Probabilities and Expectations

A. Rupam Mahmood

September 10, 2015

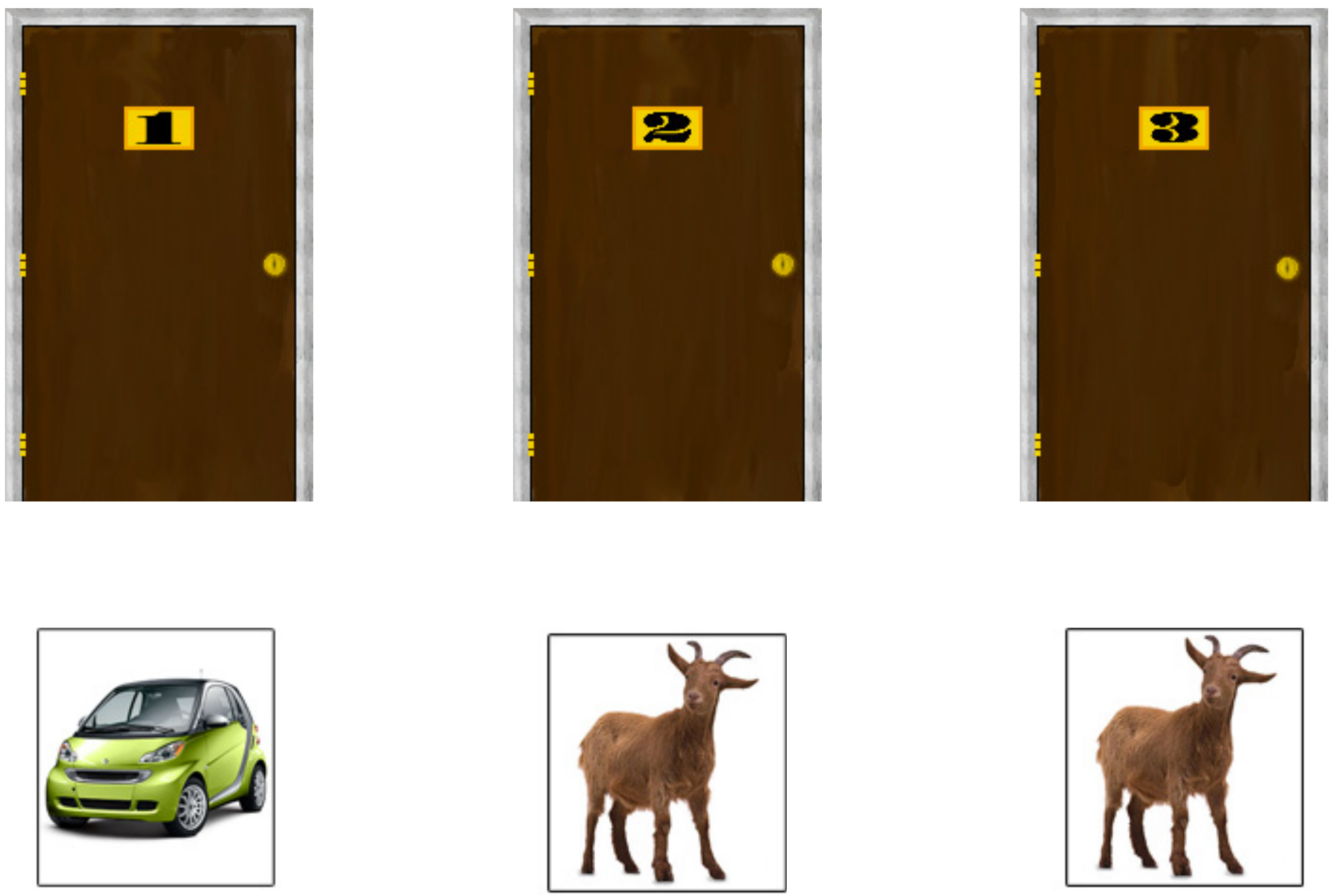
Probabilities

- Probability is a measure of uncertainty
- Being uncertain is much more than “I don’t know”
- We can make informed guesses about uncertain events

Intelligent Systems

- An intelligent system maximizes its “chances” of success
- Intelligent systems create a favorable future
- Probabilities and expectations are tools for reasoning about uncertain future events

