15-251
Great Theoretical Ideas in Computer Science

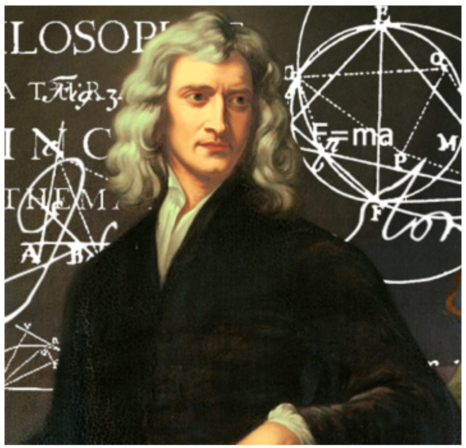
Lecture 21:
Introduction to Randomness and Probability Theory Review

April 4th, 2017

Randomness and the Universe

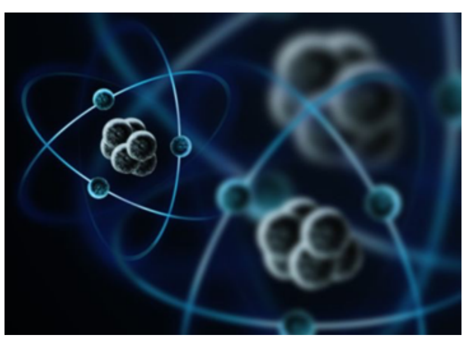
Randomness and the Universe

Does the universe have "true" randomness?



Newtonian physics:

Universe evolves deterministically.

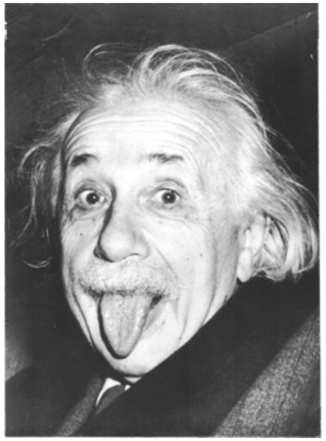


Quantum physics:

Wrong!

Randomness and the Universe

Does the universe have “true” randomness?



God does not play dice with the world.

\- *Albert Einstein*

Einstein, don’t tell God what to do.

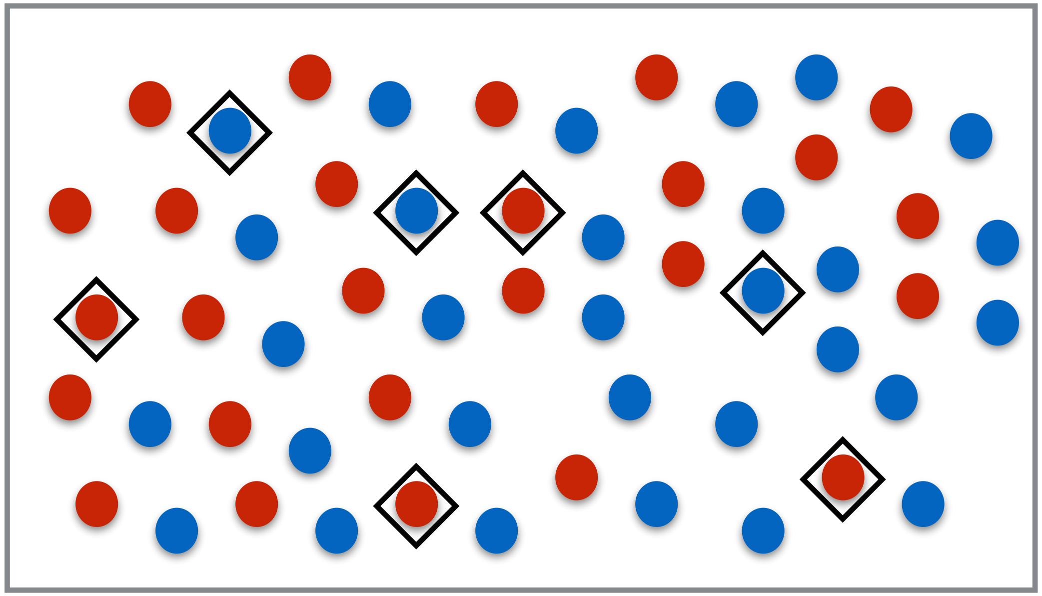
\- *Niels Bohr*

Randomness is an essential tool in
modeling and analyzing nature.

It also plays a key role in **computer science**.

Randomness and Computer Science

Statistics via Sampling



Population: 300m **Random sample size**: 2000

Theorem: With more than 99% probability,

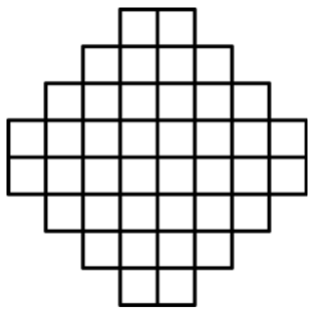
% in sample = % in population ± 2%.

Randomized Algorithms

Dimer Problem:

Given a region, in how many different ways can you tile it with 2x1 rectangles (dominoes)?

e.g.



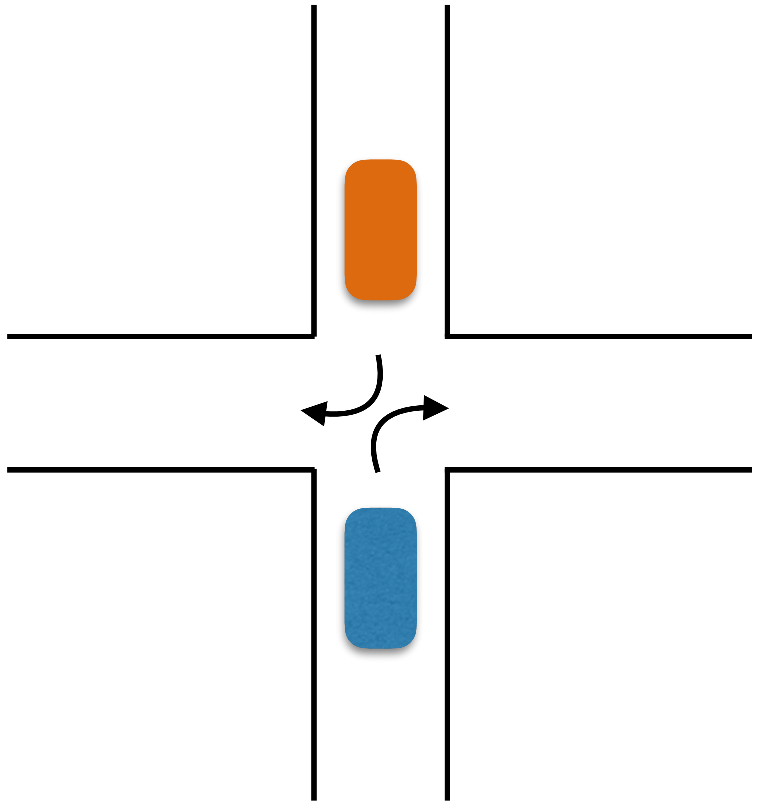
$\longrightarrow$ 1024 tilings

m x n rectangle $\longrightarrow$ $\displaystyle\prod_{j=1}^{\lceil \frac{m}{2} \rceil} \prod_{k=1}^{\lceil \frac{n}{2} \rceil} \left(4\cos^2 \frac{\pi j}{m+1} + 4\cos^2 \frac{\pi k}{n+1} \right)$ tilings

Captures thermodynamic properties of matter.

- Fast randomized algs can approximately count.
- No fast deterministic alg known.

Distributed Computing



Break symmetry with randomness.

Many more examples in the field of *distributed computing*.

Nash Equilibria in Games

The Chicken Game

	Swerve	Straight
Swerve	1 1	0 2
Straight	2 0	-3 -3

Theorem [Nash]: Every game has a Nash Equilibrium provided players can pick a randomized strategy.

Exercise: What is a NE for the game above?

