) = \Pr(\{a, c\}) = 0.35$

Conditional Probability

- The information provided by the test is that the following event has happened:

$$B = \textit{"the disease is neither b nor d"} = \{a, c\}$$

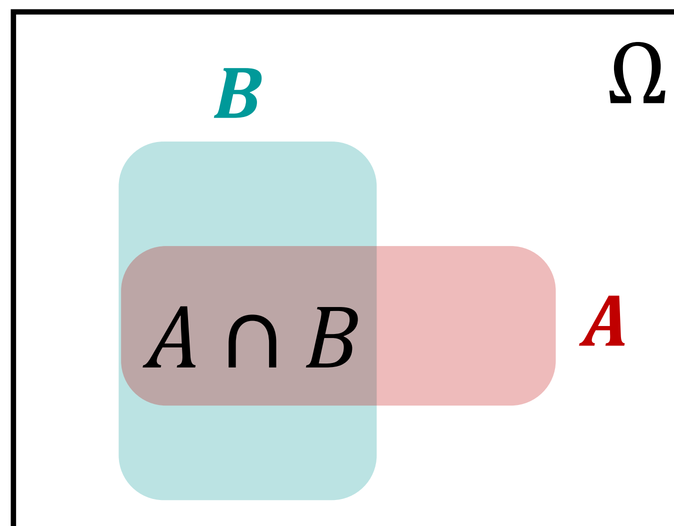
- We define a new probability function that assigns a probability $\Pr(A \mid B)$ to every event $A \subseteq \Omega$

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

Read: ``Probability of A given B''

If $A = \{\omega\}$ for an outcome ω, we also write: $Pr(\omega \mid B)$

If we specify $\Pr(\omega \mid B)$ for every outcome $\omega \in \Omega$, then $\Pr(A \mid B)$ is determined for all events A (as the sum of probabilities of all outcomes in A)



Conditional Probability

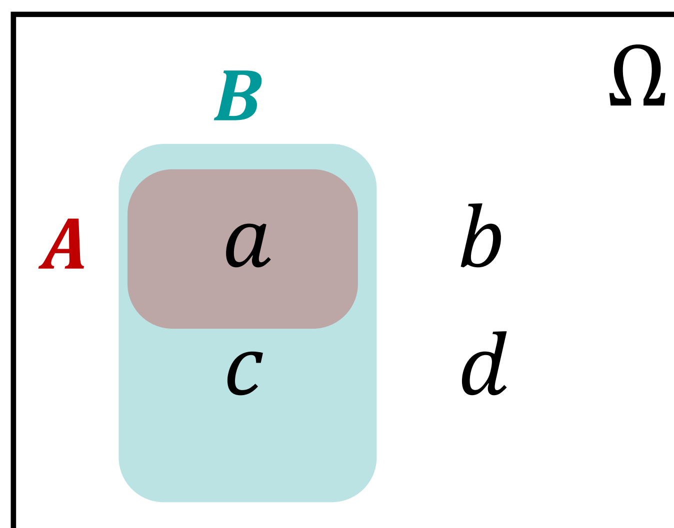
- The information provided by the test is that the following event has happened:

$$B = \text{"the disease is neither } b \text{ nor } d\text{"} = \{a, c\}$$

- We define a new probability function that assigns a probability $\Pr(A \mid B)$ to every event $A \subseteq \Omega$

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

$$\Omega = \{a, b, c, d\}$$



$$\Pr(a|B) = \frac{\Pr(a)}{\Pr(B)} = \frac{0.05}{0.35}$$

Conditional Probability

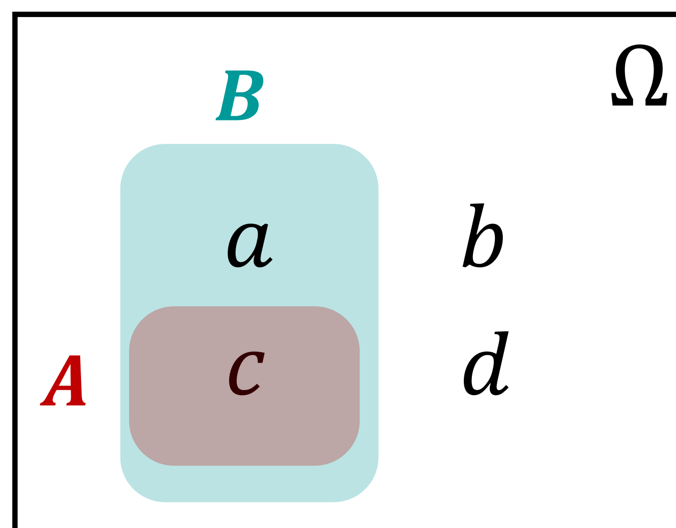
- The information provided by the test is that the following event has happened:

$$B = \text{"the disease is neither b nor d"} = \{a, c\}$$

- We define a new probability function that assigns a probability $\Pr(A \mid B)$ to every event $A \subseteq \Omega$