Example: Monty Hall Problem

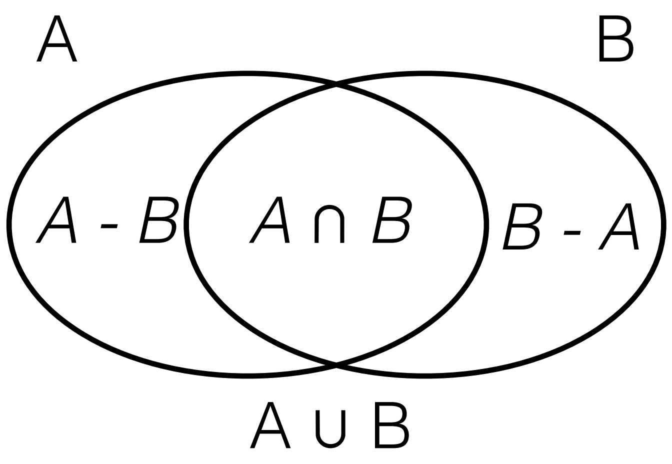


Sets

- A set is a collection of distinct of objects
- $S = \{\text{head, tail}\}$
- Element: head $\in S$, tail $\in S$
- Subsets:$\{\text{head}\} \subset S$, $S \subset S$, $\phi = \{\ \} \subset S$
- Power set: $2^S = \{\{\text{head}\}, \{\text{tail}\}, S, \phi\}$
- Union: $A = \{1, 2\}$, $B = \{2, 3\}$, $A \cup B = \{1, 2, 3\}$
- Intersection: $A = \{1, 2\}$, $B = \{2, 3\}$, $A \cap B = \{2\}$
- A complement set of A in B: $A = \{1, 2\}$, $B = \{2, 3\}$, $B - A = \{3\}$
- The Cartesian product of two sets: $A = \{1, 2\}$, $B = \{a, b\}$, $A \times B = \{(1,a), (1,b), (2,a), (2,b)\}$

Sets

- A set is a collection of distinct of objects
- S = {head, tail},
- Element: head $\in$ S, tail $\in$ S
- Subsets:{head} $\subset$ S, $S \subset S$, ϕ = { } $\subset$ S
- Power set: 2^S = {{head} , {tail} , S, ϕ}
- Union: A = {1, 2}, B = {2, 3}, $A \cup B$ = {1, 2, 3}
- Intersection: A = {1, 2}, B = {2, 3}, $A \cap B$ = {2}
- A complement set of A in B: A = {1, 2}, B = {2, 3}, B – A = {3}
- The Cartesian product of two sets: A = {1, 2}, B = {a, b}, A $\times$ B = {(1,a), (1,b), (2,a), (2,b)}



Functions

- A function is a map from one set to another
- $S = \{\text{head}, \text{tail}\}$, $V = \{+1, -1\}$, $f : S \to V$
- $f\,(\text{head}) = 1$, $f\,(\text{tail}) = -1$
- $f\,(\text{head}) = 1$ ✕
- $f\,(\text{head}) = 1$, $f\,(\text{tail}) = 1$
- $f\,(\text{head}) = 1$, $f\,(\text{head}) = -1$, $f\,(\text{tail}) = 1$ ✕

Sample space & Events

- An experiment is a repeatable process
- A sample space is the set of all possible outcomes of an experiment

  Dice-rolling: $S = \{1, 2, 3, 4, 5, 6\}$
- An event is a subset of a sample space
  the event of even number appearing: $\{2, 4, 6\}$

Probabilities

- Probability is a function that maps all possible events from a sample space to a number

$$\mathbf{Pr}: 2^{S} \rightarrow [0, 1]$$

- Probability is a measure of uncertain events

- **Non-negativity**: *A* probability is always non-negative:

$$0 \leq \mathbf{Pr}(A) \leq 1$$

- **Normalization**: Addition of probabilities of all individual outcomes of a sample space is always 1

$$\sum_{e \in S} \mathbf{Pr}(e) = 1$$

Probability distribution defines how the probability is distributed among the outcomes

- **Additivity**: $\mathbf{Pr}(A \cup B) = \mathbf{Pr}(A) + \mathbf{Pr}(B)$; $A \cap B = \phi$

Random Variables

- Random variables are a convenient way to express events
- A random variable is a function that maps a sample space to a real number