Cryptography

Adversary
Eavesdropper

"I will cut your throat"

"I will cut your throat"

Cryptography

Adversary
Eavesdropper

“I will cut your throat”

encryption

“loru23n8uladjkfb!#@”

“loru23n8uladjkfb!#@”

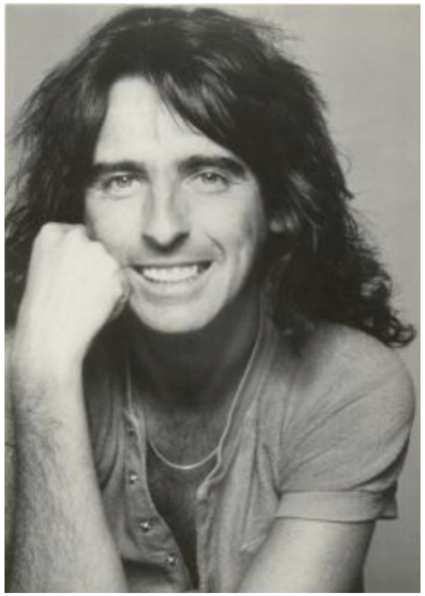
“loru23n8uladjkfb!#@”

decryption

“I will cut your throat”

Shannon: A secret is as good as the amount of entropy/uncertainty/randomness in it.

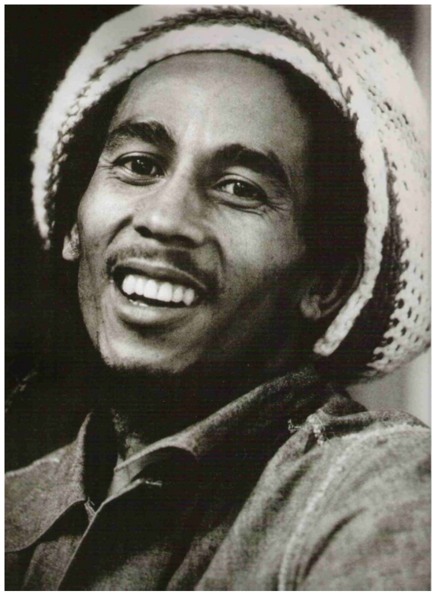
Error-Correcting Codes



Alice

"bit.ly/vrxUBN"

noisy channel



Bob

Each symbol can be corrupted with a certain probability.

How can Alice still get the message across?

Communication Complexity

Want to check if the contents of two databases are exactly the same.

How many bits need to be communicated?

Interactive Proofs

Verifier

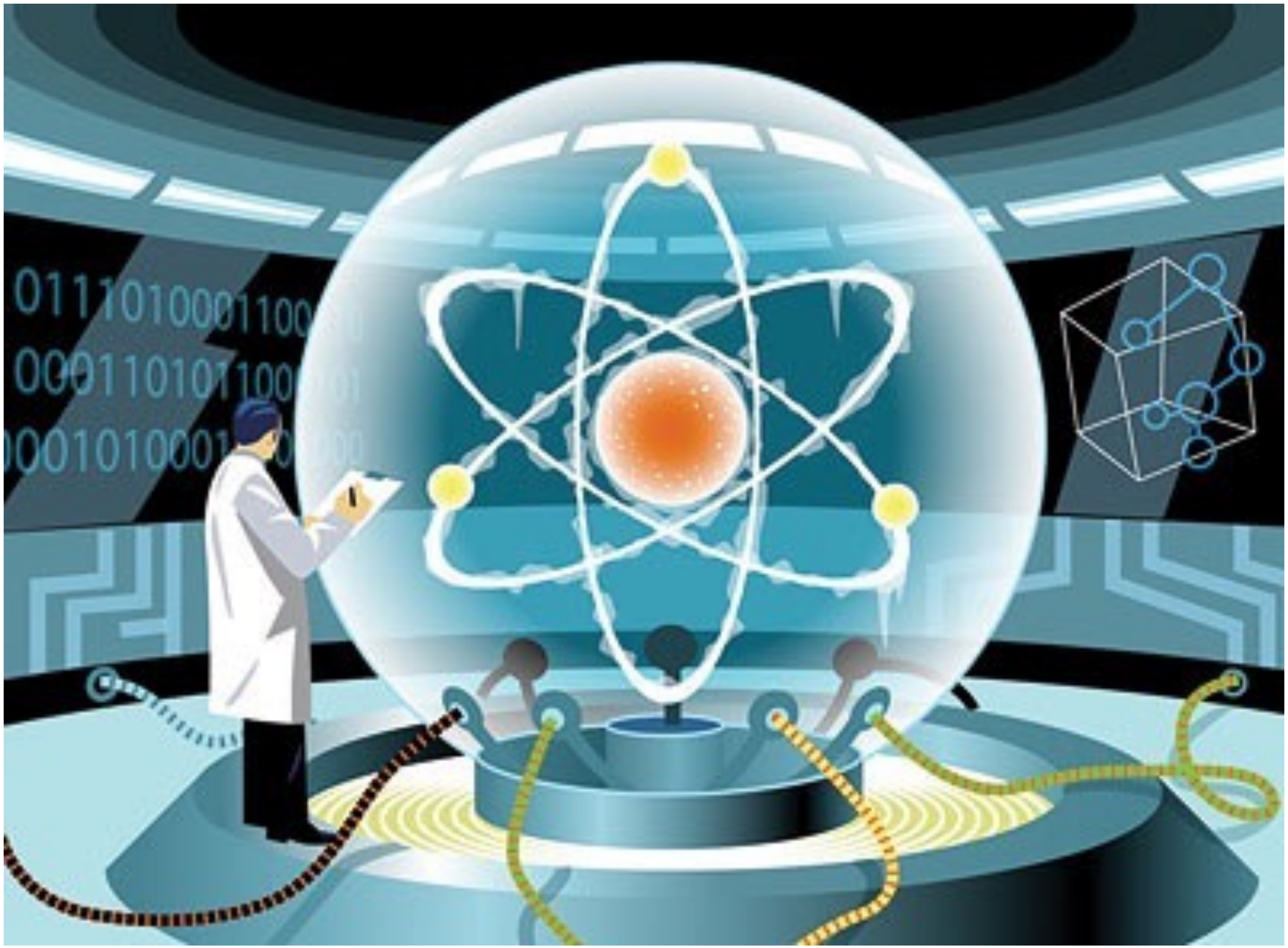
poly-time
skeptical

Prover

omniscient
untrustworthy

Can I convince you that I have proved $P \neq NP$
without revealing any information about the proof?

Quantum Computing



Some Probability Puzzles (Test Your Intuition)
and
Origins of Probability Theory

Origins of Probability Theory

France, 1654

Let's bet:

I will roll a dice four times.
I win if I get a 1.

"Chevalier de Méré"

Antoine Gombaud

Origins of Probability Theory

France, 1654

Hmm.

No one wants to take this bet anymore. :-(

“Chevalier de Méré”

Antoine Gombaud

Origins of Probability Theory

France, 1654

New bet:

I will roll two dice, 24 times.

I win if I get double-1's.

"Chevalier de Méré"

Antoine Gombaud

Origins of Probability Theory

France, 1654

Hmm.

I keep losing money! :-(

"Chevalier de Méré"

Antoine Gombaud

Origins of Probability Theory

France, 1654

"Chevalier de Méré"

Antoine Gombaud

Alice and Bob are flipping a coin.
Alice gets a point for heads.
Bob gets a point for tails.
First one to 4 points wins 100 Fr.

Alice is ahead 3-2 when gendarmes arrive to break up the game.

How should they divide the stakes?

Origins of Probability Theory

Pascal

Fermat

Probability Theory is born!

