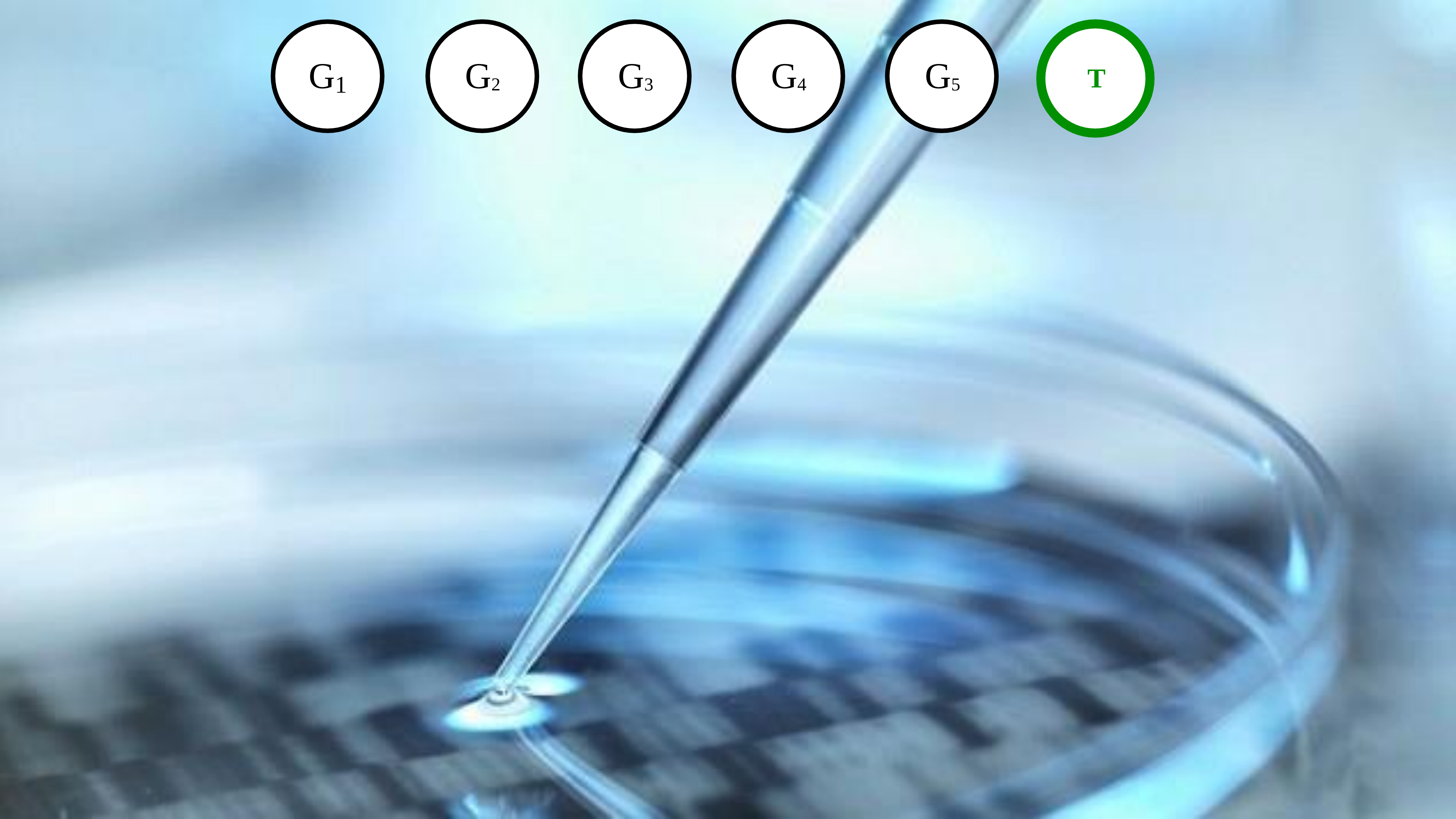
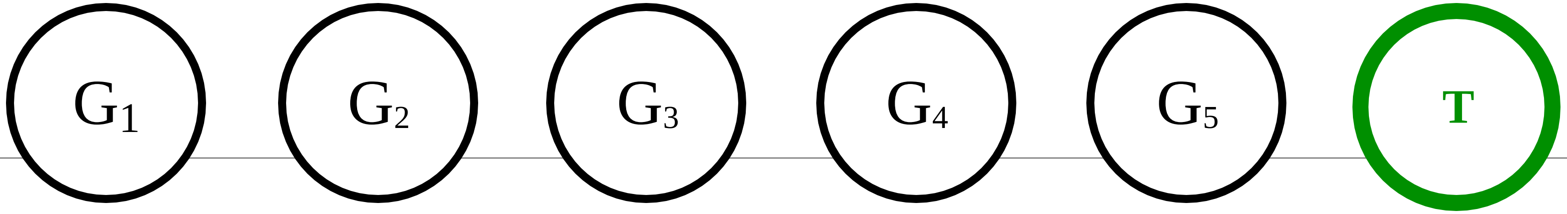
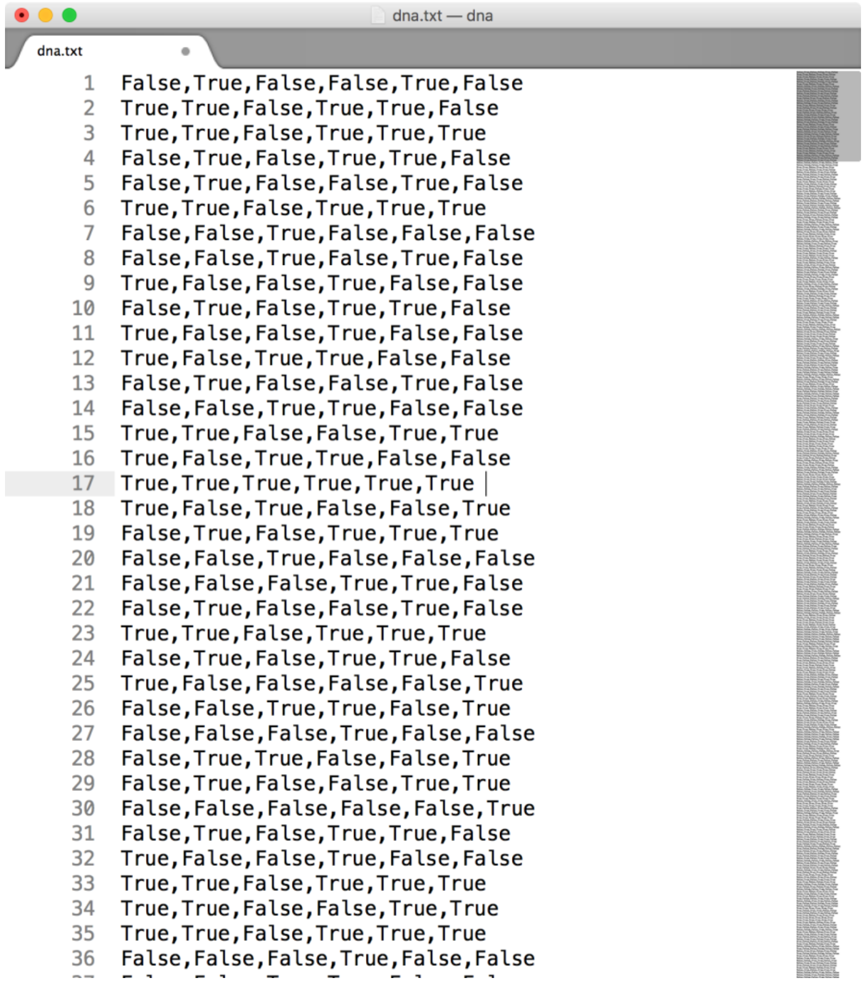
Independence
Chris Piech, CS109

Today, start with a cool program

G_1 G_2 G_3 G_4 G_5 T





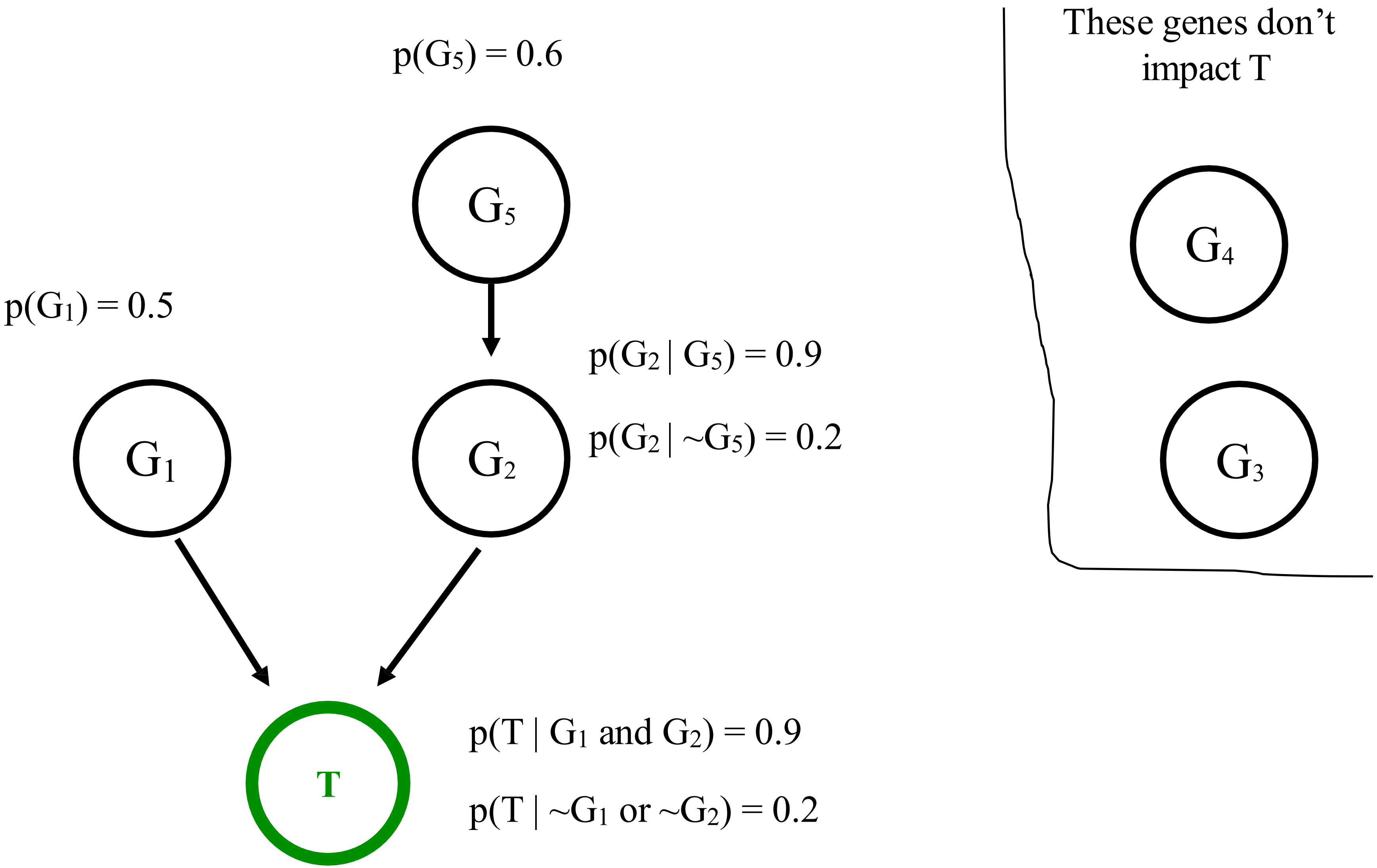


```
False,True,False,False,True,False
True,True,False,True,True,False
True,True,False,True,True,True
False,True,False,True,True,False
False,True,False,False,True,False
True,True,False,True,True,True
False,False,True,False,False,False
False,False,True,False,True,False
True,False,False,True,False,False
False,True,False,True,True,False
True,False,False,True,False,False
True,False,True,True,False,False
False,True,False,False,True,False
False,False,True,True,False,False
True,True,False,False,True,True
True,False,True,True,False,False
True,True,True,True,True,True
True,False,True,False,False,True
False,True,False,True,True,True
False,False,True,False,False,False
False,False,False,True,True,False
False,True,False,False,True,False
True,True,False,True,True,True
False,True,False,True,True,False
True,False,False,False,False,True
False,False,True,True,False,True
False,False,False,True,False,False
False,True,True,False,False,True
False,True,False,False,True,True
False,False,False,False,False,True
False,True,False,True,True,False
True,False,False,True,False,False
True,True,False,True,True,True
True,True,False,False,True,True
True,True,False,True,True,True
False,False,False,True,False,False
```