  $X : S \to \mathbb{R}$

- Dice-rolling experiment: $[X < 4]$ stands for

  $\{ \omega \in S : X(\omega) < 4 \} = \{ 1, 2, 3 \}$

Examples

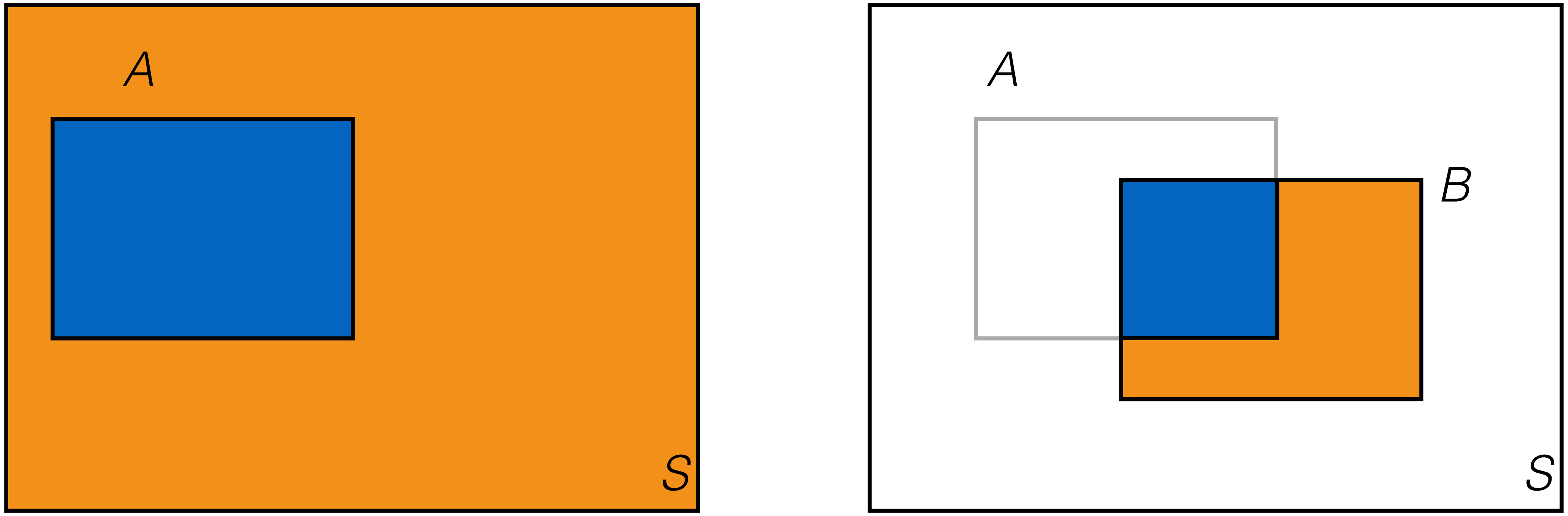
- In the dice-rolling experiment, what is the probability that the outcome is a prime number?
  - Sample space: $S = \{1, 2, 3, 4, 5, 6\}$
  - Distribution: 1/6 for each outcome
  - Event in question: $E = \{2, 3, 5\}$
  - $\mathbf{Pr}(E) = \mathbf{Pr}(\{2, 3, 5\}) = \mathbf{Pr}(2) + \mathbf{Pr}(3) + \mathbf{Pr}(5) = 3X1/6 = 1/2$.

Examples

- If we roll two dices together, what is the probability that sum of the two numbers is greater than 2?
  - Sample space: $S = \{1,\ldots,6\} \times \{1,\ldots,6\}$; (compound experiment)
    $= \{(1,1), (1,2), \ldots, (1,6),$
    $(2,1), (2,2), \ldots, (2,6), \ldots,$
    $(6,1), (6,2), \ldots, (6,6)\}$
  - Distribution: 1/36 for each outcome