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

$$\Omega = \{a, b, c, d\}$$



$$\Pr(c|B) = \frac{\Pr(c)}{\Pr(B)} = \frac{0.3}{0.35}$$

Conditional Probability

- The information provided by the test is that the following event has happened:

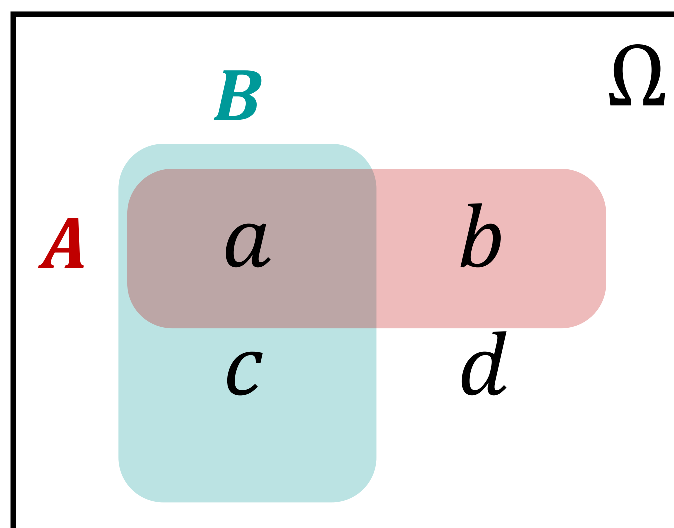
$$B = \text{"the disease is neither } b \text{ nor } d\text{"} = \{a, c\}$$

- We define a new probability function that assigns a probability $\Pr(A \mid B)$ to every event $A \subseteq \Omega$

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

$$\Omega = \{a, b, c, d\}$$

$$\Pr(\{a, b\} \mid B) = \frac{\Pr(\{a\})}{\Pr(B)} = \frac{0.05}{0.35}$$



Conditional Probability

- The information provided by the test is that the following event has happened:

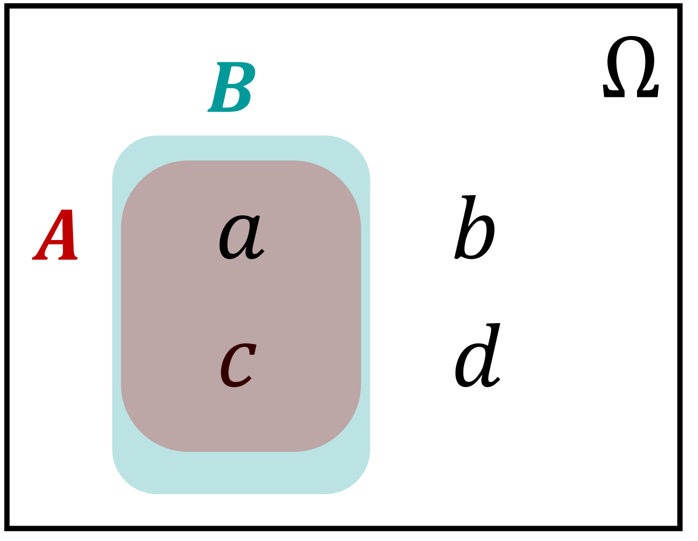
$$B = \text{"the disease is neither } b \text{ nor } d\text{"} = \{a, c\}$$

- We define a new probability function that assigns a probability $\Pr(A \mid B)$ to every event $A \subseteq \Omega$

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

$$\Pr(\{a,c\} \mid B) = \frac{\Pr(\{a,c\})}{\Pr(B)} = 1$$

$\Omega = \{a, b, c, d\}$



Conditional Probability

Definition: Conditional Probability

Let B be an event such that $\Pr(B) \neq 0$. For every event A, we define the conditional probability of event A given event B:

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

$\mathbf{Pr}(\boldsymbol{A} \mid \boldsymbol{B})$ is undefined if $\Pr(B) = 0$.

Top Hat question (Join Code: 033357)

- Experiment: roll two fair 6-sided dice
- A = "first die is 6"
- B = "sum of the two rolls is 9"

What is $\Pr(A \mid B)$?