100,000 samples

6 observations per sample

Discovered Hypothesis



$p(G_5) = 0.6$

$p(G_1) = 0.5$

$p(G_2 \mid G_5) = 0.9$

$p(G_2 \mid \sim G_5) = 0.2$

$p(T \mid G_1 \text{ and } G_2) = 0.9$

$p(T \mid \sim G_1 \text{ or } \sim G_2) = 0.2$

These genes don’t impact T

We’ve Gotten Ahead of Ourselves

Source: The Hobbit

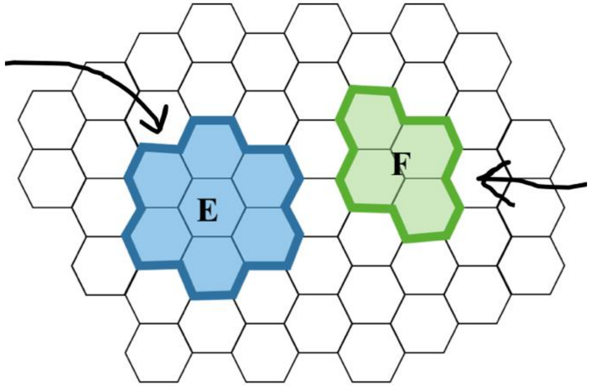
Start at the Beginning

Source: The Hobbit

Announcements

- **PSet #1** is due today 10p. Recall grace period.
- PSet #2 goes out today!

Learning Goals of Today



Mutually Exclusive

$$P(A \text{ and } B) = 0$$

Makes ***OR*** easy:

$$P(A \text{ or } B) = P(A) + P(B)$$

Independent

$$P(A) = P(A|B)$$

Makes ***AND*** easy:

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

Review

Notation

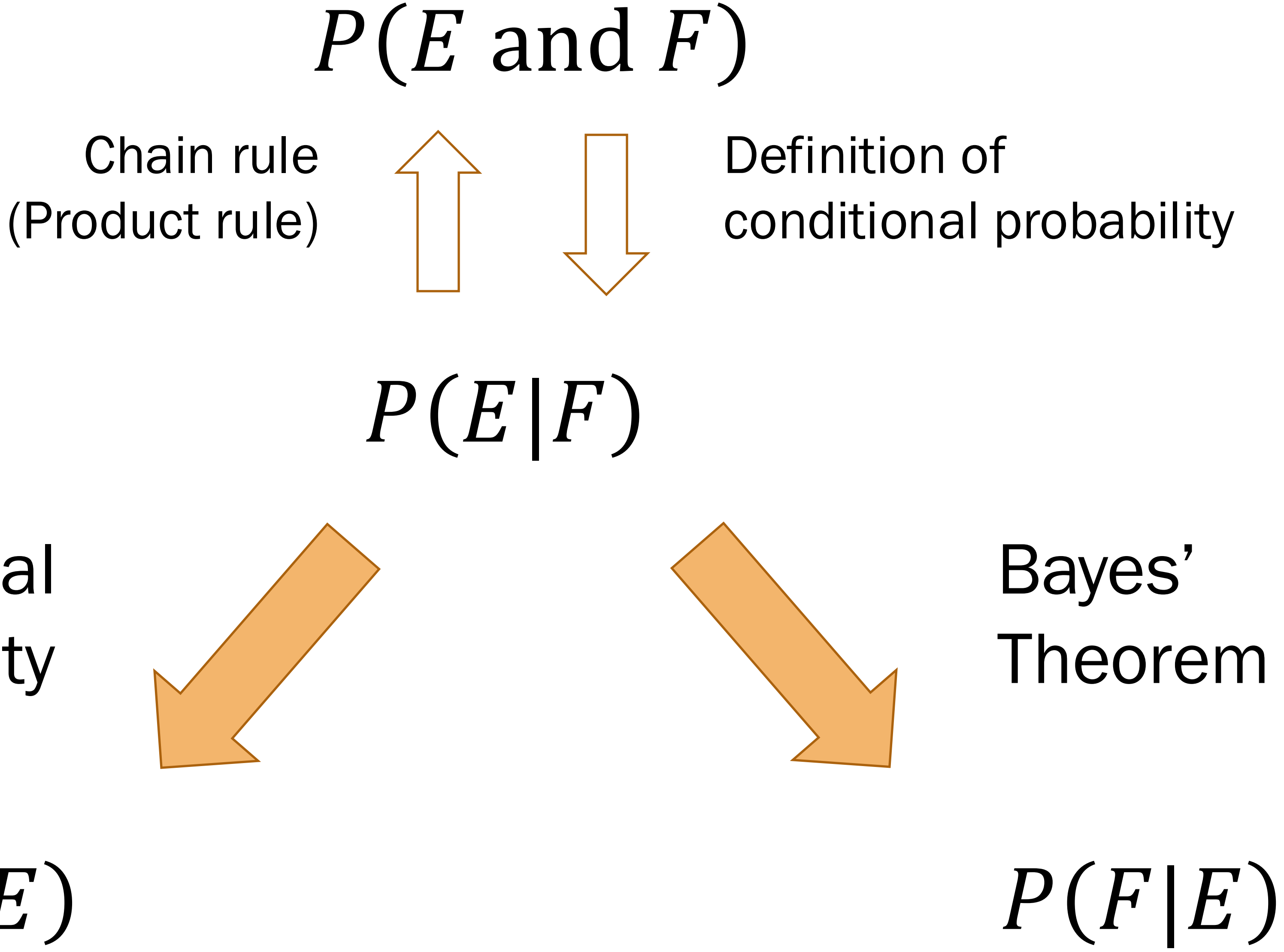
And	**Or**	**Given**
$P(E \text{ and } F)$	$P(E \text{ or } F)$	$P(E\vert F)$
$P(E, F)$	$P(E \cup F)$	
$P(EF)$		
$P(E \cap F)$		

Review: Conditional Probability

$p(AB)$ vs $P(A|B)$

$$P(AB) = P(A|B) \cdot \underline{P(B)}$$

Relationship Between Probabilities



$P(E \text{ and } F)$

Chain rule
(Product rule)

Definition of
conditional probability

$P(E|F)$

Law of Total
Probability

Bayes'
Theorem

$P(E)$

$P(F|E)$

Review: Chain Rule

Definition of conditional probability:

$$P(E|F) = \frac{P(EF)}{P(F)}$$

The Chain Rule:

$$P(EF) = P(E|F)P(F)$$

Bayes' Theorem

$P(E|F) \Rightarrow P(F|E)$

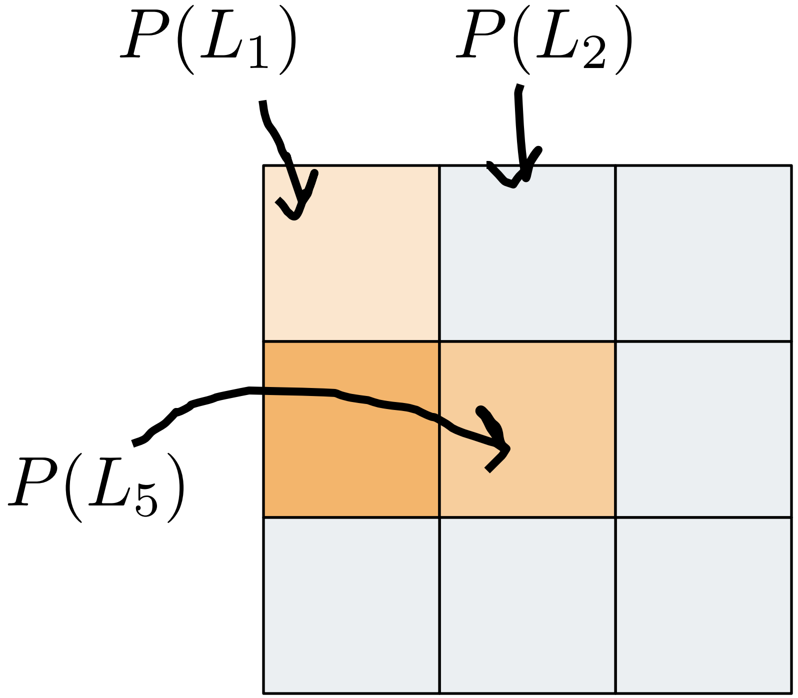
Thm For any events E and F where $P(E) > 0$ and $P(F) > 0$,

$$\boxed{P(F|E) = \frac{P(E|F)P(F)}{P(E)}}$$

Expanded form:

$$\boxed{P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}}$$

Bayes' Theorem and Location

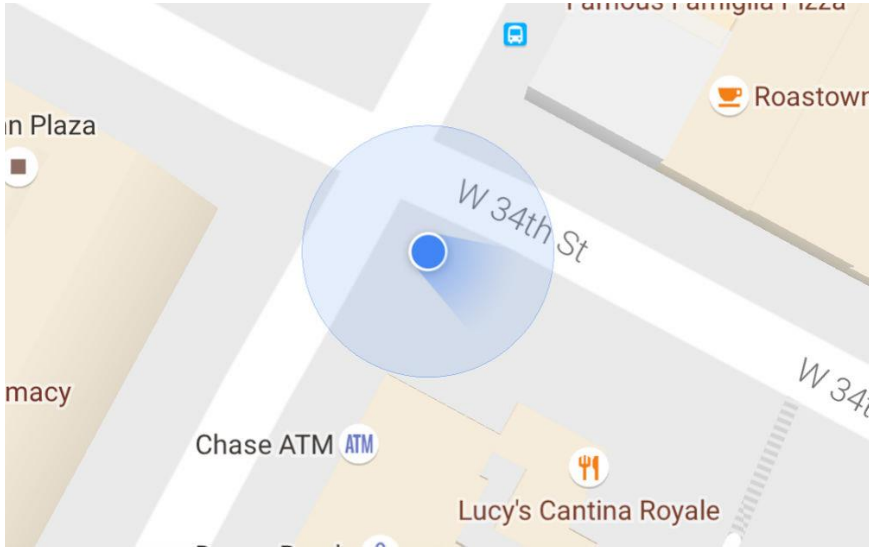


$P(L_1)$

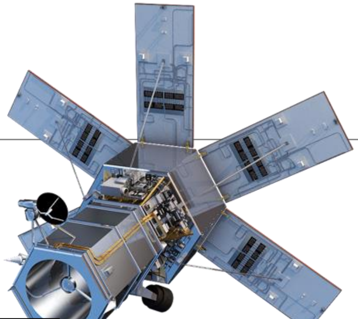
$P(L_2)$

$P(L_5)$

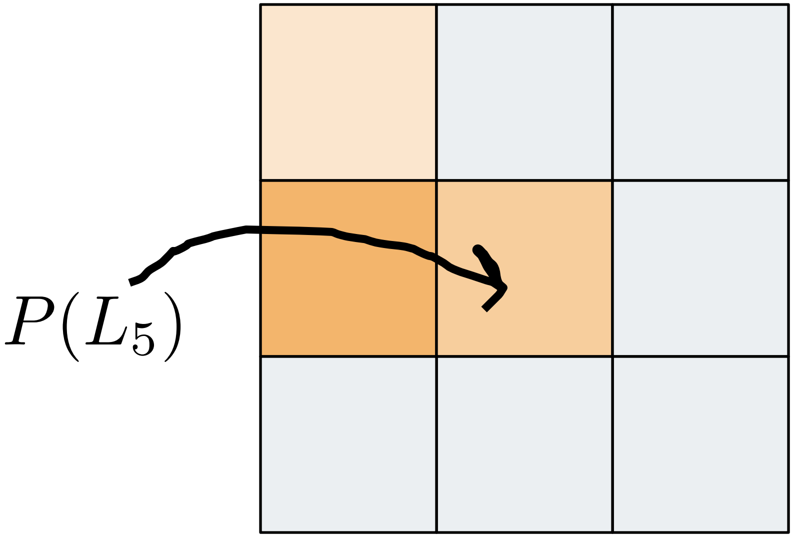
Before Observation



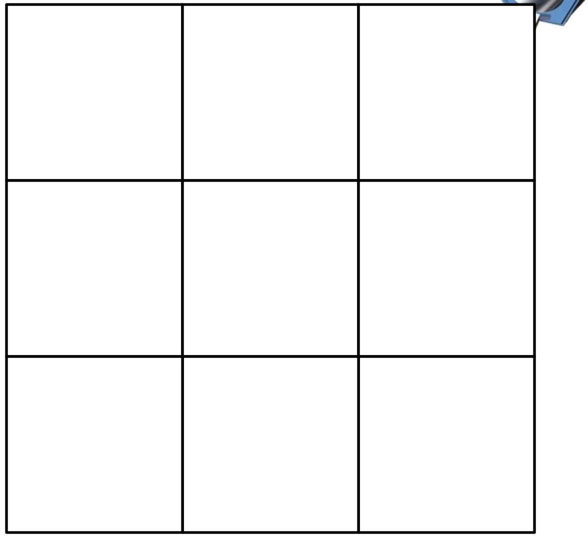
Bayes' Theorem and Location



Know: $P(O|L_i)$

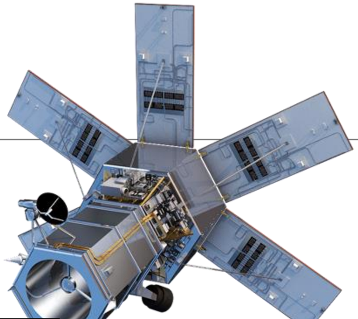


Before Observation

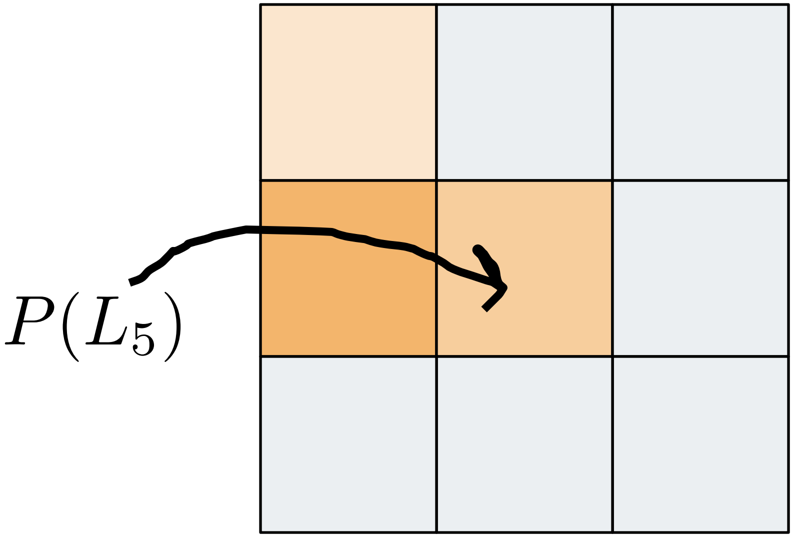


After Observation

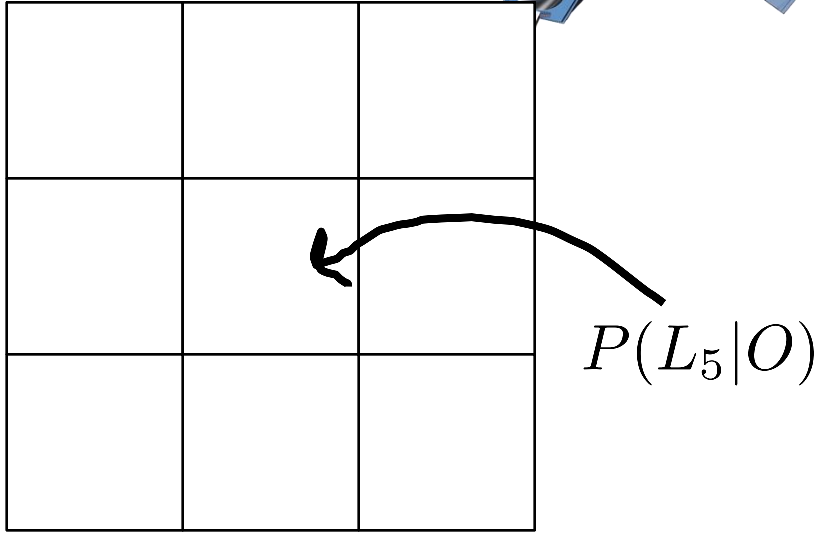
Bayes' Theorem and Location



Know: $P(O|L_i)$



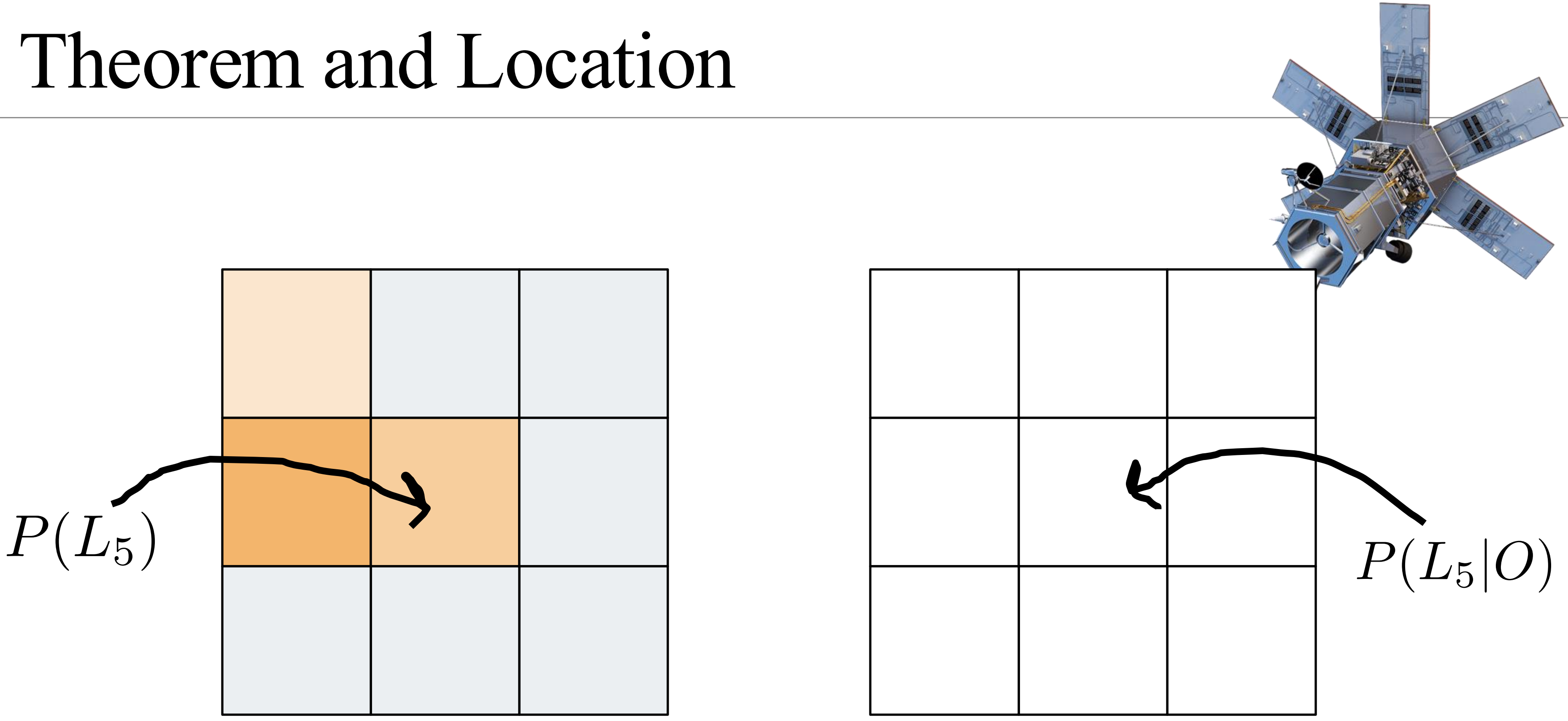
Before Observation



After Observation

$$P(L_5|O) = \frac{P(O|L_5)P(L_5)}{P(O)}$$

Bayes' Theorem and Location

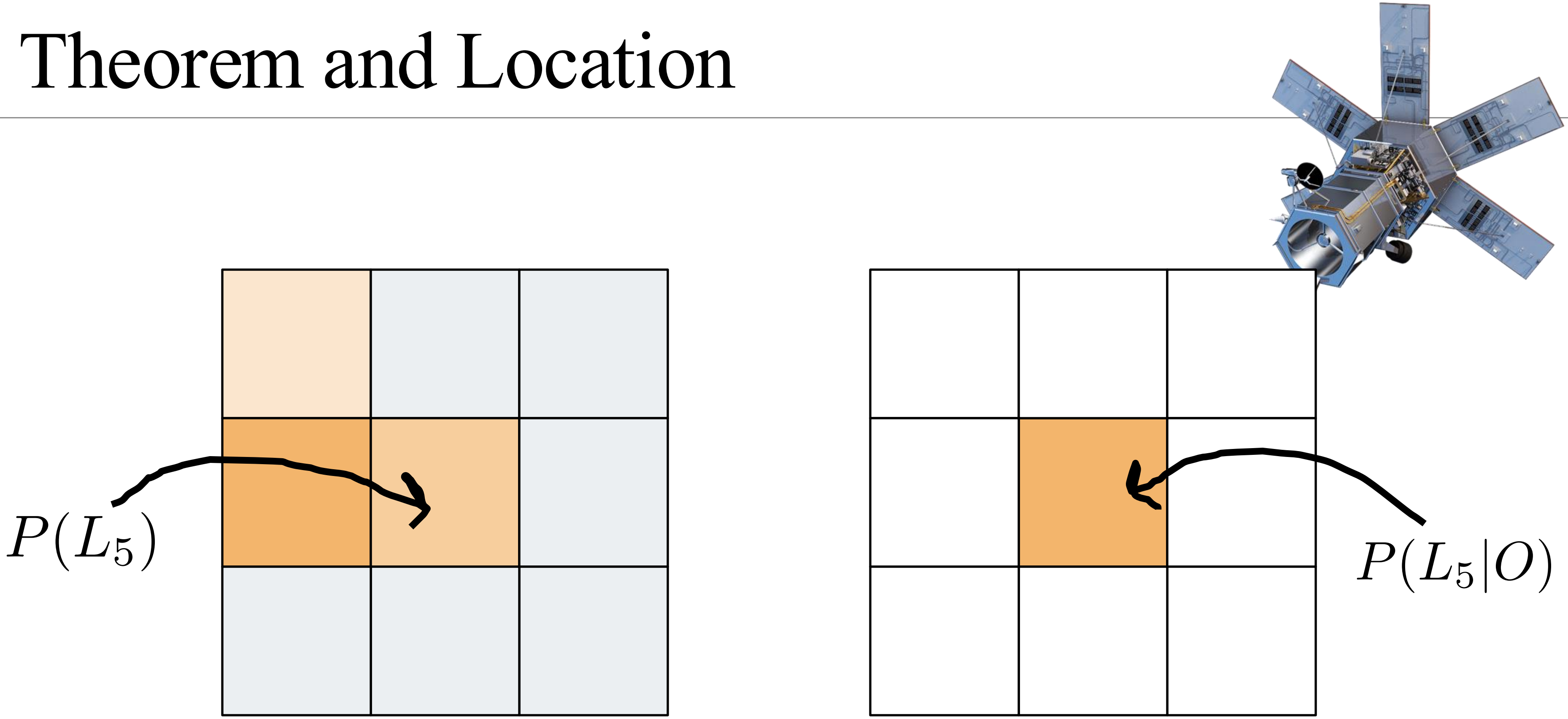


Before Observation

After Observation

$$P(L_5|O) = \frac{P(O|L_5)P(L_5)}{\sum_i P(O|L_i)P(L_i)}$$

Bayes' Theorem and Location

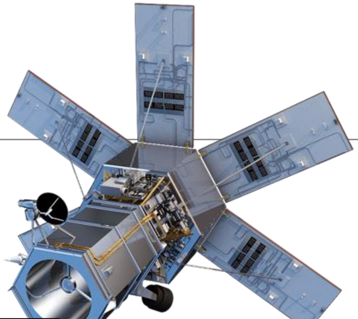


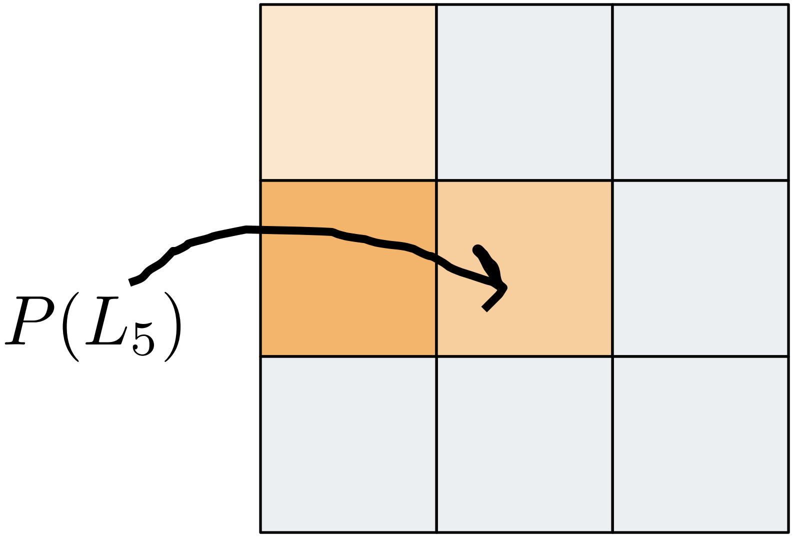
Before Observation

After Observation

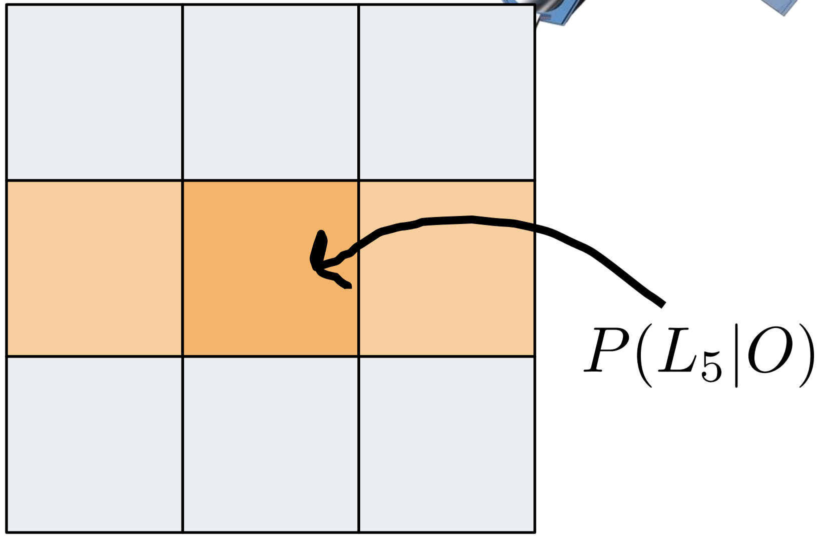
$$P(L_5|O) = \frac{P(O|L_5)P(L_5)}{\sum_i P(O|L_i)P(L_i)}$$

Bayes' Theorem and Location





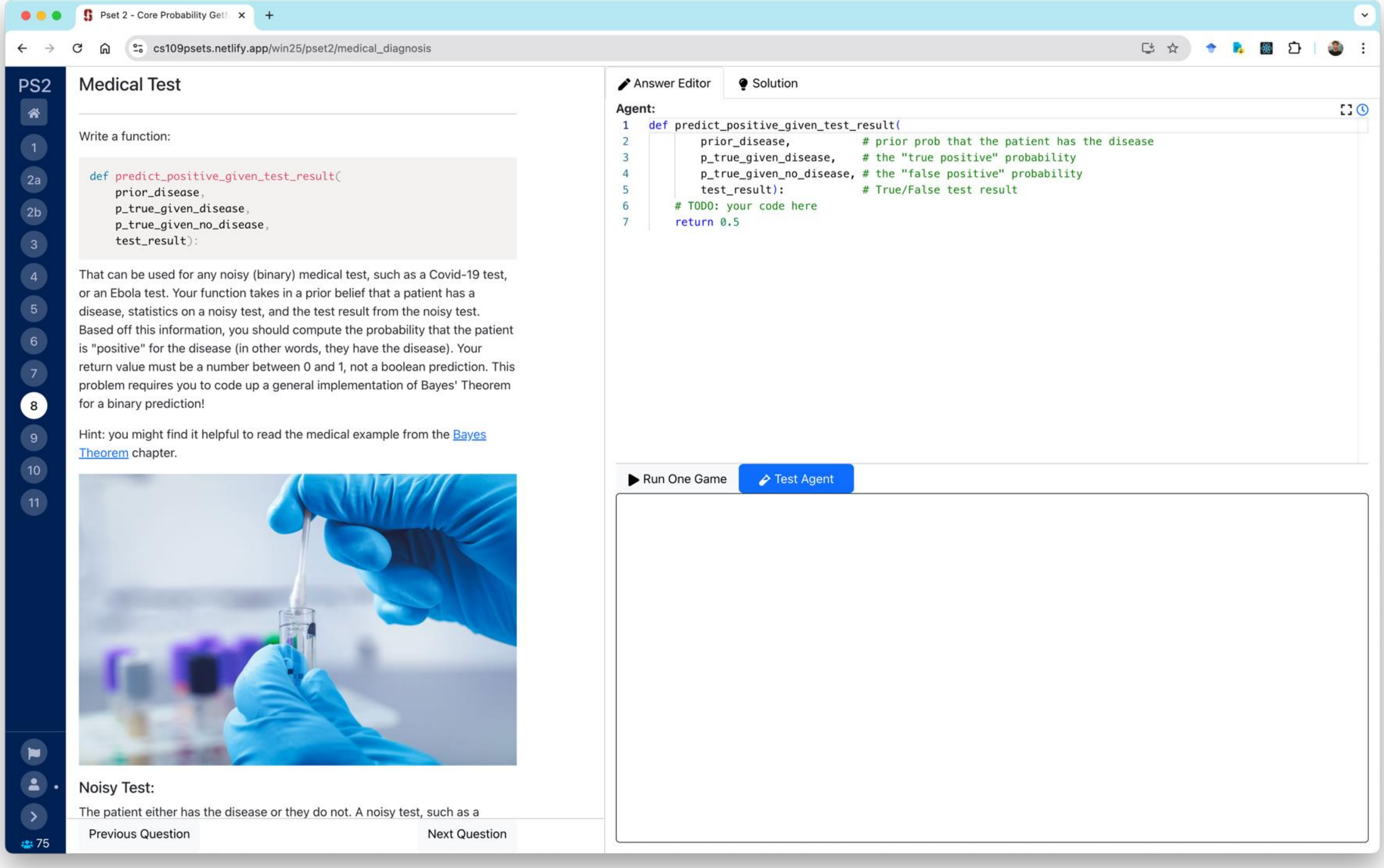
Before Observation



After Observation

$$P(L_5|O) = \frac{P(O|L_5)P(L_5)}{\sum_i P(O|L_i)P(L_i)}$$

ProTip: Try this on PSet 2



Medical Test

Write a function:

```
def predict_positive_given_test_result(
    prior_disease,
    p_true_given_disease,