  - Event in question: $E = \{(1,2), \ldots, (6,6)\}$
  - Define a random variable to be the sum of the two numbers: $X(a, b) = a + b$
  - Event: $E = [X>2]$
  - $1 = \mathbf{Pr}(S) = \mathbf{Pr}([X=2] \cup [X>2]) = \mathbf{Pr}([X=2]) + \mathbf{Pr}([X>2])$

Conditional Probabilities

- A conditional probability is a measure of an uncertain event when we know that another event has occurred
- Definition: $\mathbf{Pr}(A \mid B) = \mathbf{Pr}(A \cap B) \,/\, \mathbf{Pr}(B) \neq \mathbf{Pr}(A)$



Examples

- In the dice-rolling experiment, if a prime number appears, what is the probability that it is even?
  - Sample space: $S = \{1, 2, 3, 4, 5, 6\}$
  - Distribution: 1/6 for each outcome
  - Events: A = {2, 4, 6}, B = {2, 3, 5}
- $\mathbf{Pr}(A \mid B) = \mathbf{Pr}(A \cap B) / \mathbf{Pr}(B)$

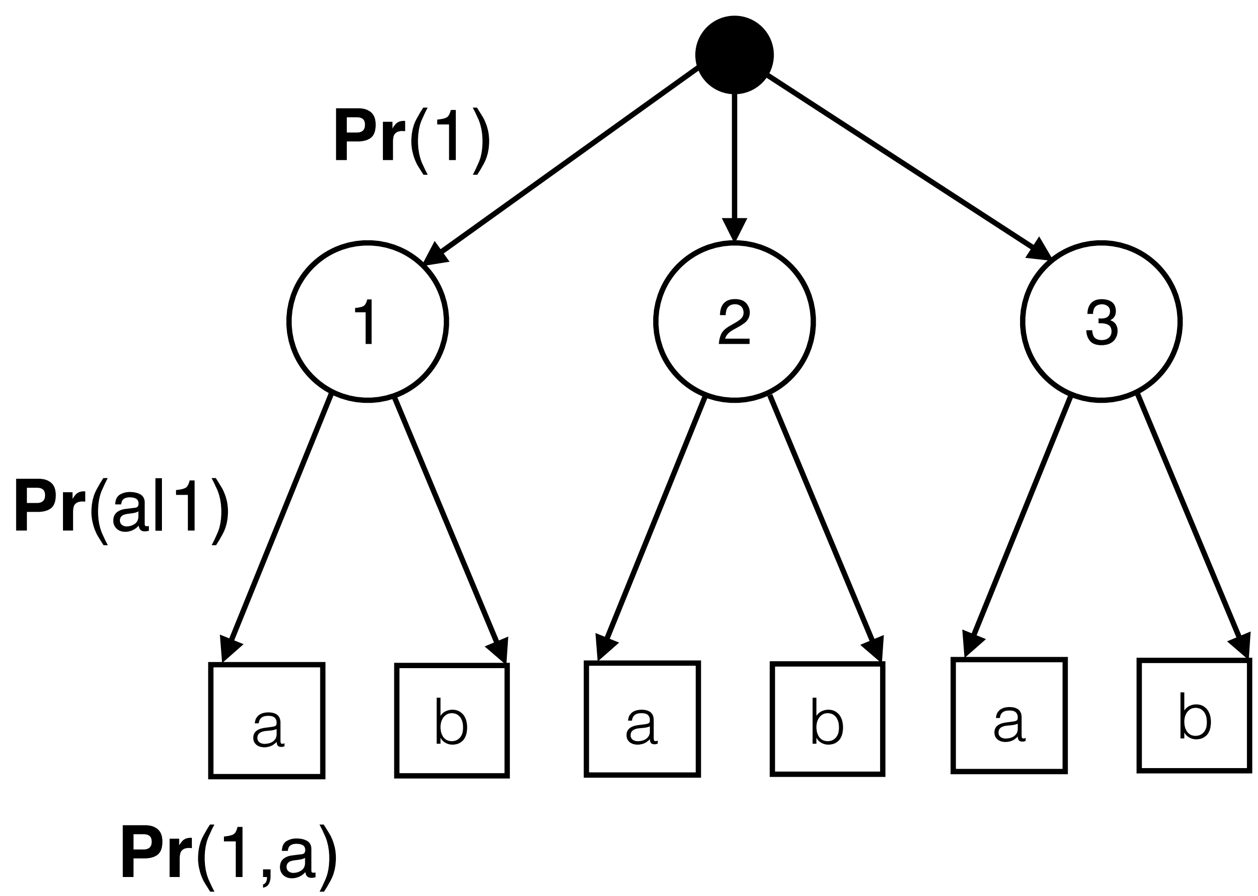
$$= \mathbf{Pr}(\{2, 4, 6\} \cap \{2, 3, 5\}) / \mathbf{Pr}(\{2, 3, 5\})$$