A. $\frac{1}{36}$

B. $\frac{1}{6}$

C. $\frac{1}{4}$

D. None of the above

Top Hat question (Join Code: 033357)

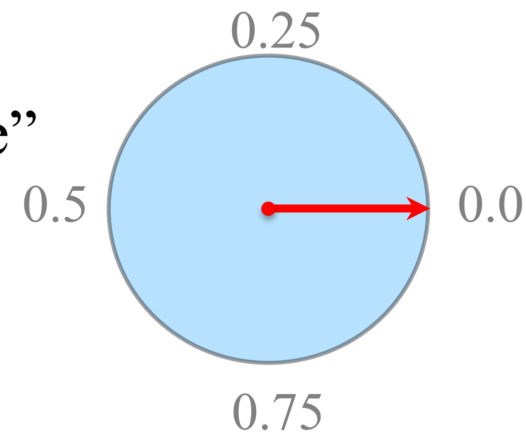
- Experiment: toss a fair coin three times
- A = “more H than T”
- B = “first toss is H”

What is $\Pr(A \mid B)$?

A. $\frac{3}{8}$
B. $\frac{1}{2}$
C. $\frac{3}{4}$
D. None of the above

Top Hat question (Join Code: 033357)

- Experiment: spin the dial of the spinner
- A = "dial stopped in the upper half of the circle"
- B = "outcome is in [1/3, 2/3]"

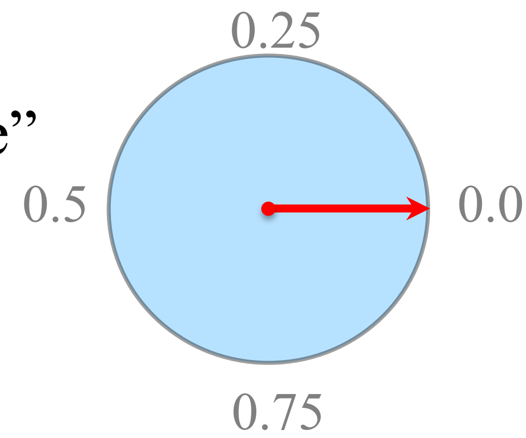


What is $\Pr(A \mid B)$?

A. $\frac{1}{3}$
B. $\frac{1}{2}$
C. $\frac{3}{4}$
D. None of the above

Top Hat question (Join Code: 033357)

- Experiment: spin the dial of the spinner
- A = “dial stopped in the upper half of the circle”
- C = “outcome is in [1/6, 1/3]”



What is $\Pr(A \mid C)$?

A. $\frac{1}{12}$
B. $\frac{1}{6}$
C. 1
D. None of the above

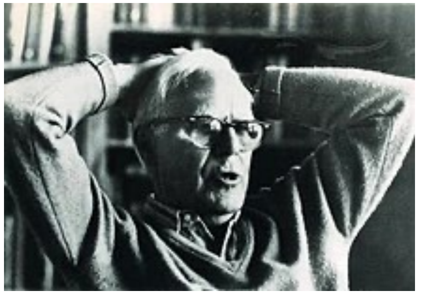
Top Hat question (Join Code: 033357)

Mr. Smith has two children
(boys and girls are equally likely)

- A = "both children are boys"
- B = "at least one child is a boy"
- C = "older child is a boy"

True or false: $\Pr(A \mid B) = \Pr(A \mid C)$?

A. True
B. False



[Martin Gardner]
[1914-2010]

CS 237 Axioms of Probability

Conditional probabilities satisfy the axioms of probability

- Non-negativity: $\Pr(A \mid E) \geq 0$ for all events $A \subseteq \Omega$
- Additivity: if A and B are *disjoint* events then
  $$\Pr(A \cup B \mid E) = \Pr(A \mid E) + \Pr(B \mid E)$$
- Normalization: $\Pr(\Omega \mid E) = 1$

- This applies only to conditional probabilities with the same conditioning event: $\Pr(\cdot \mid E)$