    p_true_given_no_disease,
    test_result):
```

That can be used for any noisy (binary) medical test, such as a Covid-19 test, or an Ebola test. Your function takes in a prior belief that a patient has a disease, statistics on a noisy test, and the test result from the noisy test. Based off this information, you should compute the probability that the patient is "positive" for the disease (in other words, they have the disease). Your return value must be a number between 0 and 1, not a boolean prediction. This problem requires you to code up a general implementation of Bayes' Theorem for a binary prediction!

Hint: you might find it helpful to read the medical example from the Bayes Theorem chapter.

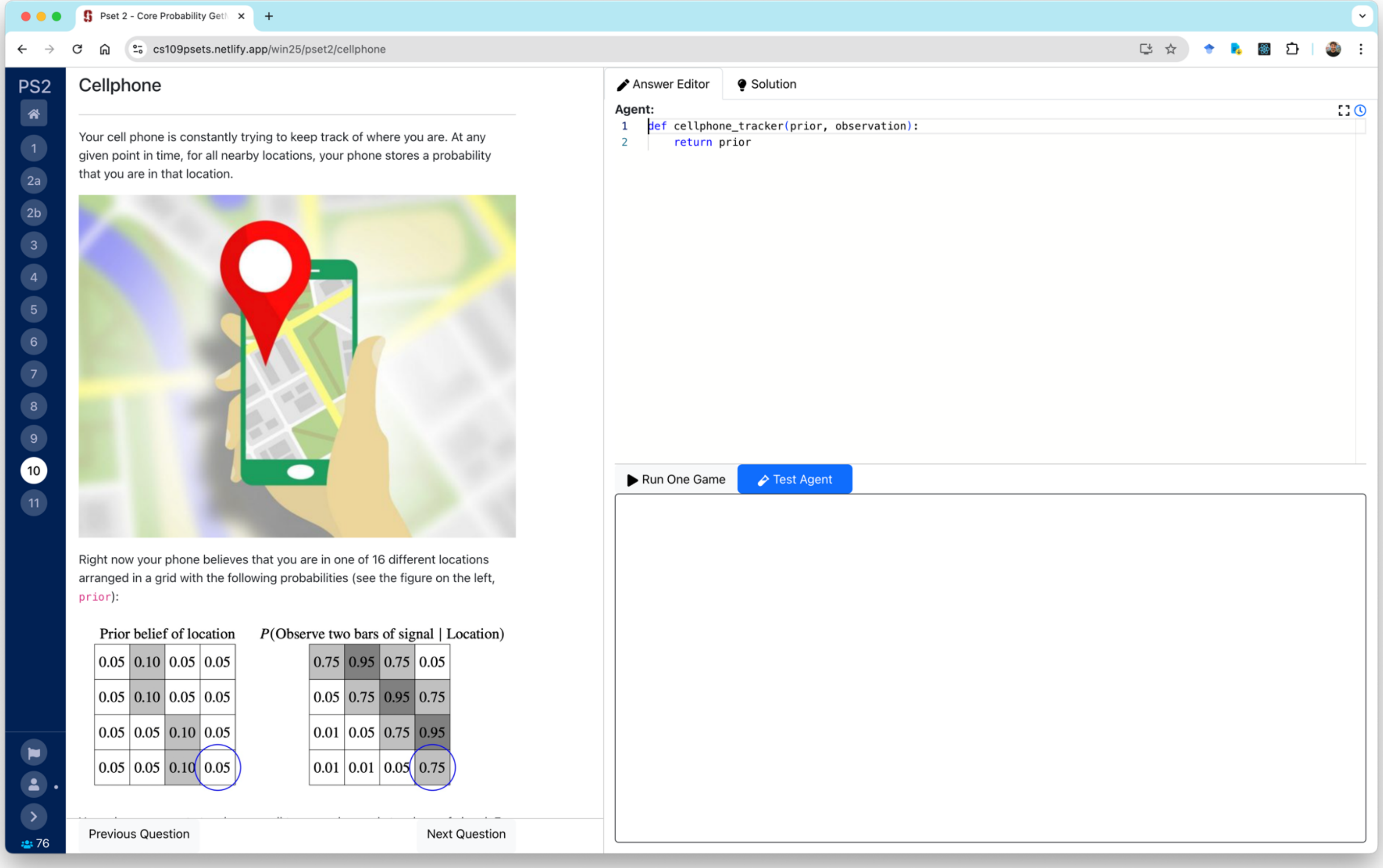
Noisy Test:

The patient either has the disease or they do not. A noisy test, such as a

Agent:

```
def predict_positive_given_test_result(
        prior_disease,           # prior prob that the patient has the disease
        p_true_given_disease,    # the "true positive" probability
        p_true_given_no_disease, # the "false positive" probability
        test_result):            # True/False test result
    # TODO: your code here
    return 0.5
```

ProTip: Try this on PSet 2



Cellphone

Your cell phone is constantly trying to keep track of where you are. At any given point in time, for all nearby locations, your phone stores a probability that you are in that location.