$$= \mathbf{Pr}(2) / \mathbf{Pr}(\{2, 3, 5\})$$

$$= (1/6)/(1/2) = 1/3$$

Probability Trees

- Often in compound experiments, outcome of one depends on the other (unlike the double dice-rolling experiment)
- It is convenient in that case to calculate probabilities using probability trees

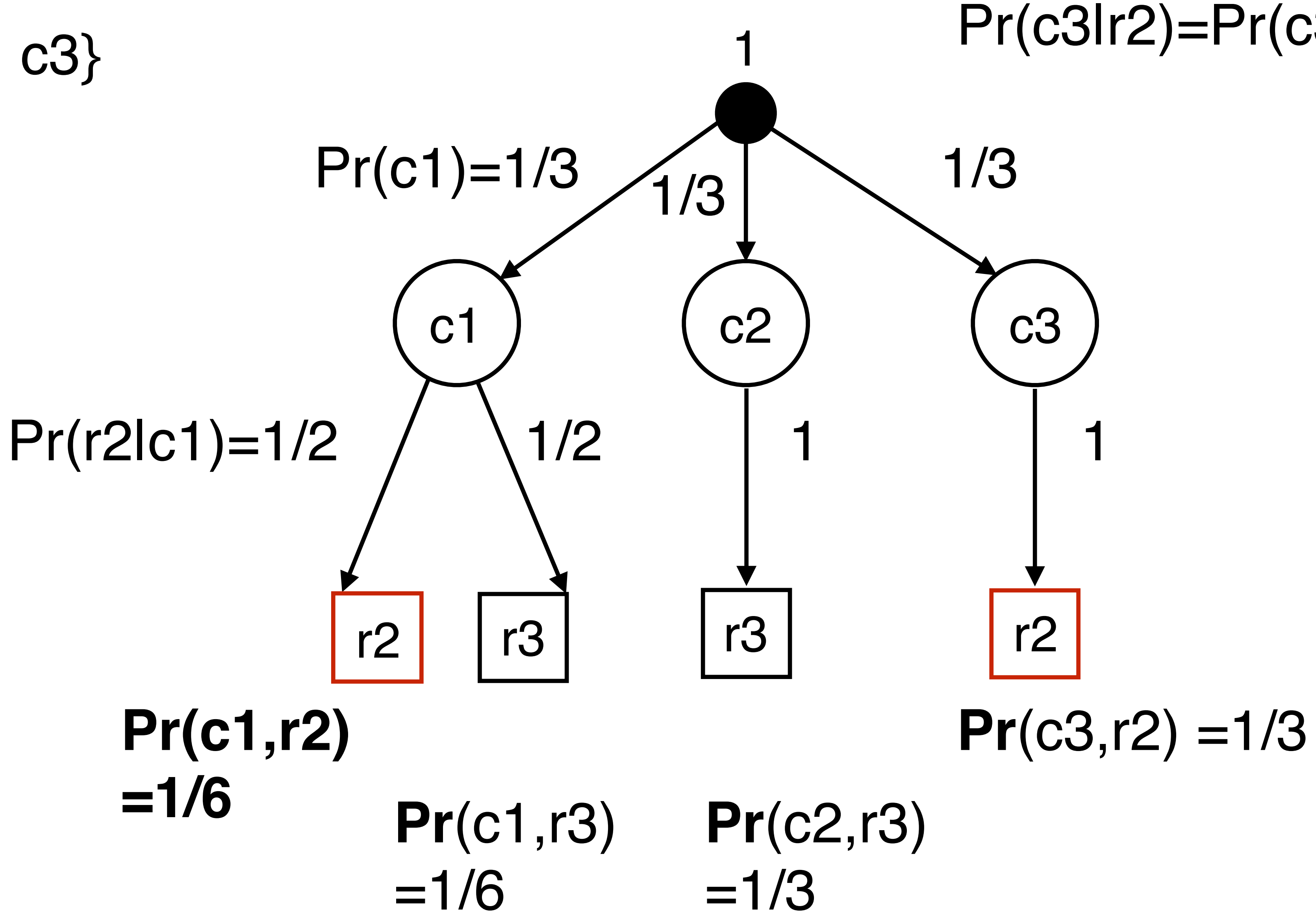


Examples: Monty Hall Problem

- In the Monty Hall problem, we chose the 1st door and the Host revealed 2nd.

Pr(c1|r2)=Pr(c1,r2)/Pr(r2)
Pr(c3|r2)=Pr(c3,r2)/Pr(r2)

$S_1 = \{c1, c2, c3\}$
$S_2 = \{r2, r3\}$
$S = S_1 \times S_2$

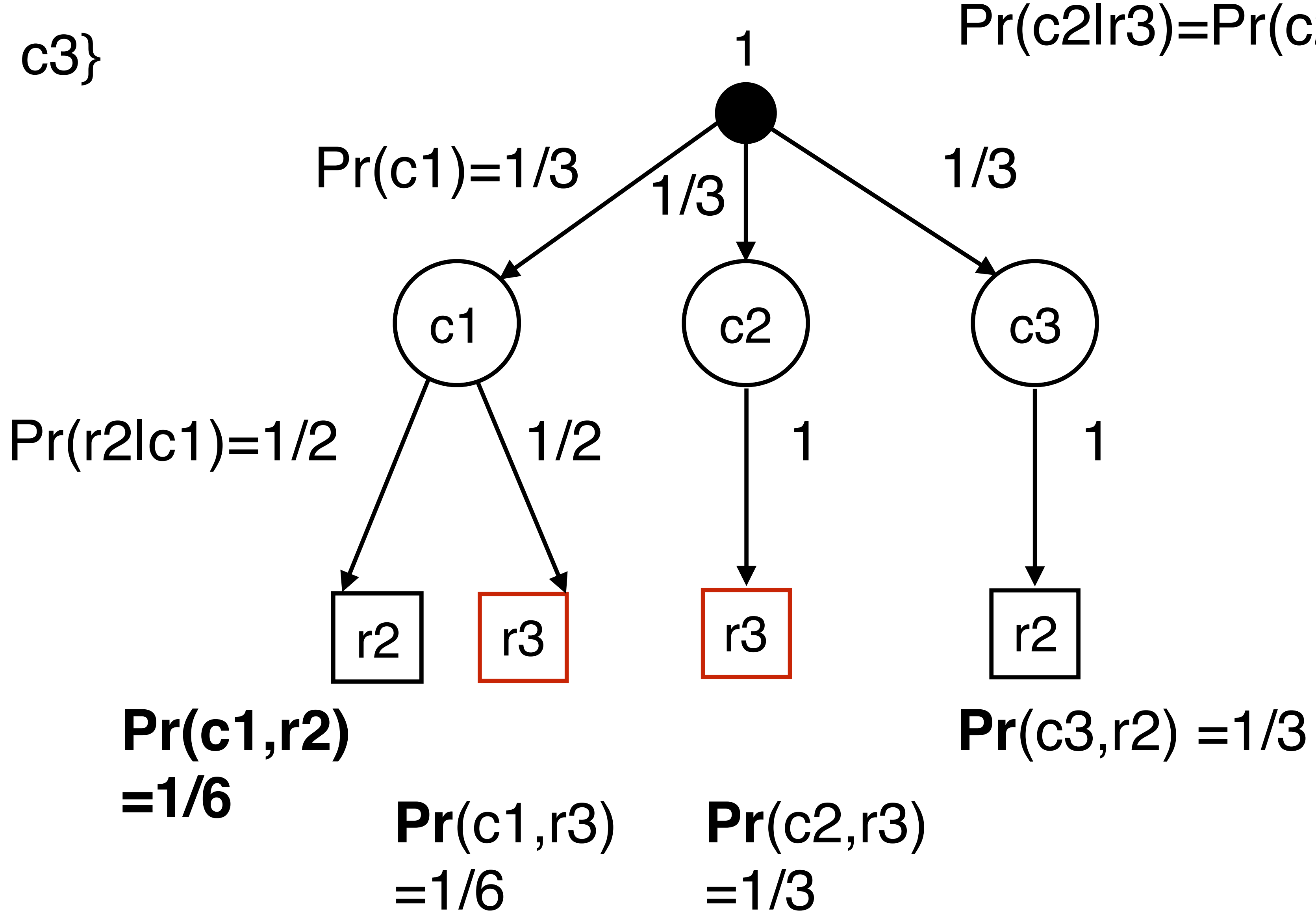


Examples: Monty Hall Problem

- In the Monty Hall problem, we chose the 1st door and the Host revealed 3rd.