Probability Theory:
The CS Approach

The Big Picture

The Non-CS Approach

Real World → Mathematical Model

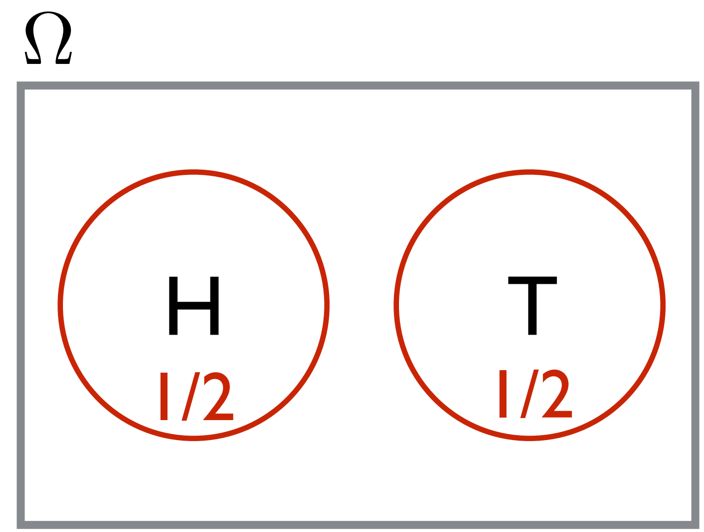
(random) experiment/process

probability space

The Big Picture

Real World → Mathematical Model

Flip a coin.



Ω = "sample space"

= set of all possible outcomes

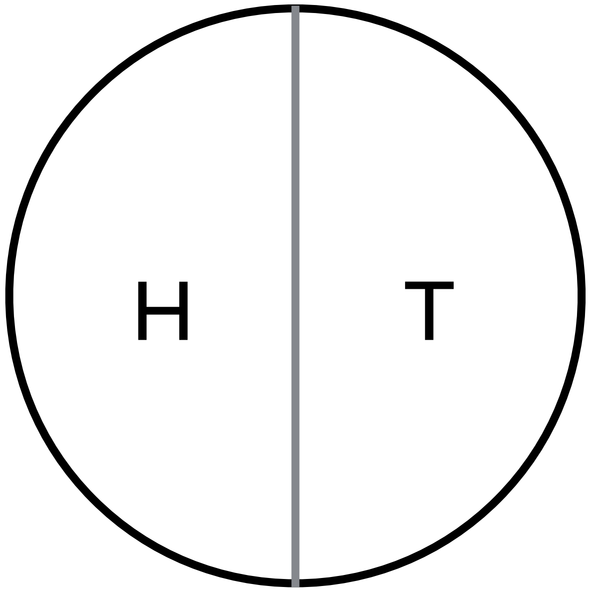
$\mathrm{Pr} : \Omega \to [0, 1]$ prob. distribution

$$\sum_{\ell \in \Omega} \mathrm{Pr}[\ell] = 1$$

The Big Picture

Real World → Mathematical Model

Flip a coin.



unit pie, area = 1

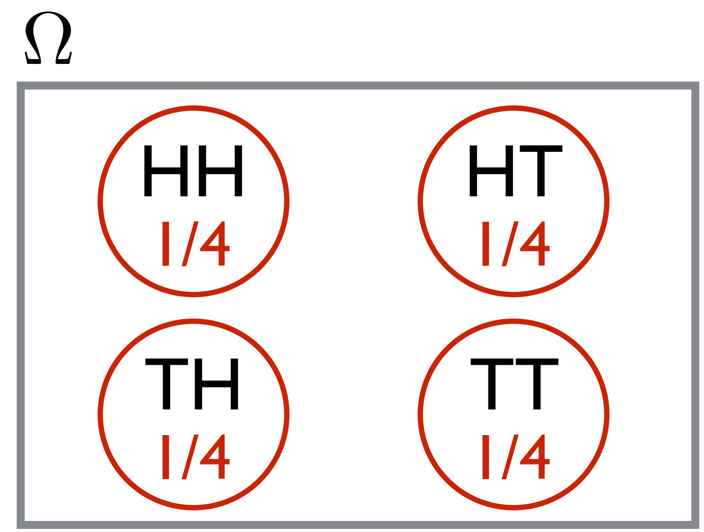
$$\Pr[\text{outcome}] = \text{area of outcome}$$

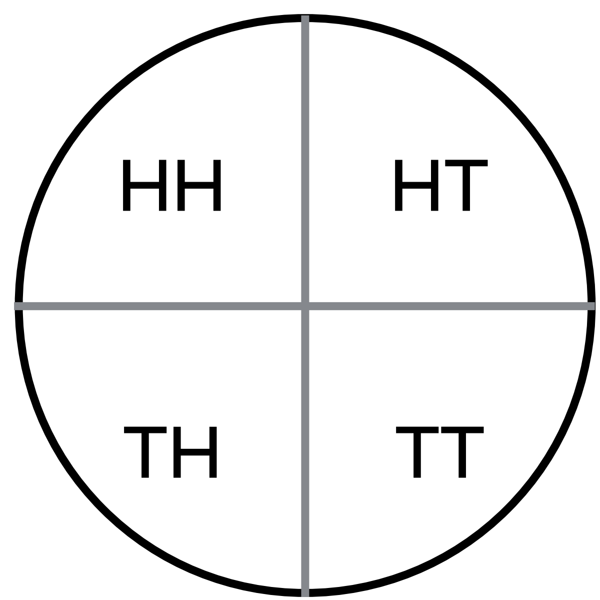
$$= \frac{\text{area of outcome}}{\text{area of pie}}$$

The Big Picture

Real World → Mathematical Model

Flip two coins.

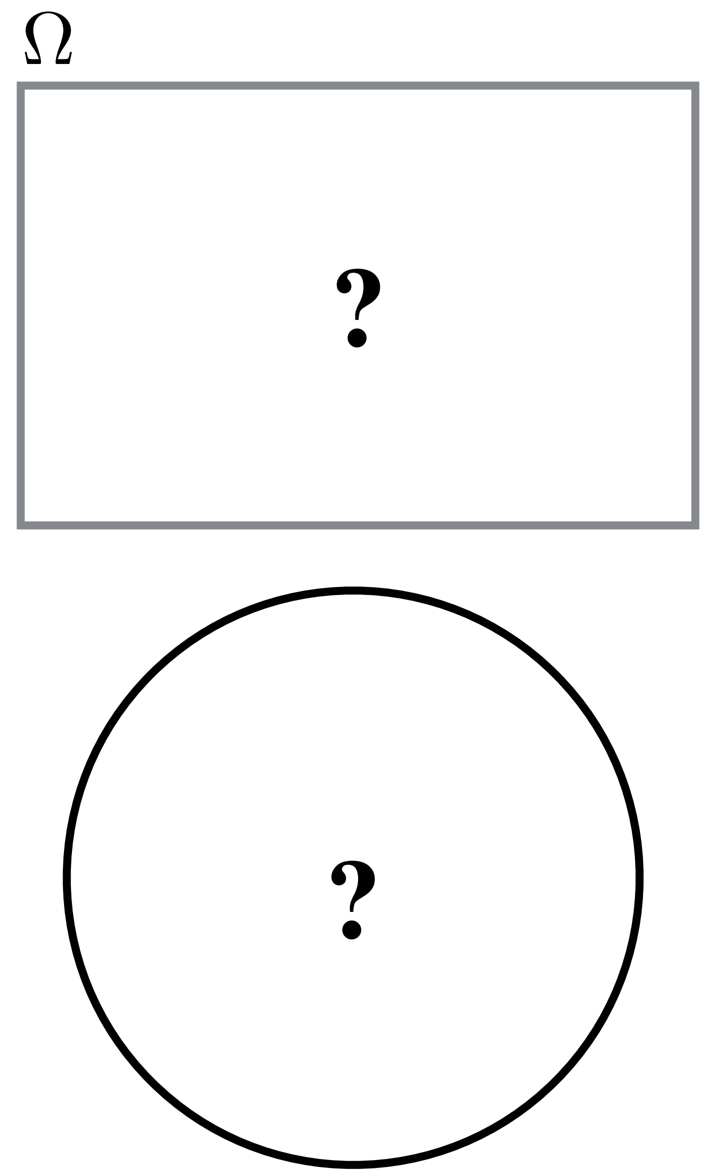




The Big Picture

Real World $\longrightarrow$ Mathematical Model

Flip a coin.
If it is Heads, throw a 3-sided die.
If it is Tails, throw a 4-sided die.