Right now your phone believes that you are in one of 16 different locations arranged in a grid with the following probabilities (see the figure on the left, `prior`):

Prior belief of location

0.05	0.10	0.05	0.05
0.05	0.10	0.05	0.05
0.05	0.05	0.10	0.05
0.05	0.05	0.10	0.05

P(Observe two bars of signal | Location)

0.75	0.95	0.75	0.05
0.05	0.75	0.95	0.75
0.01	0.05	0.75	0.95
0.01	0.01	0.05	0.75

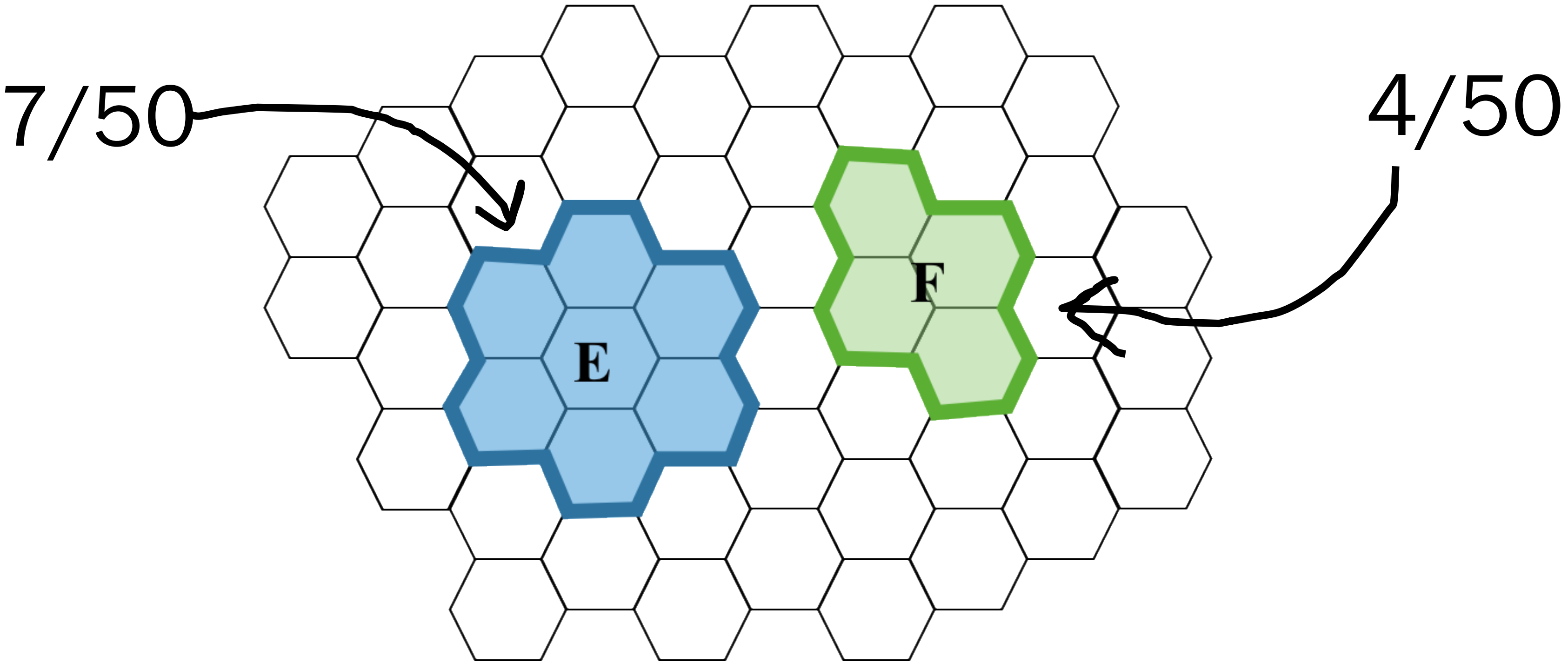
Agent:

```
def cellphone_tracker(prior, observation):
    return prior
```

End Review

Probability of “OR”

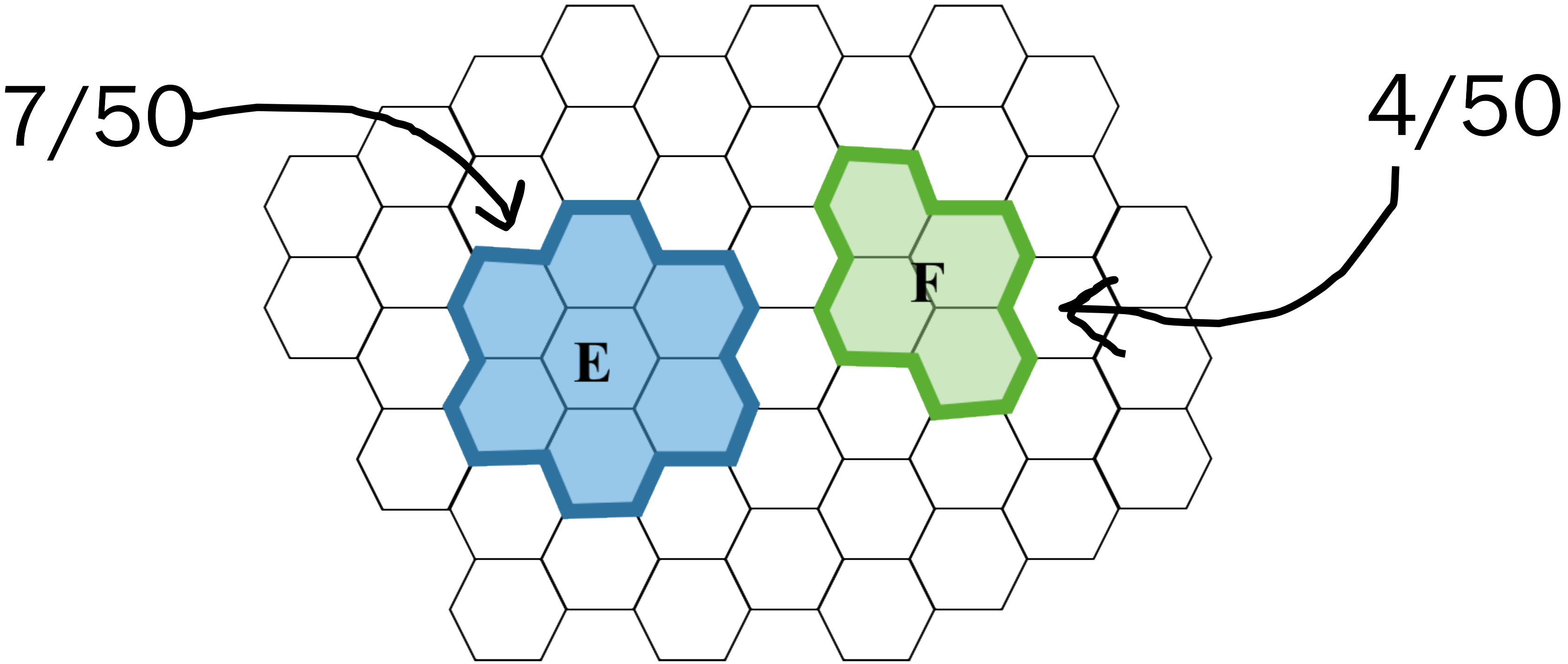
Review: OR with Mutually Exclusive Events



If events are mutually exclusive, probability of OR is simple:

$$P(E \text{ or } F) = P(E) + P(F)$$

Review: OR with Mutually Exclusive Events



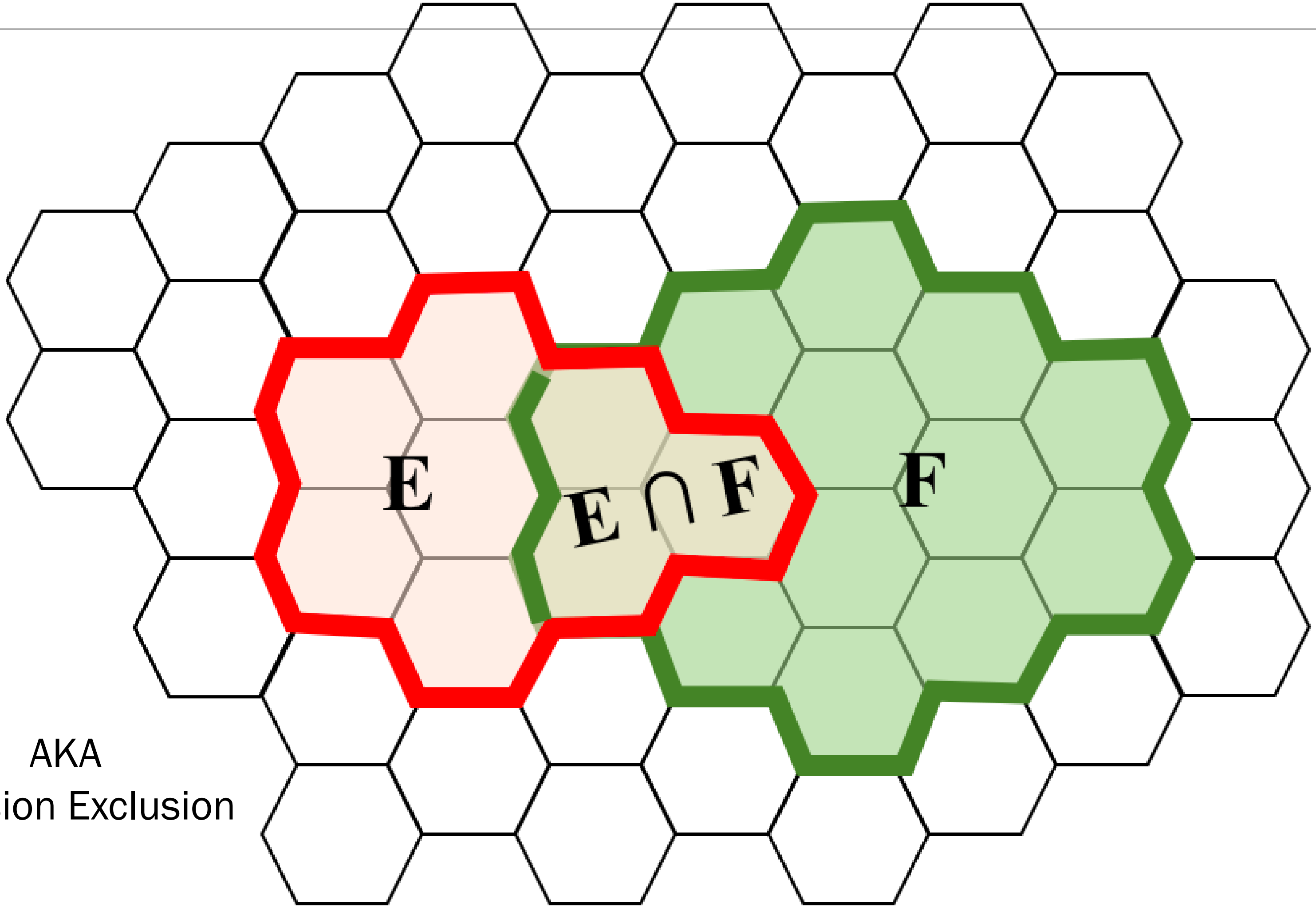
If events are mutually exclusive, probability of OR is simple:

$$P(E \text{ or } F) = \frac{7}{50} + \frac{4}{50} = \frac{11}{50}$$

What about when they are not

Mutually exclusive?