$S_1 = \{c1, c2, c3\}$
$S_2 = \{r2, r3\}$
$S = S_1 \times S_2$

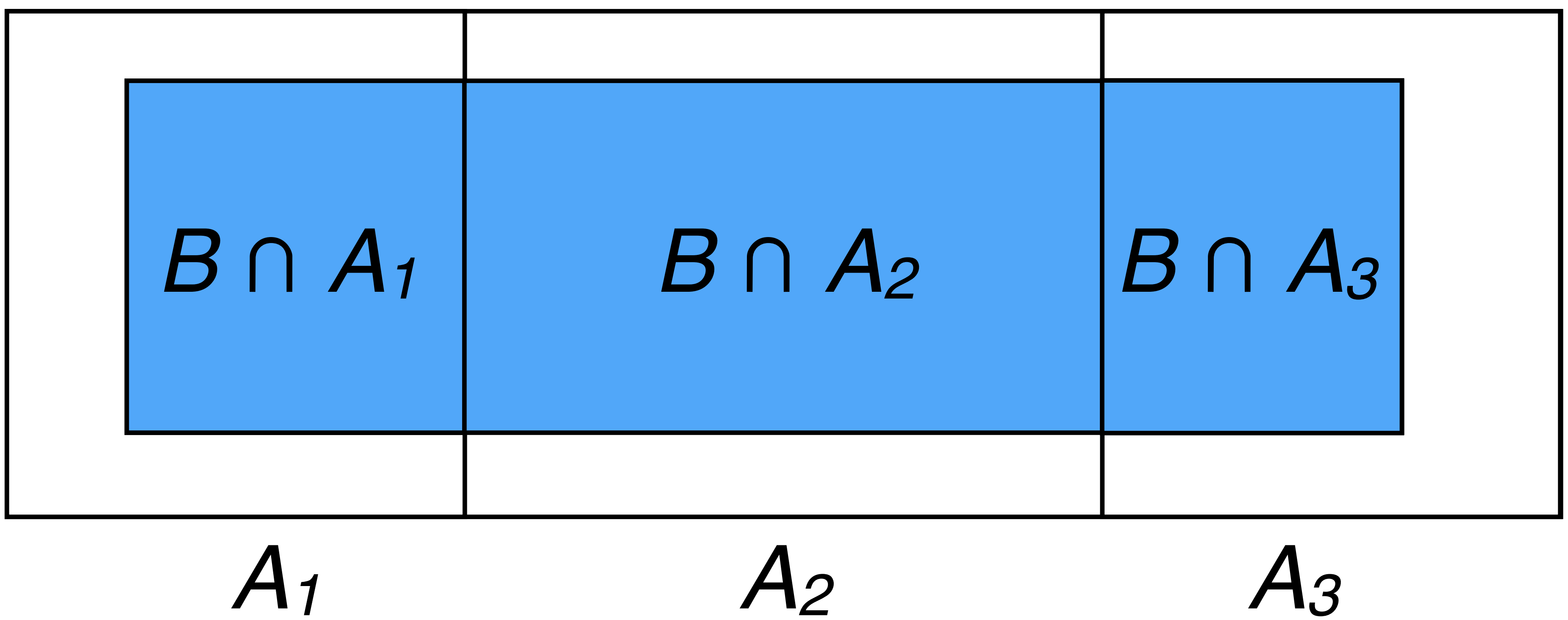
Pr(c1|r3)=Pr(c1,r3)/Pr(r3)
Pr(c2|r3)=Pr(c2,r3)/Pr(r3)



Law of Total Probability

$$\mathbf{Pr}(B) = \sum_j \mathbf{Pr}(B \cap A_j)$$
$$= \sum_j \mathbf{Pr}(B \mid A_j)\, \mathbf{Pr}(A_j)$$

$$A_i \cap A_j = \phi,\ i \neq j, \quad \cup_i A_i = S$$



Bayes Theorem

$$\mathbf{Pr}(A_1|B) = \frac{\mathbf{Pr}(A_1 \cap B)}{\mathbf{Pr}(B)} = \frac{\mathbf{Pr}(B|A_1)\mathbf{Pr}(A_1)}{\sum_j \mathbf{Pr}(B|A_j)\mathbf{Pr}(A_j)}$$

Examples

- A drug test returns positive for a drug user 99% of the time and returns negative for a non-user 95% of the time. Suppose that 1% of the population uses drug. Then what is the probability that an individual is a drug user given that she tests positive?