CS 237 Axioms of Probability: Derivation

Conditional probabilities satisfy the axioms of probability

– Non-negativity: $\Pr(A \mid E) \geq 0$ for all events $A \subseteq \Omega$

Proof:

$$\Pr(A \mid E) = \frac{\Pr(A \cap E)}{\Pr(E)}$$

by definition of conditional probability

– Normalization: $\Pr(\Omega \mid E) = 1$

Proof:

$$\Pr(\Omega \mid E) = \frac{\Pr(\Omega \cap E)}{\Pr(E)}$$

by definition of conditional probability

$$= \frac{\Pr(E)}{\Pr(E)} = 1$$

since $E \subseteq \Omega$

Axioms of Probability: Derivation

Conditional probabilities satisfy the axioms of probability

– Additivity: if A and B are *disjoint* events then

$$\Pr(A \cup B \mid E) = \Pr(A \mid E) + \Pr(B \mid E)$$

Proof:

$$\Pr(A \cup B \mid E) = \frac{\Pr((A \cup B) \cap E)}{\Pr(E)}$$

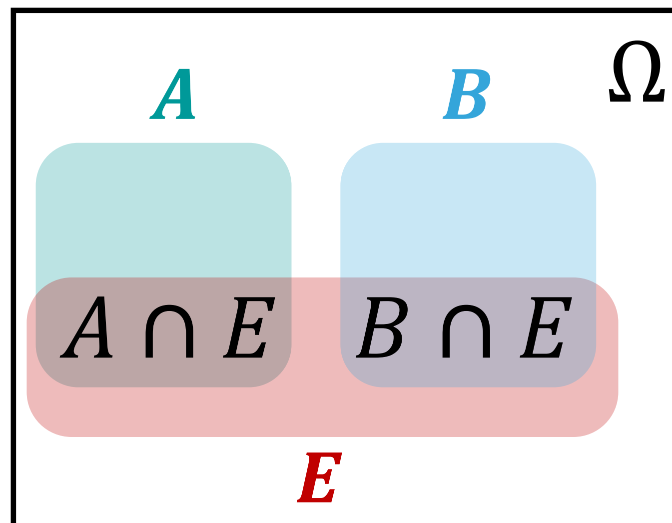
by definition of conditional probability

$$= \frac{\Pr((A \cap E) \cup (B \cap E))}{\Pr(E)}$$

by distributive law

$$= \frac{\Pr(A \cap E) + \Pr(B \cap E))}{\Pr(E)}$$

$$= \Pr(A \mid E) + \Pr(B \mid E)$$



Conditional Probability Rules

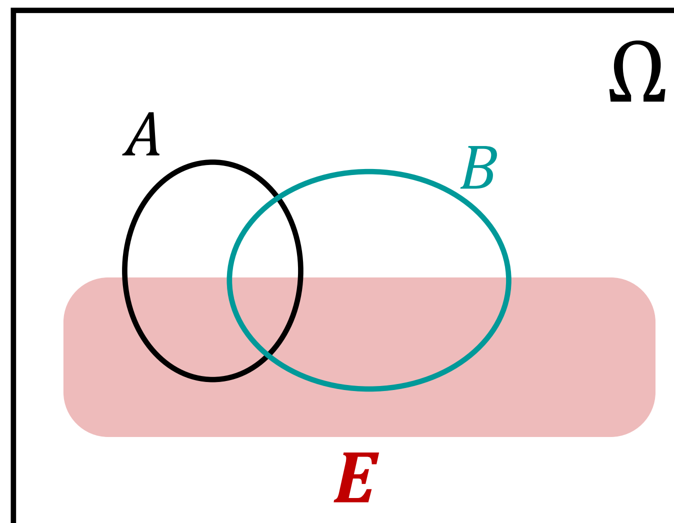
Since the conditional probability function is a valid probability function, all the probability rules remain valid