The Big Picture

The CS Approach

Real World → Code → Probability Tree = Mathematical Model

The Big Picture

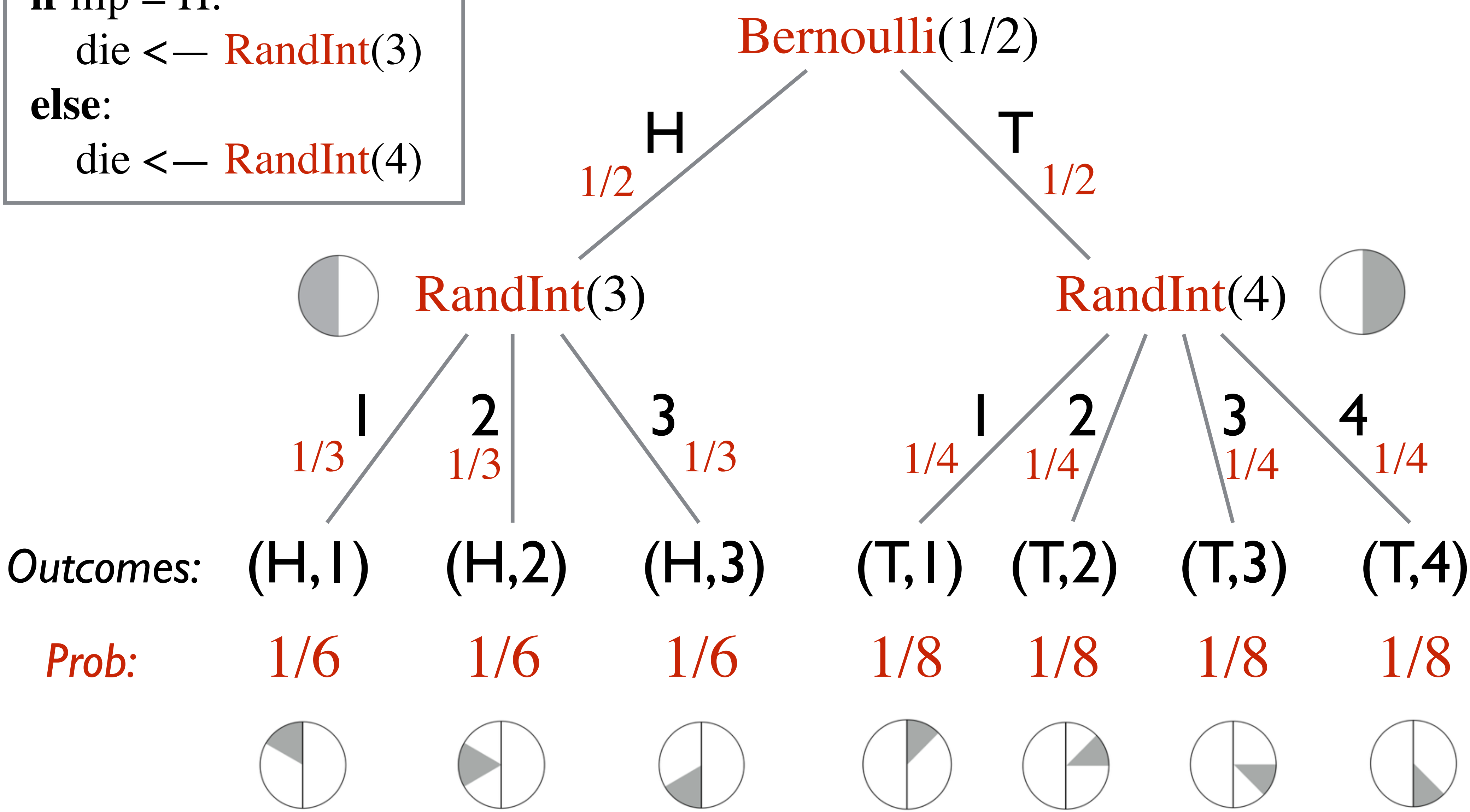
Real World → Code → Probability Tree

Flip a coin.
If it is Heads, throw a 3-sided die.
If it is Tails, throw a 4-sided die.

```
flip <— Bernoulli(1/2)
if flip = 1: # i.e. Heads
    die <— RandInt(3)
else:
    die <— RandInt(4)
```

Probability Tree

```
flip <— Bernoulli(1/2)
if flip = H:
    die <— RandInt(3)
else:
    die <— RandInt(4)
```



Events

Real World → Code → Probability Tree

Flip a coin.
If it is Heads, throw a 3-sided die.
If it is Tails, throw a 4-sided die.

→

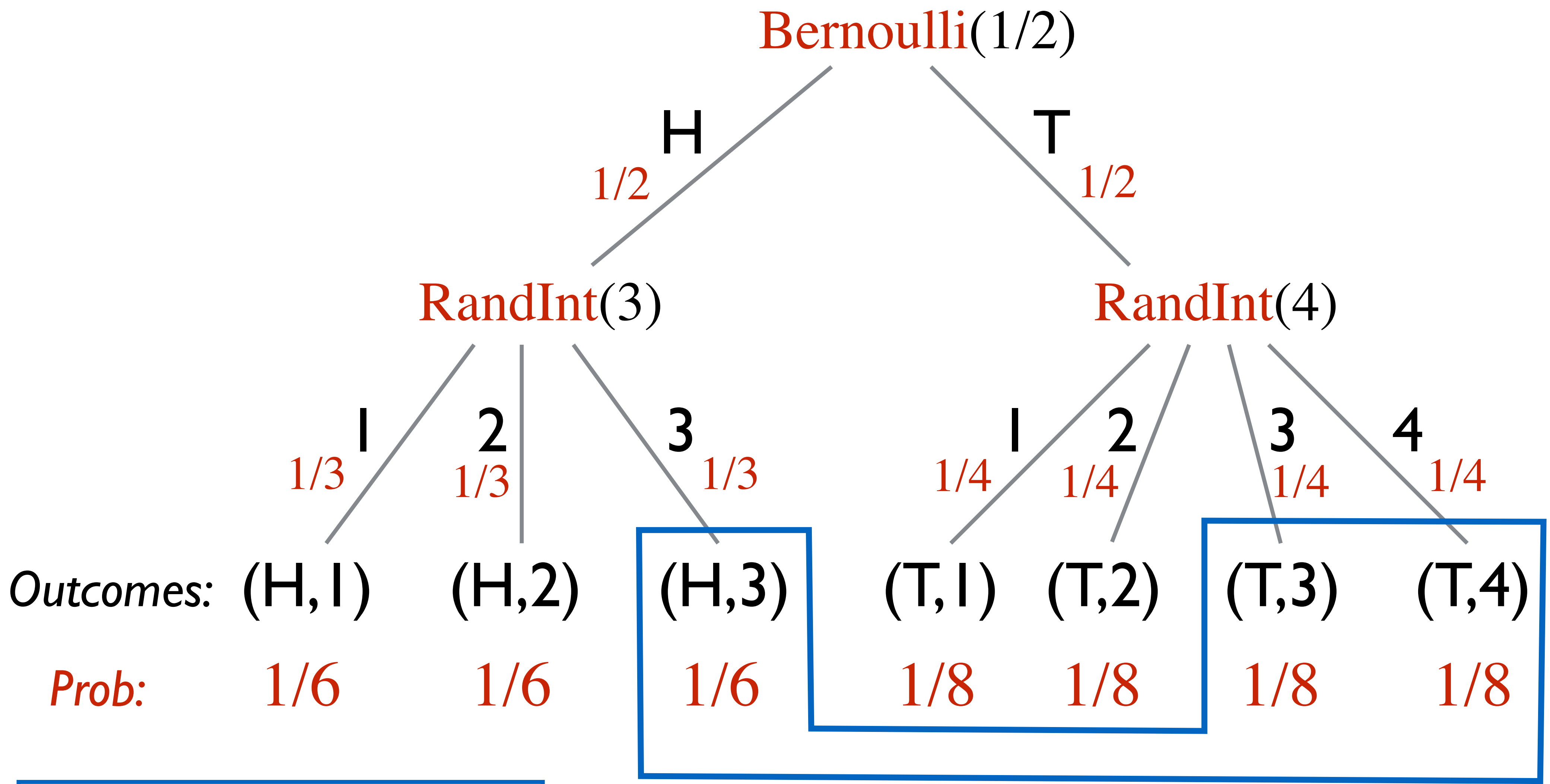
```
flip <— Bernoulli(1/2)
if flip = H:
    die <— RandInt(3)
else:
    die <— RandInt(4)
```

→

What is the probability $\boxed{\text{die roll is} \geq 3}$?