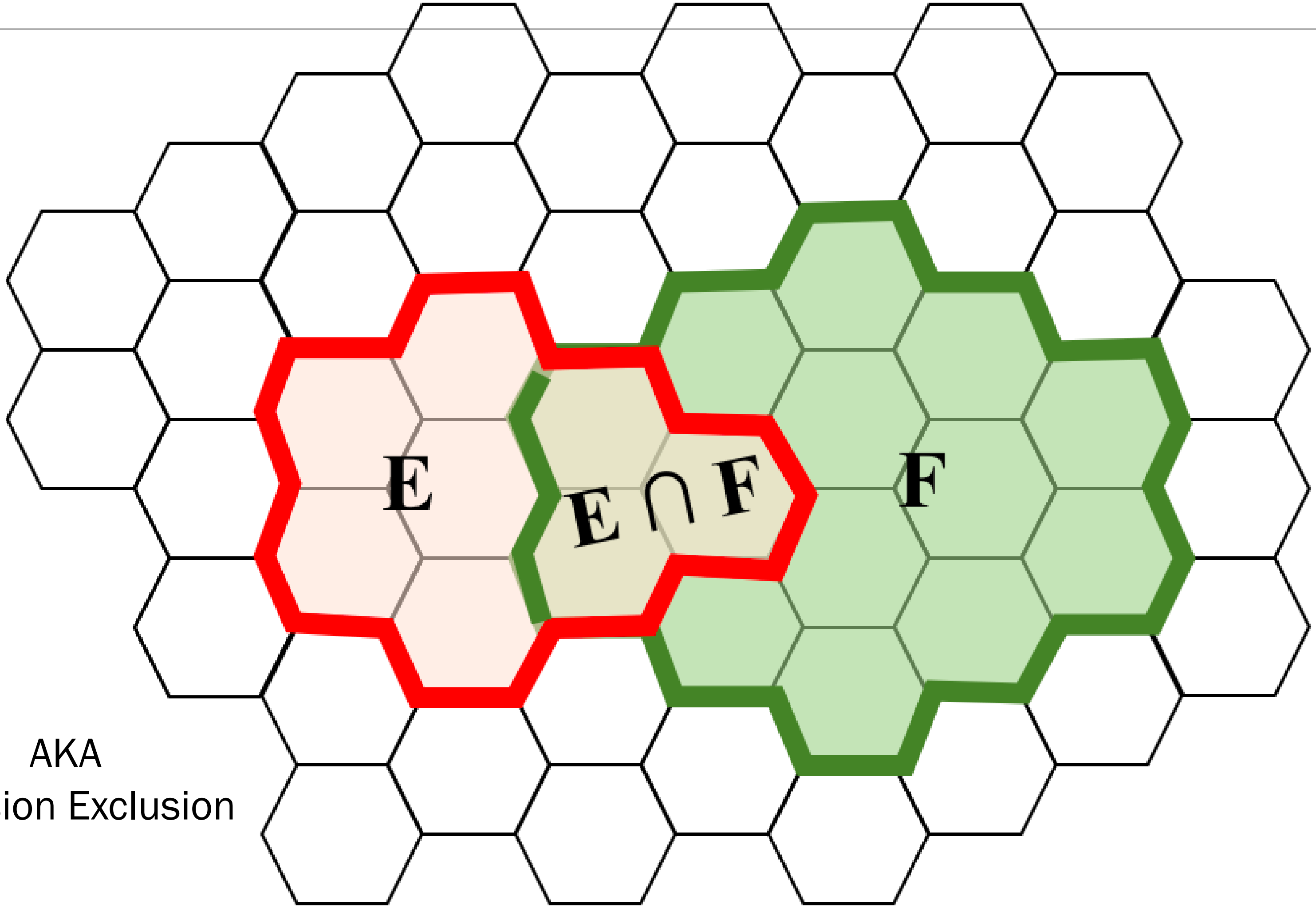
OR **without** Mutually Exclusive Events



AKA
Inclusion Exclusion

$$P(E \text{ or } F) = P(E) + P(F) - P(EF)$$

OR **without** Mutually Exclusive Events



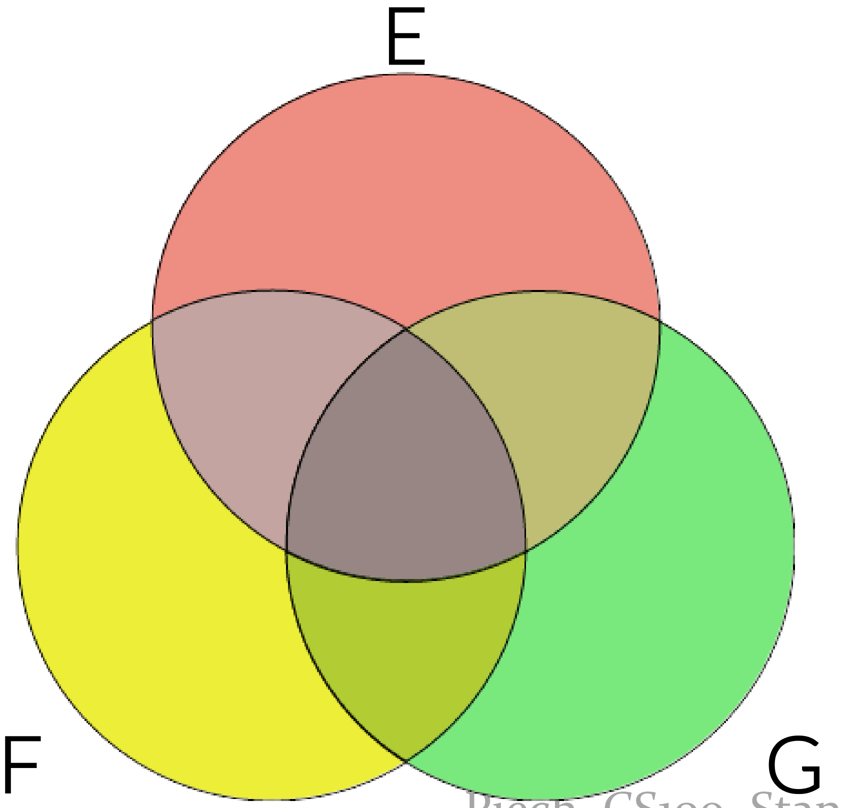
AKA
Inclusion Exclusion

$$P(E \text{ or } F) = \frac{8}{50} + \frac{14}{50} - \frac{3}{50}$$

More than two sets?