- Sample space: { user+, user-, nonuser+, nonuser-}

- **Pr**(+luser) = 0.99
- **Pr**(-lnonuser) = 0.95
- **Pr**(user) = 0.01
- **Pr**(userl+) = ?

$$
\begin{aligned}
\Pr(\texttt{user}|+) &= \frac{\Pr(+|\texttt{user})\Pr(\texttt{user})}{\Pr(+|\texttt{user})\Pr(\texttt{user}) + \Pr(+|\texttt{nonuser})\Pr(\texttt{nonuser})} \\
&= \frac{0.99 \times 0.01}{0.99 \times 0.01 + (1 - \Pr(-|\texttt{nonuser})) \times (1 - \Pr(\texttt{user}))} \\
&= \frac{0.0099}{0.0099 + (1 - 0.95) \times (1 - 0.01)} \\
&= \frac{0.0099}{0.0099 + 0.05 \times 0.99} \\
&= \frac{0.0099}{0.0099 + 0.0495} \\
&\approx 0.167.
\end{aligned}
$$

Expectations & Conditional Expectations

- An expected value of a random variable is a weighted average of possible outcomes, where the weights are the probabilities of those outcomes

$$\mathbf{E}[X] = \sum_{x \in S} x\, \mathbf{Pr}(X=x)$$

- An expected value of a random variable conditional on another event is a weighted average of possible outcomes, where the weights are the conditional probabilities of those outcomes given the event

$$\mathbf{E}[X \mid Y=y] = \sum_{x \in S} x\, \mathbf{Pr}(X=x \mid Y=y)$$

- Law of total expectation: $\mathbf{E}[X] = \sum_y \mathbf{E}[X \mid Y=y]\, \mathbf{Pr}(Y=y)$

Examples

- In a certain lottery, it is 0.01% likely to win, and the prize is 1000 dollars. The ticket price is 10 dollars. What is the expected monetary gain?

- Sample space: $S = \{ 990, -10 \}$

- Expected value: $\mathbf{E}[X] = 990\ \mathbf{Pr}(X{=}990) + (-10)\ \mathbf{Pr}(X{=}-10)$