"event" → subset of outcomes/leaves

Events



E = die roll is 3 or higher

$$\Pr[E] = 1/6 + 1/8 + 1/8 = 5/12$$

Extend $\Pr$ to:

$$\Pr : \mathcal{P}(\Omega) \to [0, 1]$$

Conditional Probability

Real World → Code → Probability Tree

Flip a coin.
If it is Heads, throw a 3-sided die.
If it is Tails, throw a 4-sided die.

```
flip <— Bernoulli(1/2)
if flip = H:
    die <— RandInt(3)
else:
    die <— RandInt(4)
```

What is the probability of flipping Heads $\boxed{\text{given the die roll is } \geq 3}$? → conditional probability

conditioning on partial information

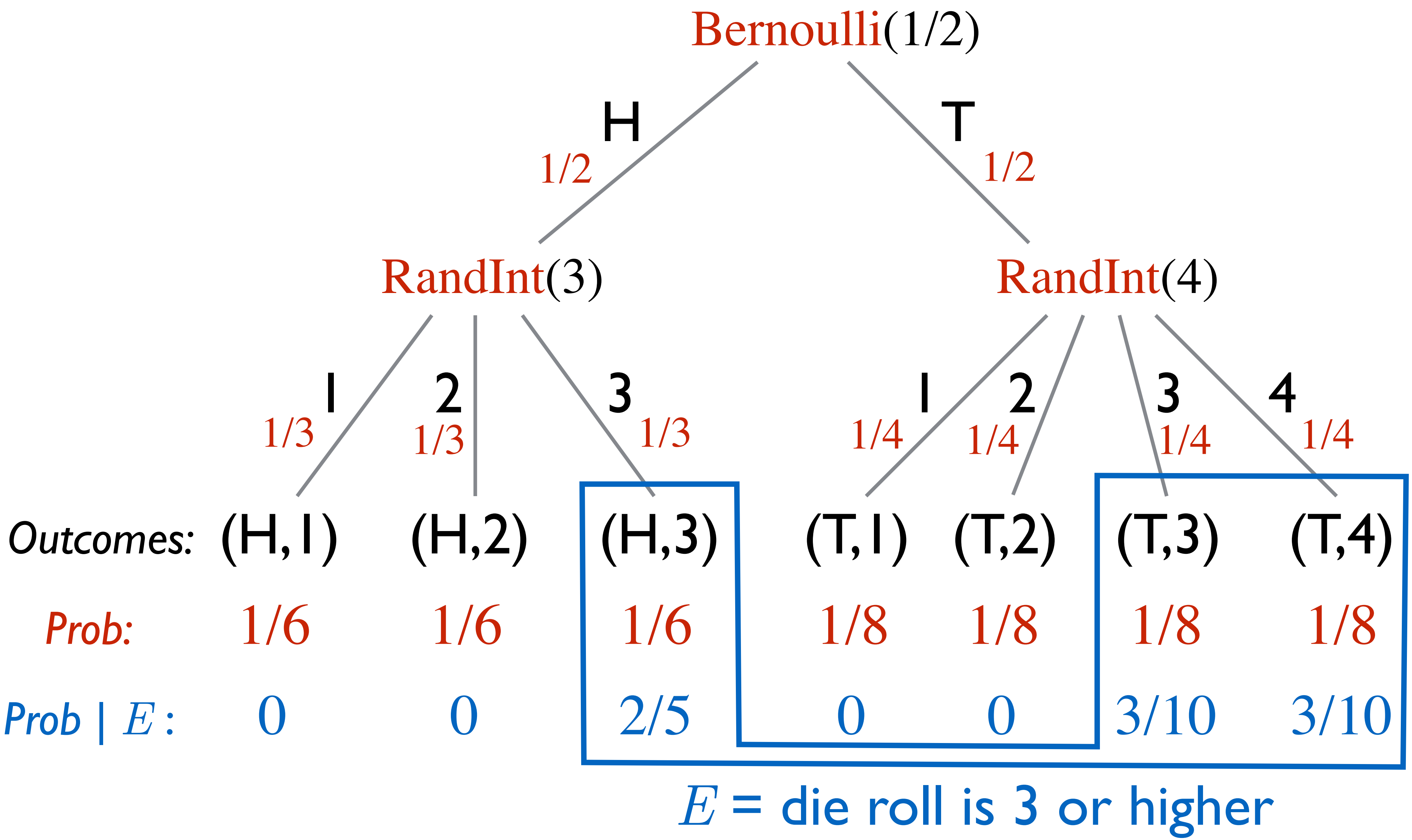
Conditional Probability

Revising probabilities based on '*partial information*'.

'*partial information*' = event E

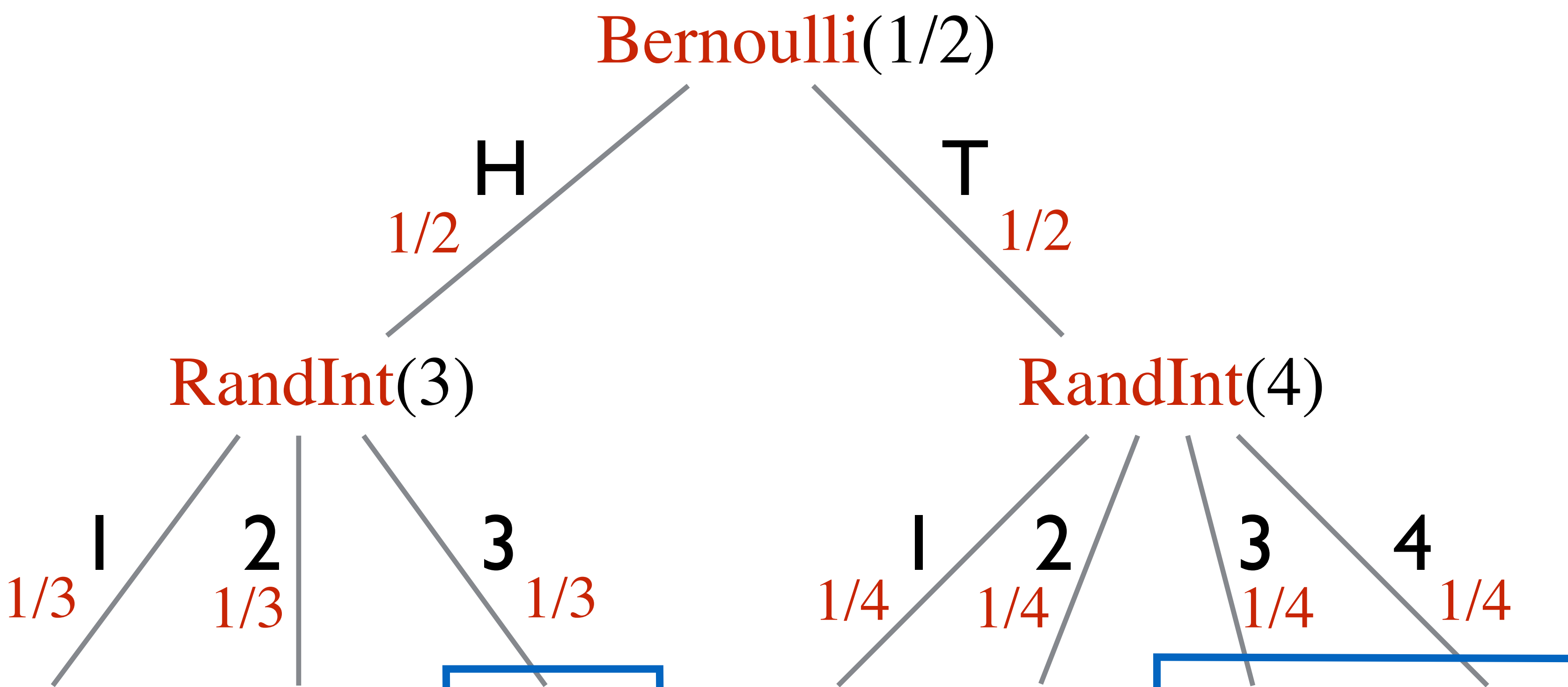
Conditioning on E = Assuming/promising E has happened