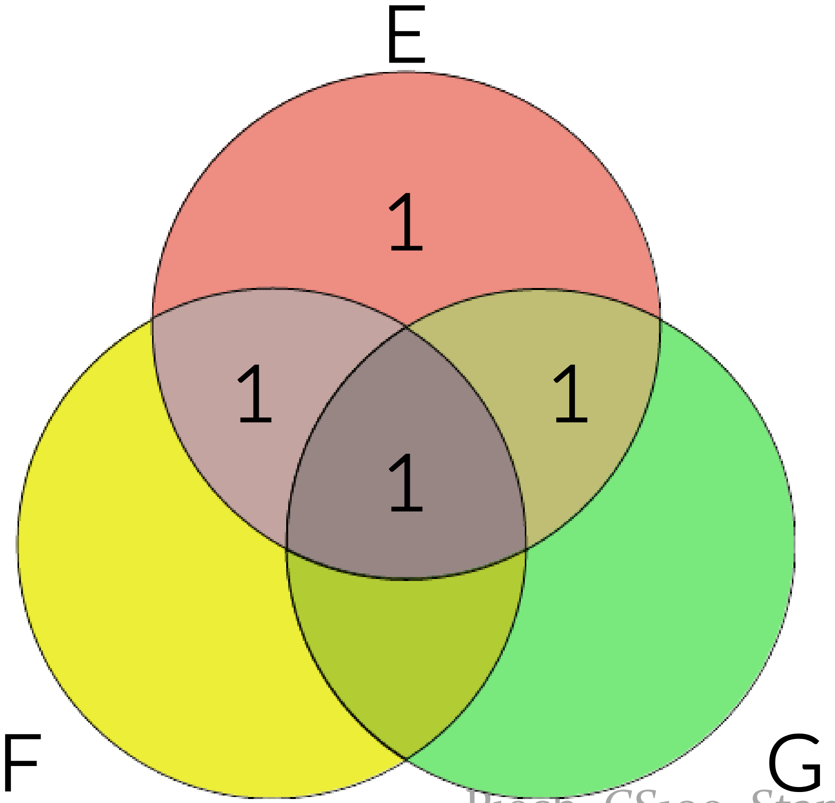
Inclusion / Exclusion with Three Events

$$P(E \text{ or } F \text{ or } G) =$$



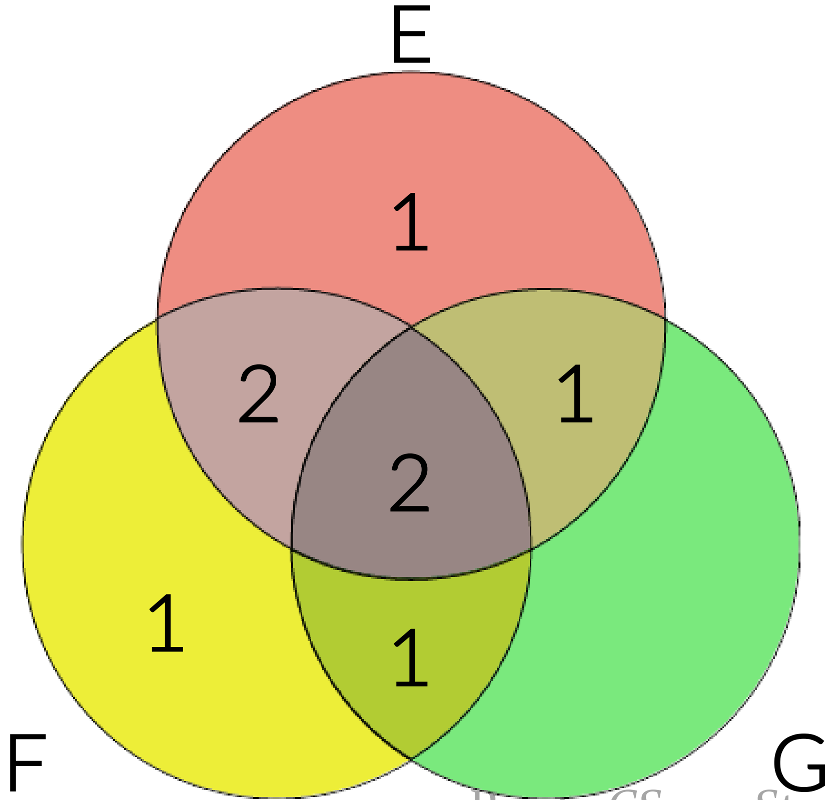
Inclusion / Exclusion with Three Events

$$P(E \text{ or } F \text{ or } G) = P(E)$$



Inclusion / Exclusion with Three Events

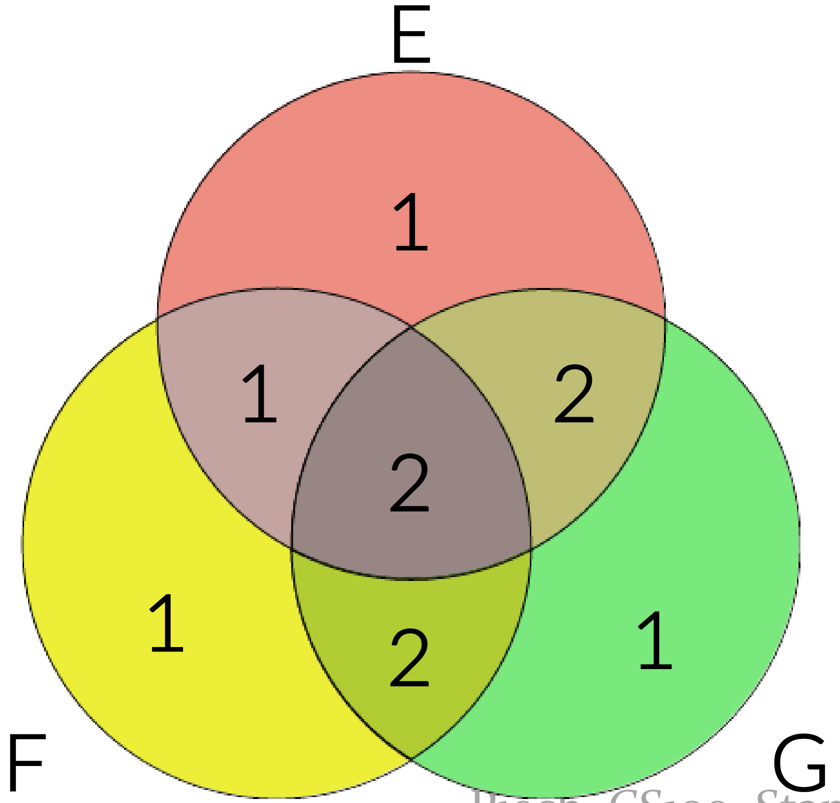
$$P(E \text{ or } F \text{ or } G) = P(E) + P(F)$$



Inclusion / Exclusion with Three Events

$$P(E \text{ or } F \text{ or } G) = P(E) + P(F) + P(G)$$

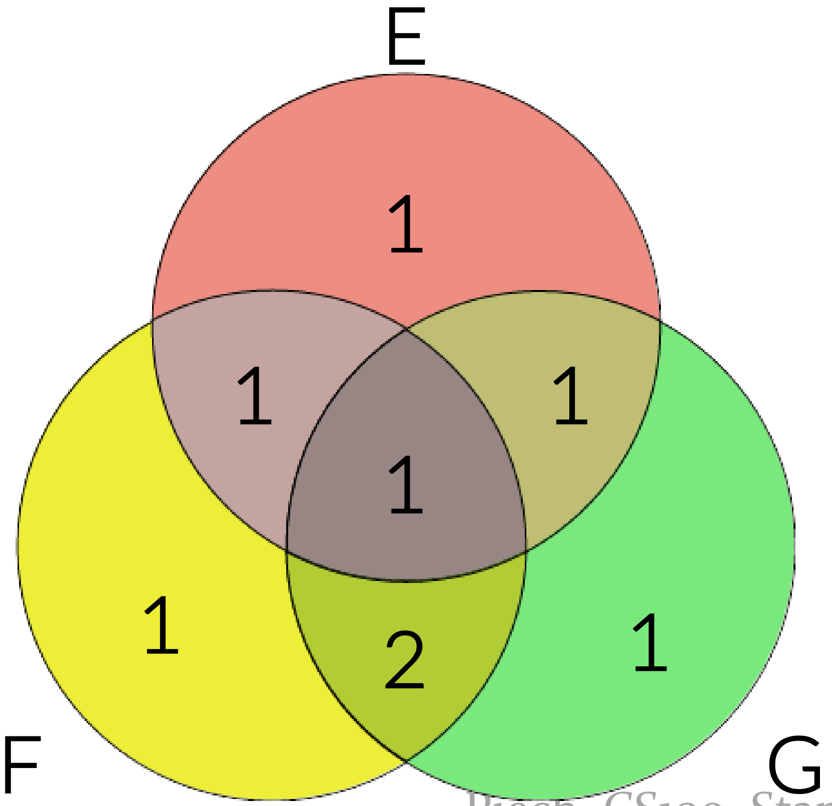
$$-P(EF)$$



Inclusion / Exclusion with Three Events

$$P(E \text{ or } F \text{ or } G) = P(E) + P(F) + P(G)$$

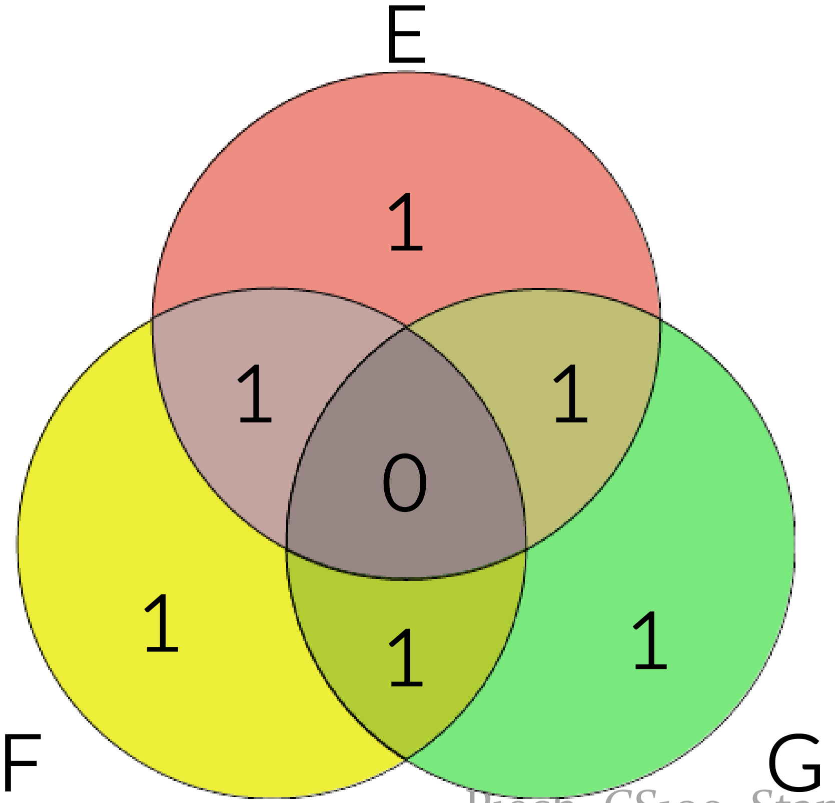
$$- P(EF) - P(EG)$$



Inclusion / Exclusion with Three Events

$$P(E \text{ or } F \text{ or } G) = P(E) + P(F) + P(G)$$

$$- P(EF) - P(EG) - P(FG)$$

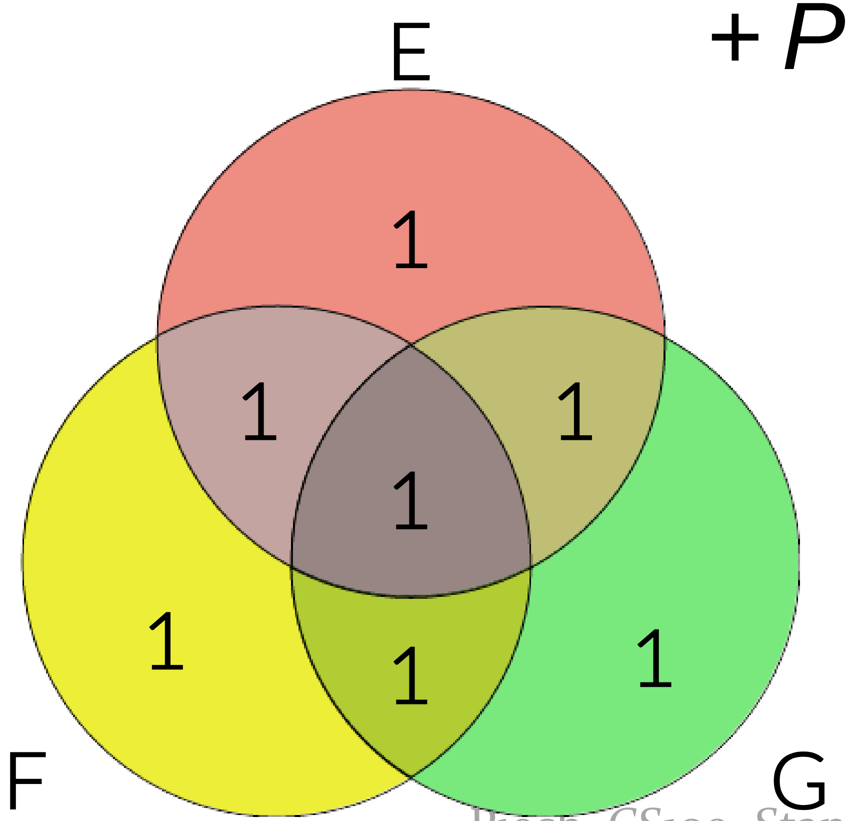


Inclusion / Exclusion with Three Events

$$P(E \text{ or } F \text{ or } G) = P(E) + P(F) + P(G)$$

$$- P(EF) - P(EG) - P(FG)$$

$$+ P(EFG)$$



General Inclusion / Exclusion

$$P(E_1 \text{ or } E_2 \text{ or } \cdots \quad E_n) = \sum_{r=1}^{n} (-1)^{r+1} Y_r$$