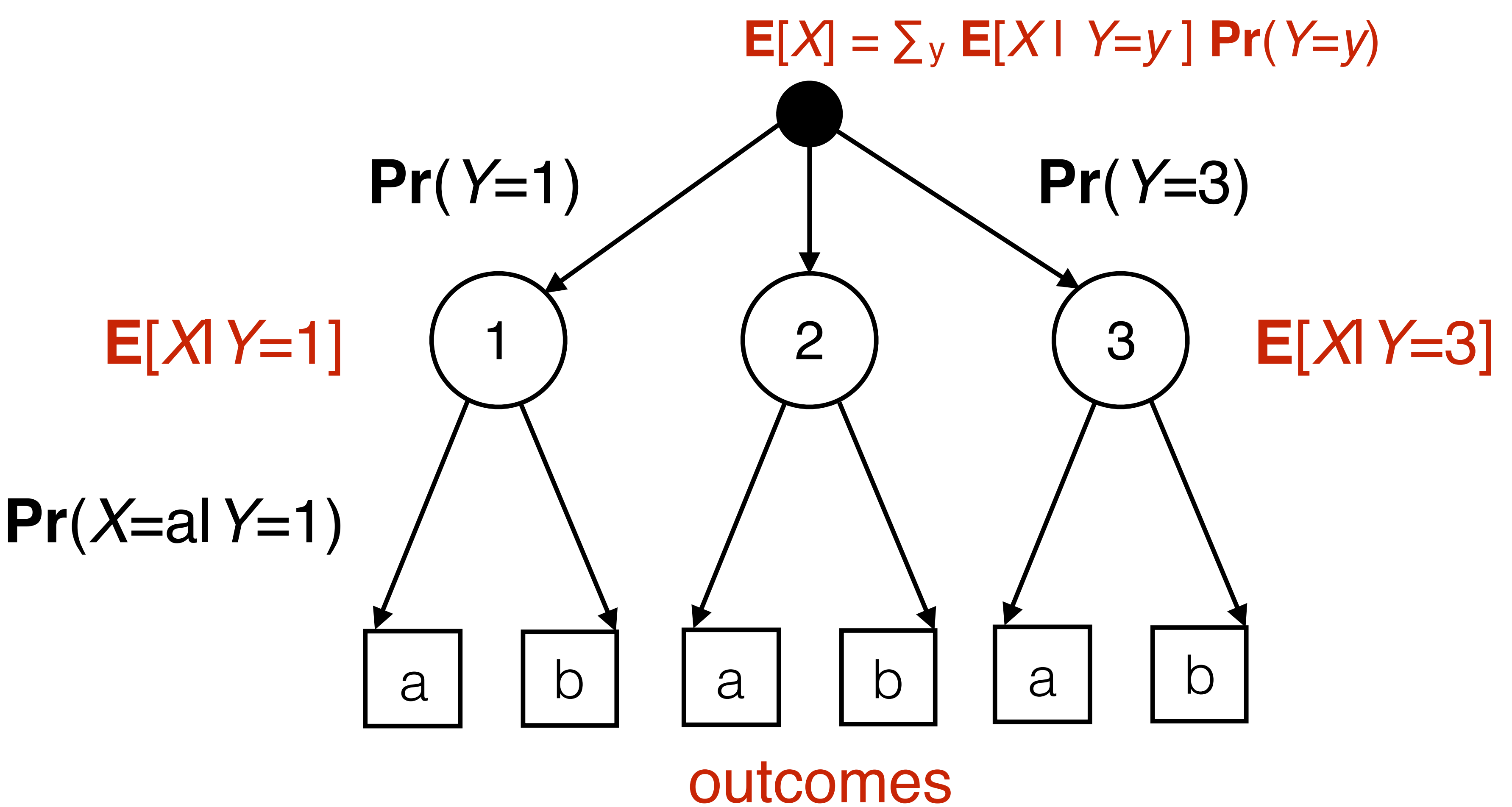
  $= 990 * 0.0001 + (-10) * 0.9999$

  $= 0.099 - 9.999$

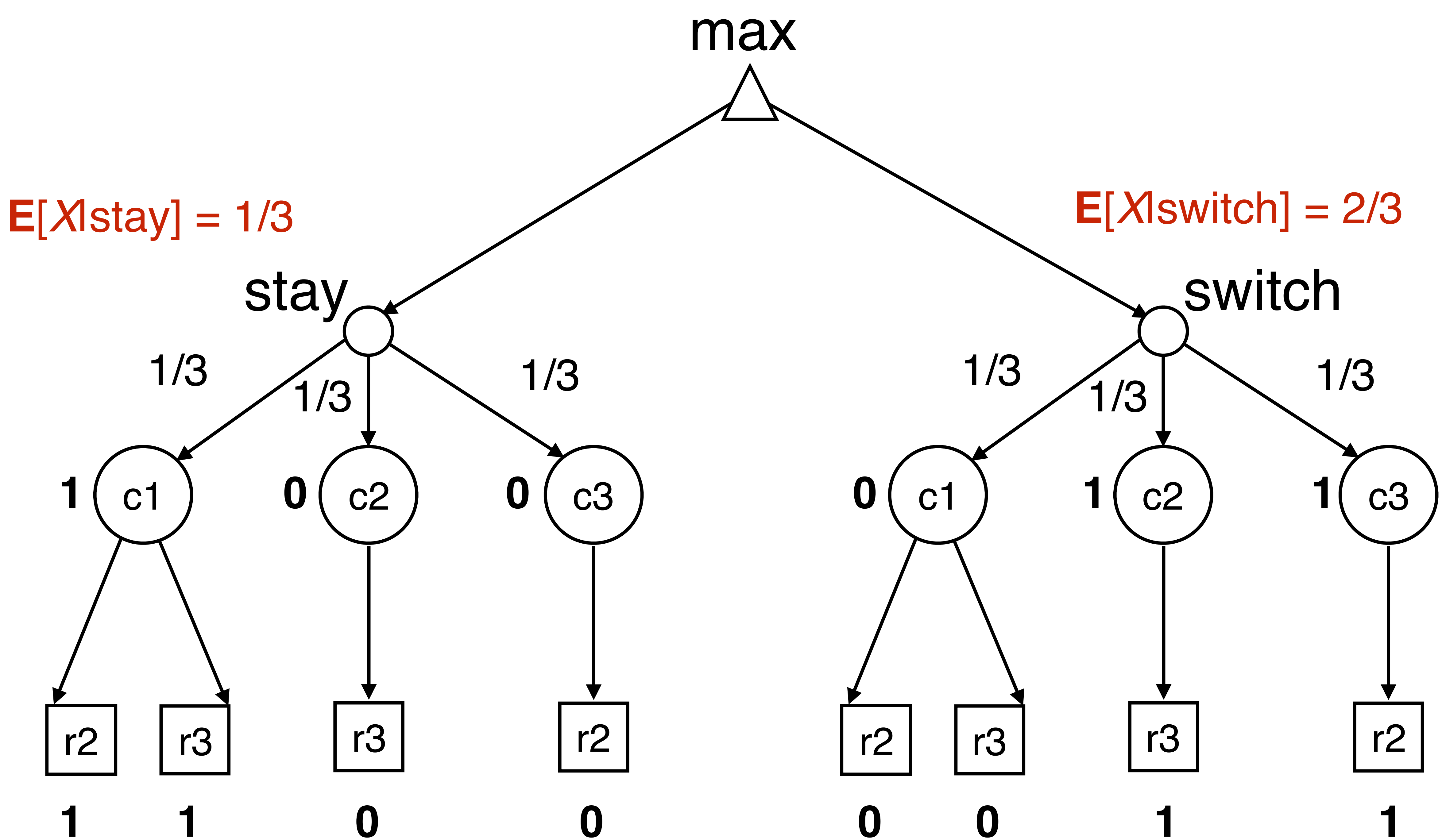
  $= -9.9.$

Expectation Trees

- Often in compound experiments, outcome of one depends on the other (unlike the double dice-rolling experiment)
- It is convenient in that case to calculate probabilities using probability trees



Examples: Monty Hall Problem



Two Ways of Calculation

- Model-based calculation
  - We know the probability model
- Model-free or empirical estimation
  - Learn from experience!

Concluding Remarks

- Probabilities and expectations let us make favorable choices
- There are two ways of calculating them
- If we know the model, we can make intelligent systems by feeding them the model and automating the calculation
- If we do not know the model, we can let the intelligent try things out!
- In either case, intelligent systems can make favorable choices by dealing with probabilities and expectations