Conditional Probability



$$\Pr[(H,1) \mid E] = 0 \qquad \Pr[(H,3) \mid E] = 2/5$$

Conditional Probability



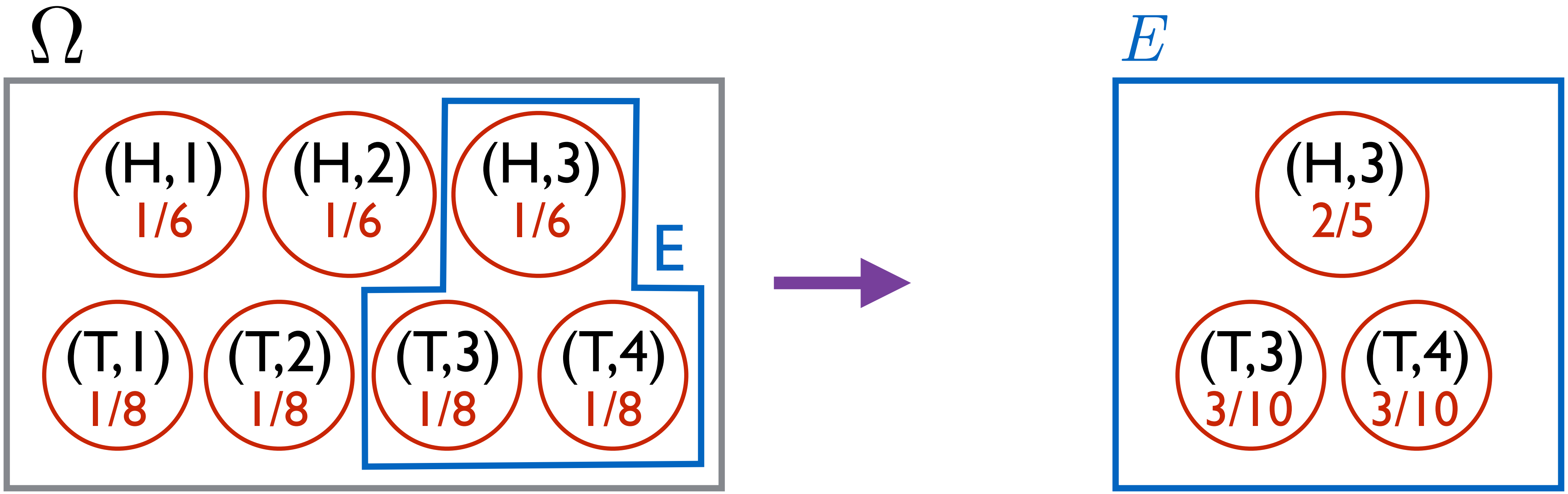
Outcomes:	(H,1)	(H,2)	(H,3)	(T,1)	(T,2)	(T,3)	(T,4)
Prob:	1/6	1/6	1/6	1/8	1/8	1/8	1/8
Prob | E :	0	0	2/5	0	0	3/10	3/10