- Example: Inclusion-Exclusion Principle

Standard: $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$

Conditional: $\Pr(A \cup B \mid E)$

$$= \Pr(A \mid E) + \Pr(B \mid E) - \Pr(A \cap B \mid E)$$



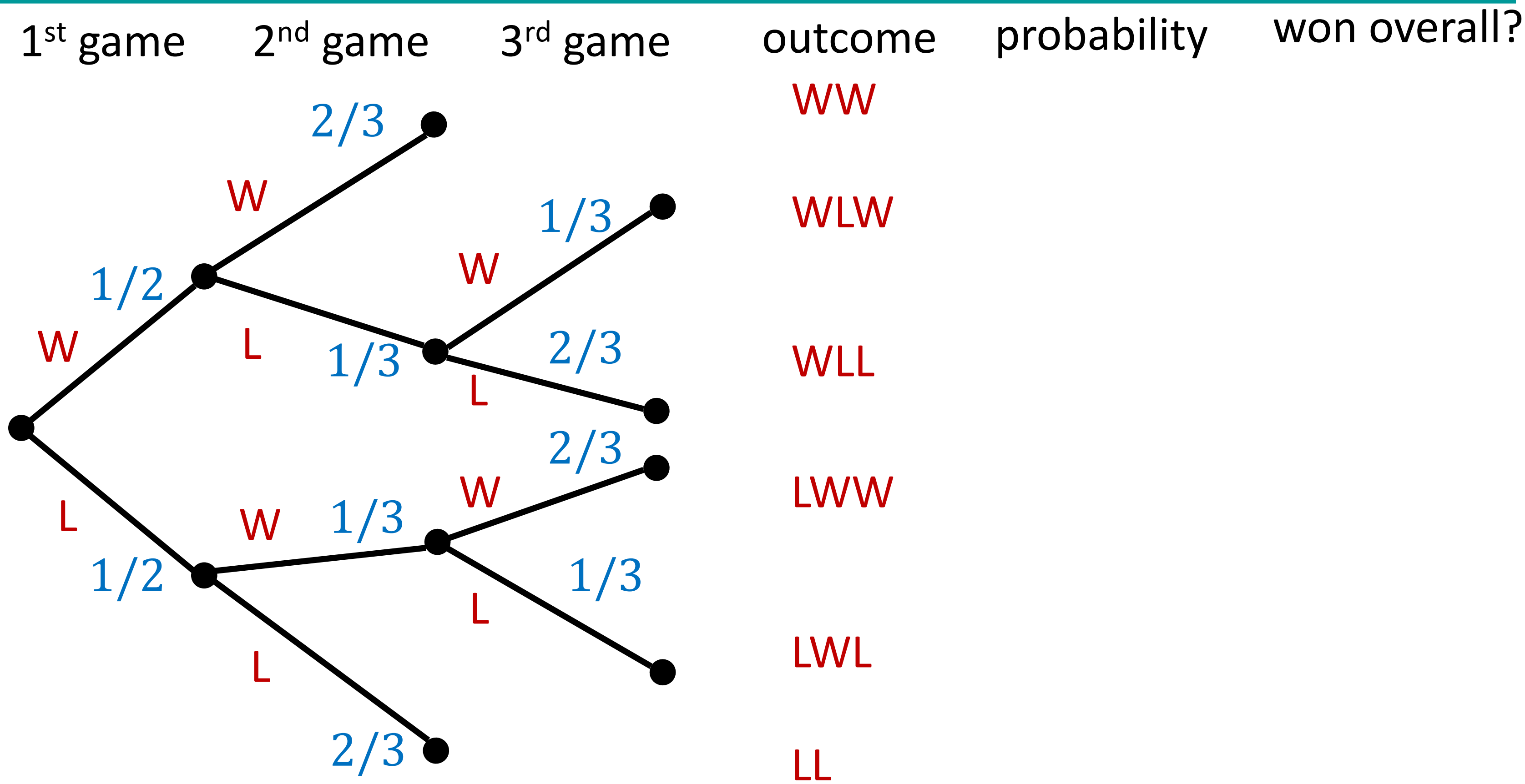
CS 237 Tree Diagrams: Example

- Best-of-three tournament between BU Terriers and Harvard Huskies
- First game: Terriers win with probability $\frac{1}{2}$
- Subsequent games:

$$\text{Terriers win a game with probability } \begin{cases} \frac{2}{3} & \text{if they won the previous game} \\ \frac{1}{3} & \text{if they lost the previous game} \end{cases}$$

What is the probability that the Terriers win the tournament **given** that they win the first game?

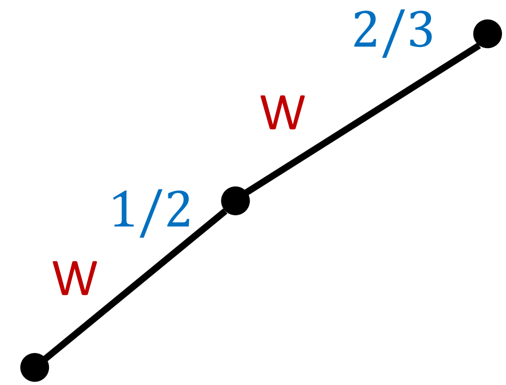
Tree Diagram: Tournament



- Event T =``Terriers win the tournament''
- Event E_1 =``Terriers win the first game"

Tree Diagram: Justification

- Why do we multiply probabilities along each branch in a tree diagram?
- Example:
  - Event E_1 =``Terriers win the first game"
  - Event E_2 =``Terriers win the second game"



- We calculated $\Pr(WW) = \Pr(E_1 \cap E_2)$
  $= \Pr(E_1) \cdot \Pr(E_2 \mid E_1)$
- Product rule follows from the definition of conditional probability:

$$\Pr(E_2 \mid E_1) = \frac{\Pr(E_1 \cap E_2)}{\Pr(E_1)}$$