E = die roll is 3 or higher

A = Tails was flipped

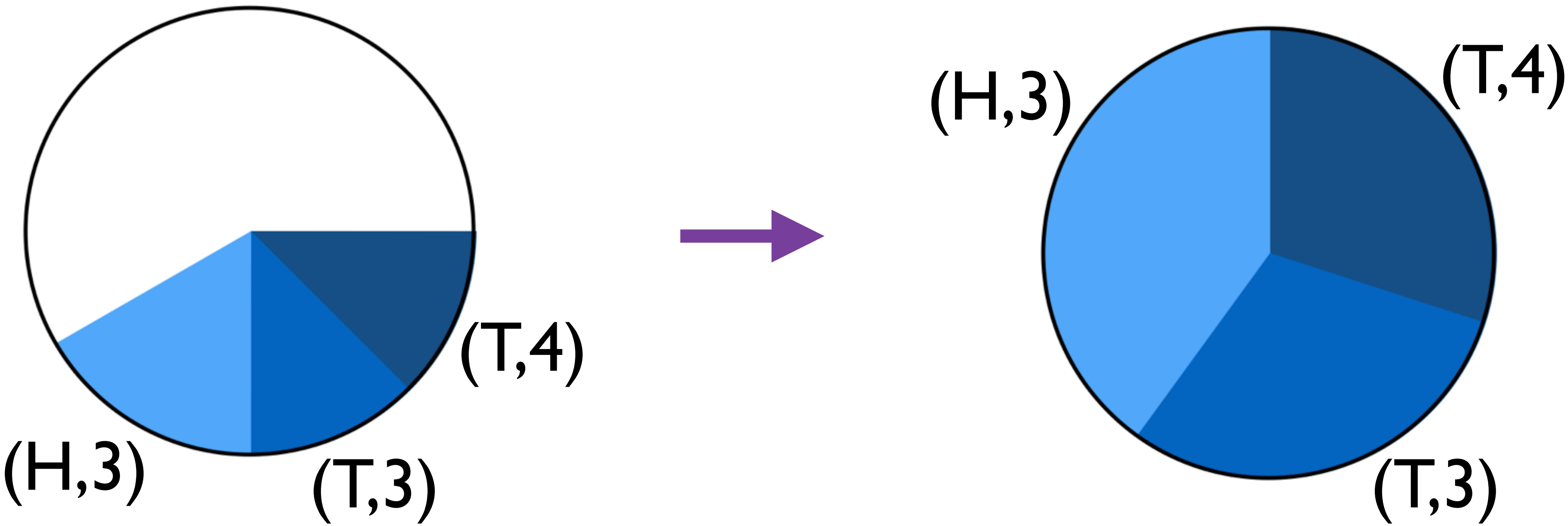
$\Pr[A \mid E] = 3/5$

Conditioning



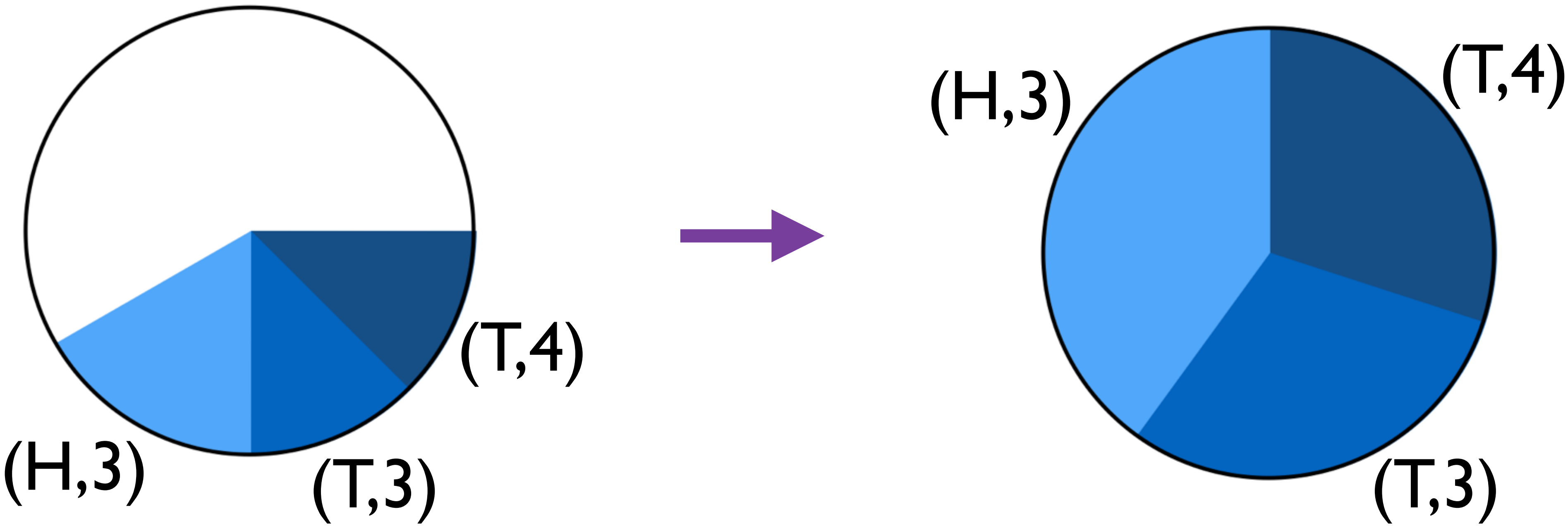
$$\Pr : \Omega \to [0,1]$$

$$\Pr_E : E \to [0,1]$$



Conditioning

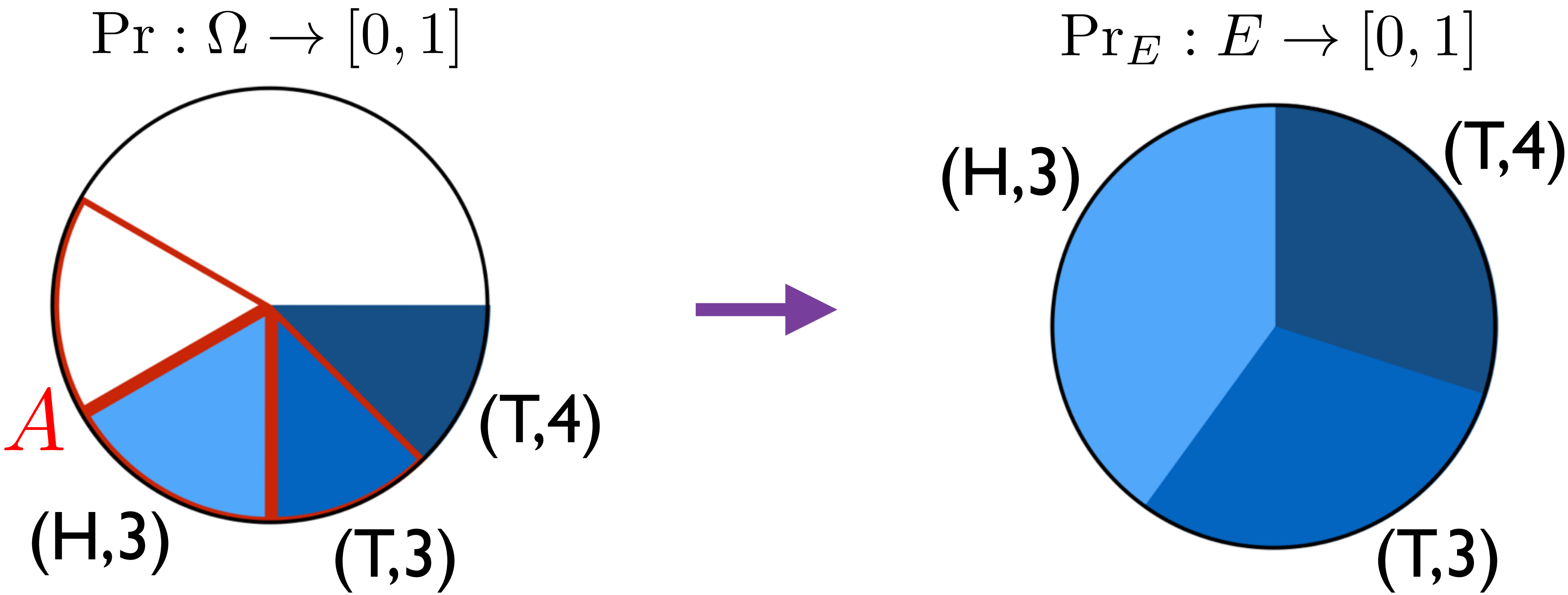
$\Pr : \Omega \to [0,1]$ $\qquad\qquad$ $\Pr_E : E \to [0,1]$



$$\Pr[\ell \mid E] \overset{\text{def}}{=} \Pr_E[\ell]$$

$$= \begin{cases} 0 & \text{if } \ell \notin E \\ \Pr[\ell]/\Pr[E] & \text{if } \ell \in E \end{cases}$$

Conditioning



$$\Pr[A \mid E] = \frac{\Pr[A \cap E]}{\Pr[E]}$$

(cannot condition on an event with prob. 0)

Conditional Probability —> Chain Rule

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B \mid A]$$

"For A and B to occur:

- first A must occur (probability $\Pr[A]$)
- then B must occur given that A occured (probability $\Pr[B \mid A]$)."

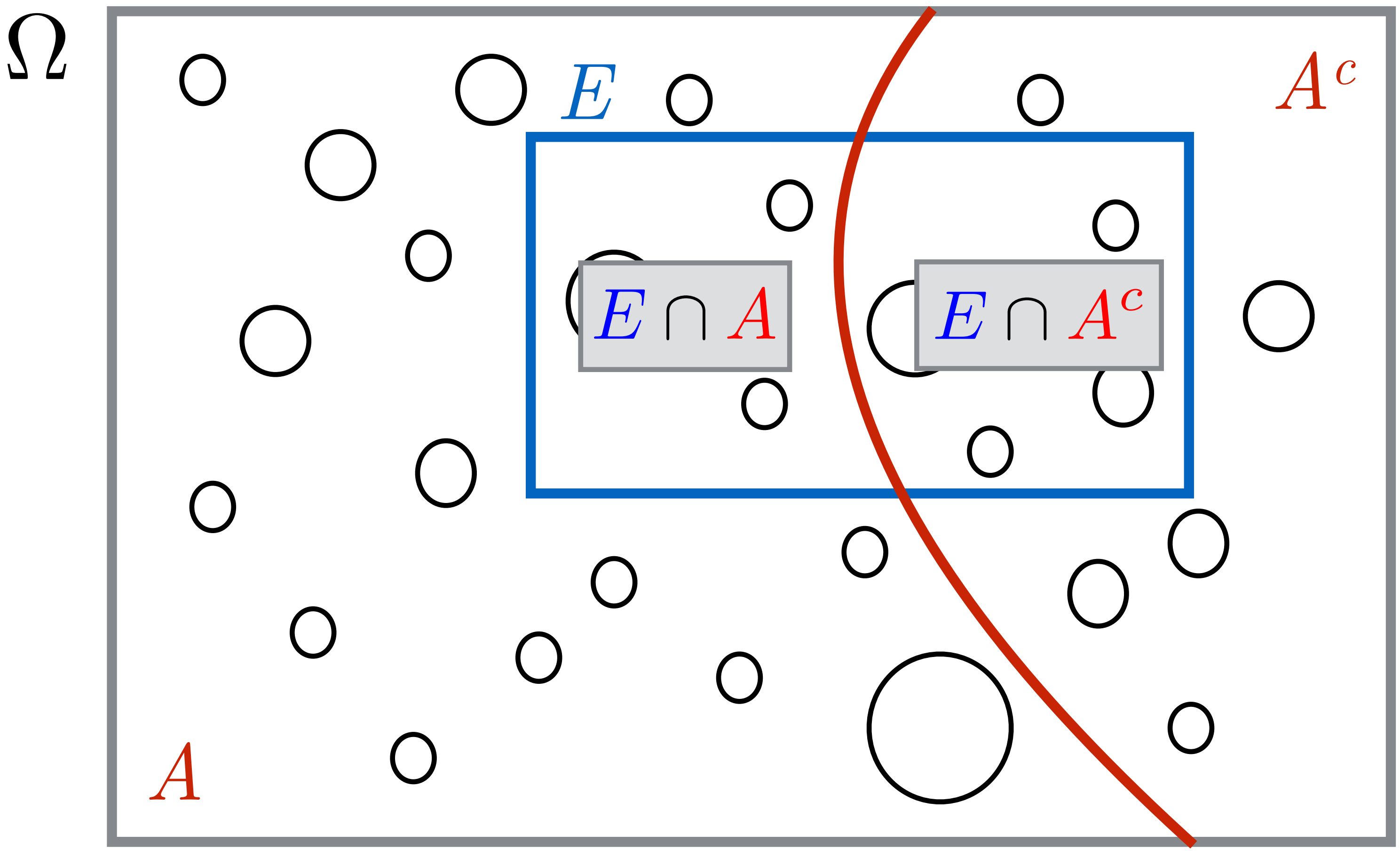
Generalizes to more than two events.

e.g.

$$\Pr[A \cap B \cap C] = \Pr[A] \cdot \Pr[B \mid A] \cdot \Pr[C \mid A \cap B]$$

Conditional Probability —> LTP

LTP = Law of Total Probability



$$\begin{aligned}\Pr[E] &= \Pr[E \cap A] + \Pr[E \cap A^c] \\ &= \Pr[A] \cdot \Pr[E \mid A] + \Pr[A^c] \cdot \Pr[E \mid A^c]\end{aligned}$$

Conditional Probability —> LTP

LTP = Law of Total Probability

If $A_1, A_2, \ldots, A_n$ partition Ω, then

$$
\begin{aligned}
\Pr[E] = {} & \Pr[A_1] \cdot \Pr[E \mid A_1] + \\
& \Pr[A_2] \cdot \Pr[E \mid A_2] + \\
& \ldots \\
& \Pr[A_n] \cdot \Pr[E \mid A_n].
\end{aligned}
$$

Conditional Probability —> Independence

Two events A and B are independent if

$$\Pr[A \mid B] = \Pr[A].$$

This is equivalent to:

$$\Pr[B \mid A] = \Pr[B].$$

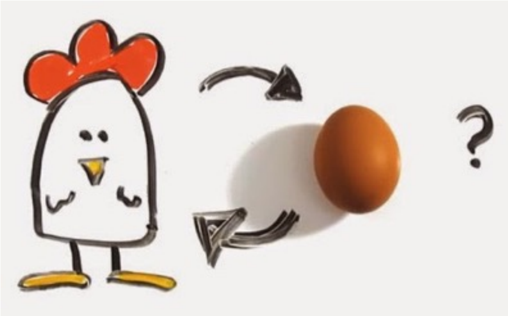
This is equivalent to:

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$$

(except that this equality can be used even when $\Pr[A] = 0,$ or $\Pr[B] = 0$.)

So this is actually used for the definition of independence.

Problem with Independence Definition



Want to calculate $\Pr[A \cap B]$.

If they are independent, we can use $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.
(but we need to show this equality to show independence)

Argue independence by informally arguing:
if B happens, this cannot affect the probability of A happening.

Then use $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.

Problem with Independence Definition

Real World → Mathematical Model

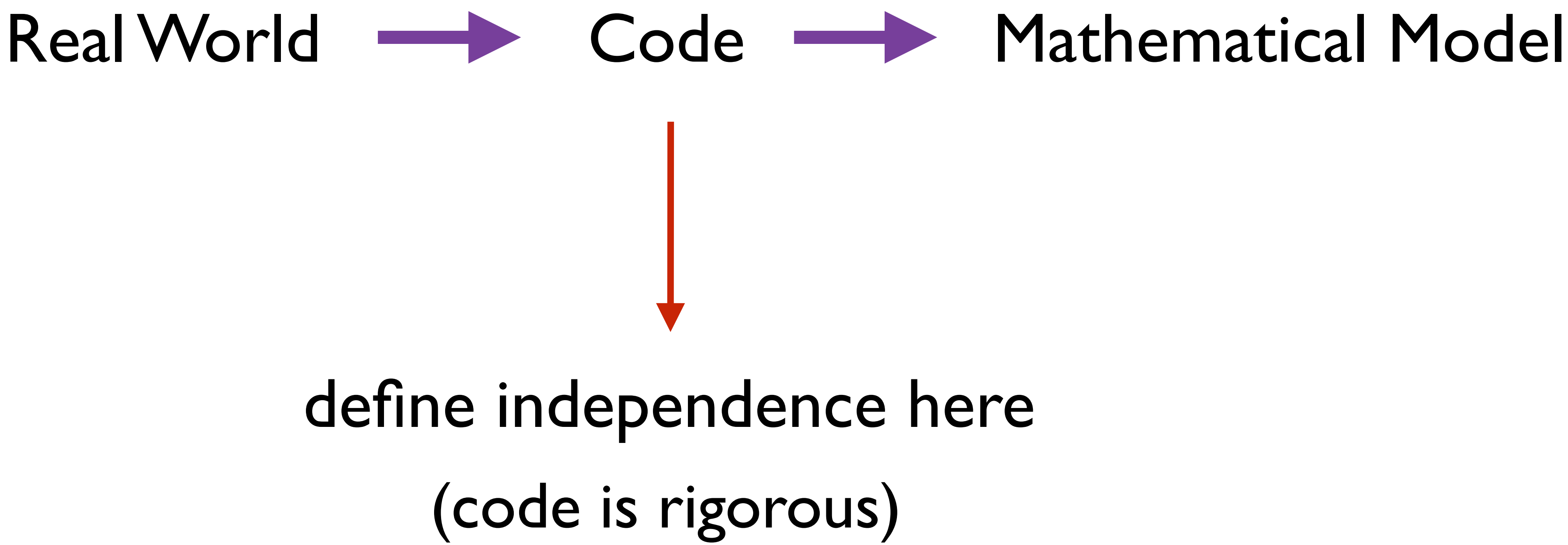
some notion of
independence
of A and B

(the secret definition
of independence)

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$$

<u>problem</u>: real-world description not always very rigorous.

Fixing the Problem



Fixing the Problem

<u>Randomized code:</u>

part 1

⋮

part 2

Suppose A is an event that depends only on part 1.

Suppose B is an event that depends only on part 2.

Suppose you prove two parts cannot affect each other.
(i.e., could run them in opposite order.)

Then A and B are independent.

You may conclude $\Pr[A \mid B] = \Pr[A]$.

Independence of More Events

Events $A_1, A_2, \ldots, A_n$ are independent if

for every $S \subseteq \{1, 2, \ldots, n\}$:

$$\Pr\left[\bigcap_{i \in S} A_i\right] = \prod_{i \in S} \Pr[A_i].$$

We can define it also in the "Code World"
(with n blocks of code that don't affect each other).

<u>**Consequence**</u>: anything like

$$\Pr[A_1 \mid (A_2 \cup A_3) \cap (A_4^c \cup A_5)] = \Pr[A_1]$$

SUMMARY SO FAR

Real World → Code → Probability Tree

=

Mathematical Model

- set of outcomes Ω
- a prob. associated with each outcome.

Events

Conditional probability:

$$\Pr[A \mid B] = \Pr[A \cap B] / \Pr[B]$$

Chain rule:

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B \mid A]$$

Law of total probability:

$$\Pr[B] = \Pr[A] \cdot \Pr[B \mid A] + \Pr[A^c] \cdot \Pr[B \mid A^c]$$

Independent events:

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$$

Union bound:

$$\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$$

Next Time:

Random Variables and

Introduction to Randomized Algorithms