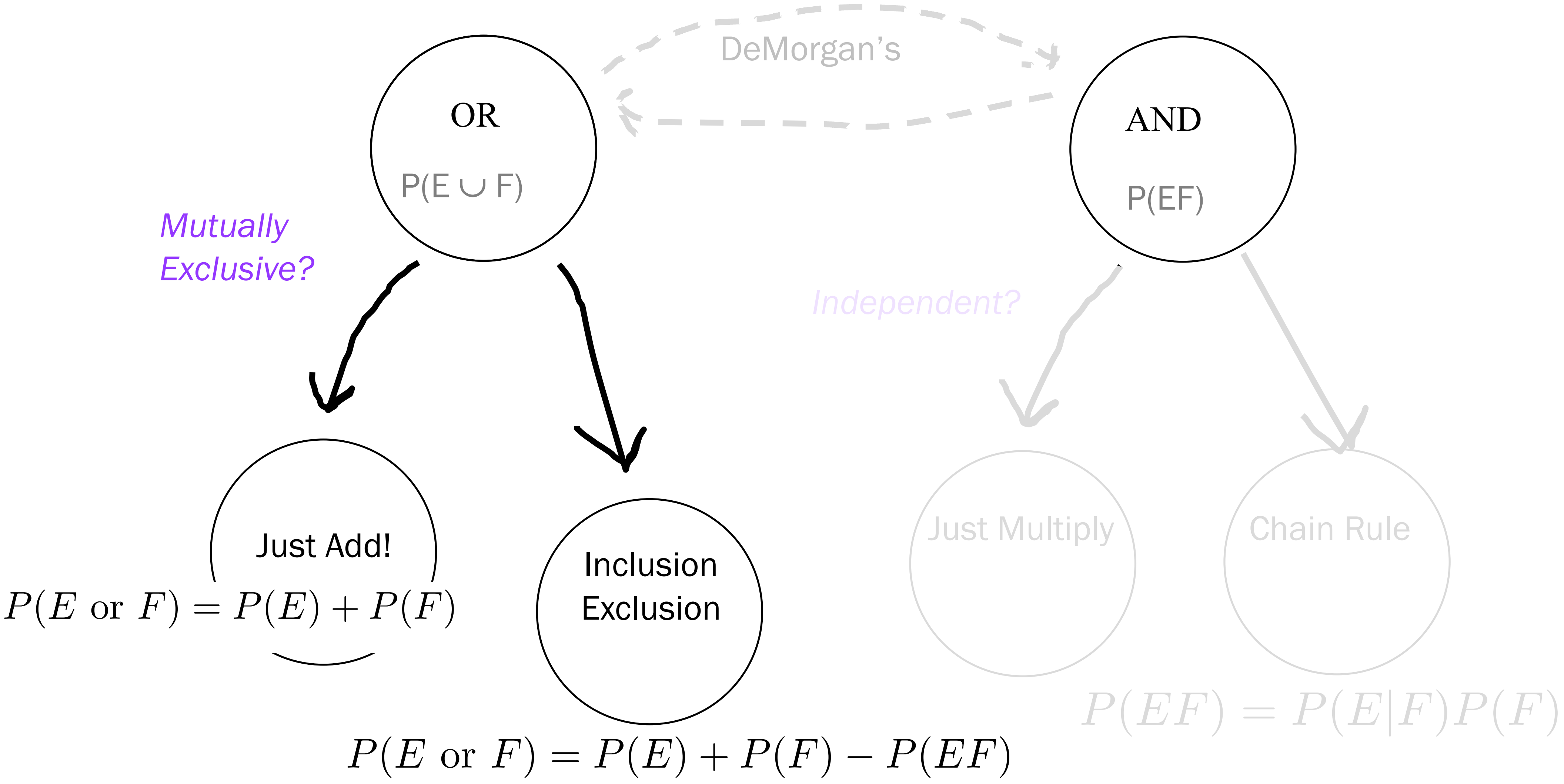
Y_1 = Sum of all events on their own $\sum_i P(E_i)$

Y_2 = Sum of all pairs of events $\sum_{i,j} P(E_i \text{ and } E_j)$ $\text{s.t.} i \neq j$

Y_3 = Sum of all triples of events $\sum_{i,j,k} P(E_i \text{ and } E_j \text{ and } E_k)$ $\text{s.t.} i \neq j, j \neq k, i \neq k$

Where Y_r is the sum, for all combinations of r events, of the probability of the intersection of those events

Today



OR

P(E ∪ F)

DeMorgan's

AND

P(EF)

Mutually Exclusive?

Just Add!

$$P(E \text{ or } F) = P(E) + P(F)$$

Inclusion Exclusion

$$P(E \text{ or } F) = P(E) + P(F) - P(EF)$$

Independent?

Just Multiply

Chain Rule

$$P(EF) = P(E|F)P(F)$$

Probability of “AND”

We the People of the
insure domestic Tranquility, provide for the common defence
and our Posterity, do ordain and establish this Constitution
Article I

Independence

Two events A and B are called **independent** if:

$$P(A) = P(A|B)$$

Knowing that event B happened, doesn’t change our belief that A will happen.

Otherwise, they are called **dependent** events

Alternative Definition of Independence

Notation for *and*

$$\begin{aligned} P(A, B) &= P(A) \cdot P(B|A) && \text{Chain rule} \\ &= P(A) \cdot P(B) && \text{Since B is independent of A} \end{aligned}$$

If you show this is true, you have proved the two events are independent!

If events are *independent* probability of AND is easy!

*You will need to use this “trick” with high probability

Dice, our misunderstood friends

Roll two 6-sided dice, yielding values D_1 and D_2

- Let E be event: $D_1 = 1$
- Let F be event: $D_2 = 1$

What is P(E), P(F), and P(EF)?

- $P(E) = 1/6, \quad P(F) = 1/6, \quad P(EF) = 1/36$
- $P(EF) = P(E)\,P(F)$ → E and F ***independent***

Let G be event: $D_1 + D_2 = 5$ {(1, 4), (2, 3), (3, 2), (4, 1)}

What is P(E), P(G), and P(EG)?

- $P(E) = 1/6, \quad P(G) = 4/36 = 1/9, \quad P(EG) = 1/36$
- $P(EG) \neq P(E)\,P(G)$ → E and G ***dependent***

Intuition through proofs:

Independence is reciprocal

If A is independent of B, then B is independent of A

$$P(A) = P(A|B) \qquad P(B|A) = P(B)$$

Proof:

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

Bayes' Thm.

$$= \frac{P(A)P(B)}{P(A)}$$

Because A is independent of B

$$= P(B)$$

Independence of a complement

Given independent events A and B, prove that A and B^C are independent

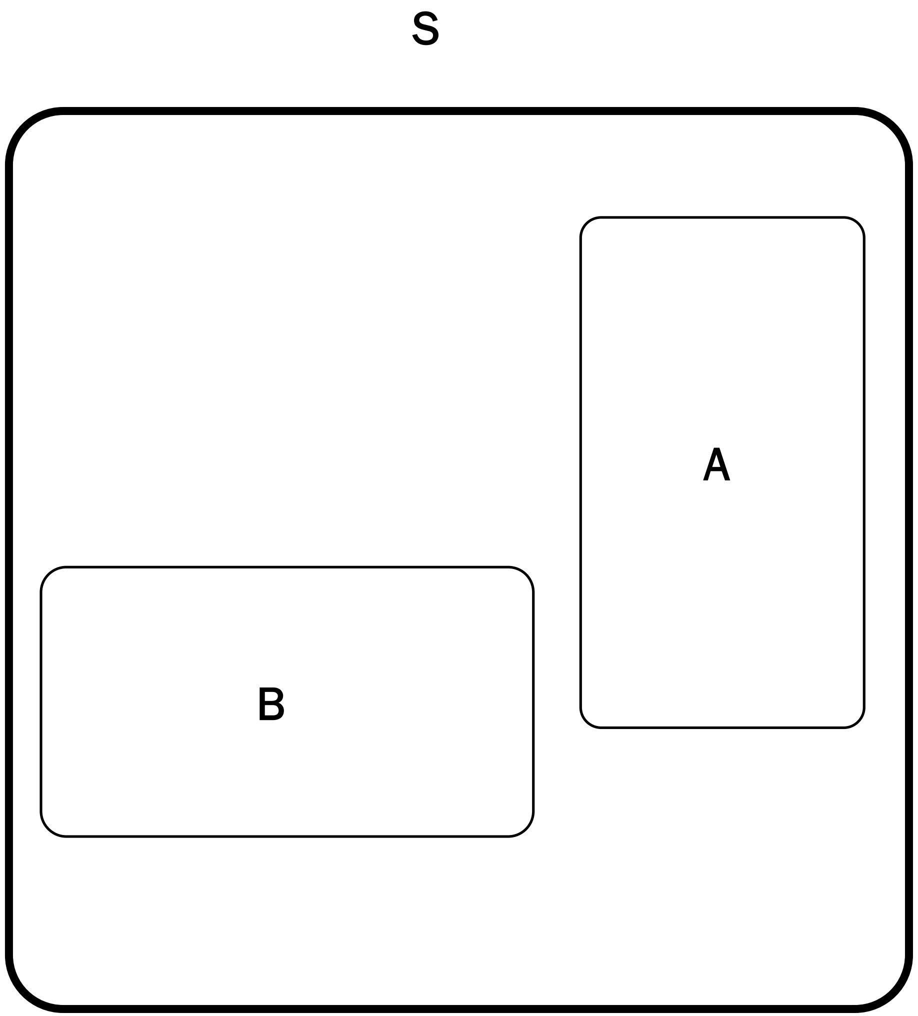
We want to show that $P(AB^C) = P(A)P(B^C)$

$$
\begin{aligned}
P(AB^C) &= P(A) - P(AB) && \text{By Total Law of Prob.} \\
&= P(A) - P(A)P(B) && \text{By independence} \\
&= P(A)[1 - P(B)] && \text{Factoring} \\
&= P(A)P(B^C) && \text{Since } P(B) + P(B^C) = 1
\end{aligned}
$$

So if A and B are independent A and B^C are also independent

What does independence look like?

Independence

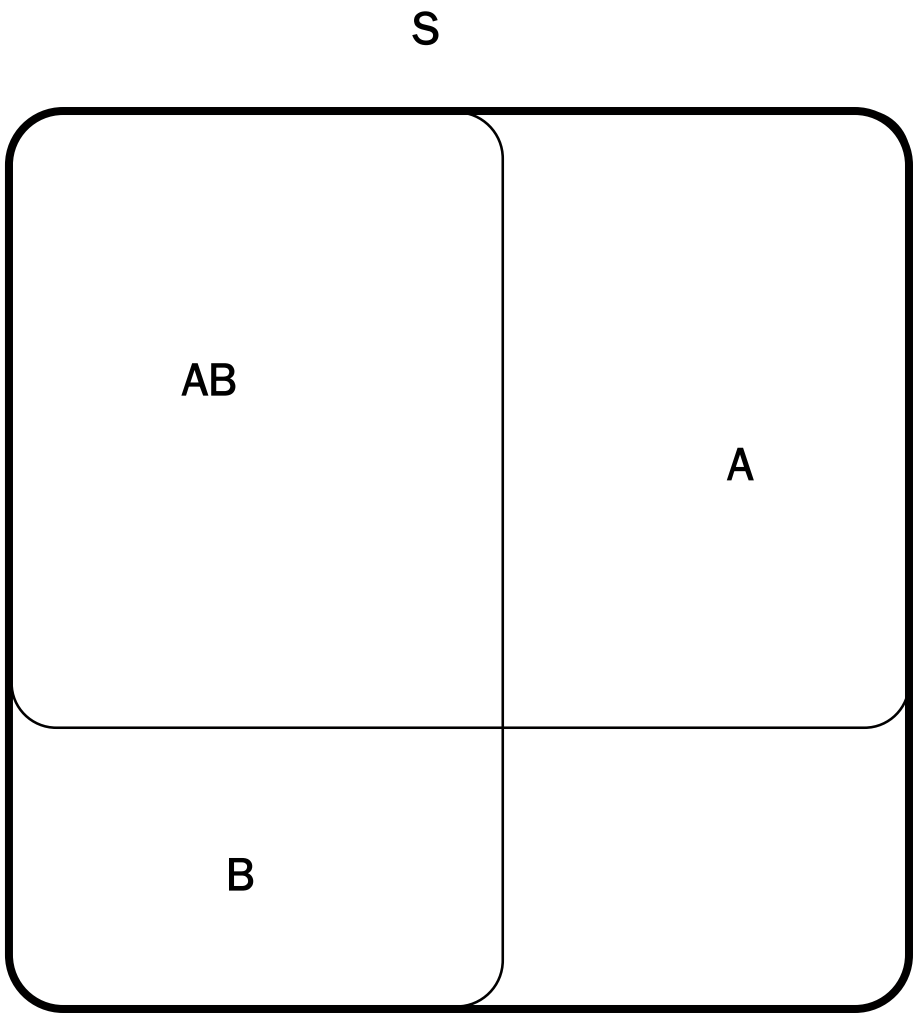


Independence Definition 1:

$$P(AB) = P(A)P(B)$$

$$\cancelto{0}{\frac{|AB|}{|S|}} = \frac{|A|}{|S|} \nrightarrow \frac{|B|}{|S|}$$

Independence



Independence Definition 1:

$$P(AB) = P(A)P(B)$$

$$\frac{|AB|}{|S|} = \frac{|A|}{|S|} \nrightarrow \frac{|B|}{|S|}$$

Independence Definition 2:

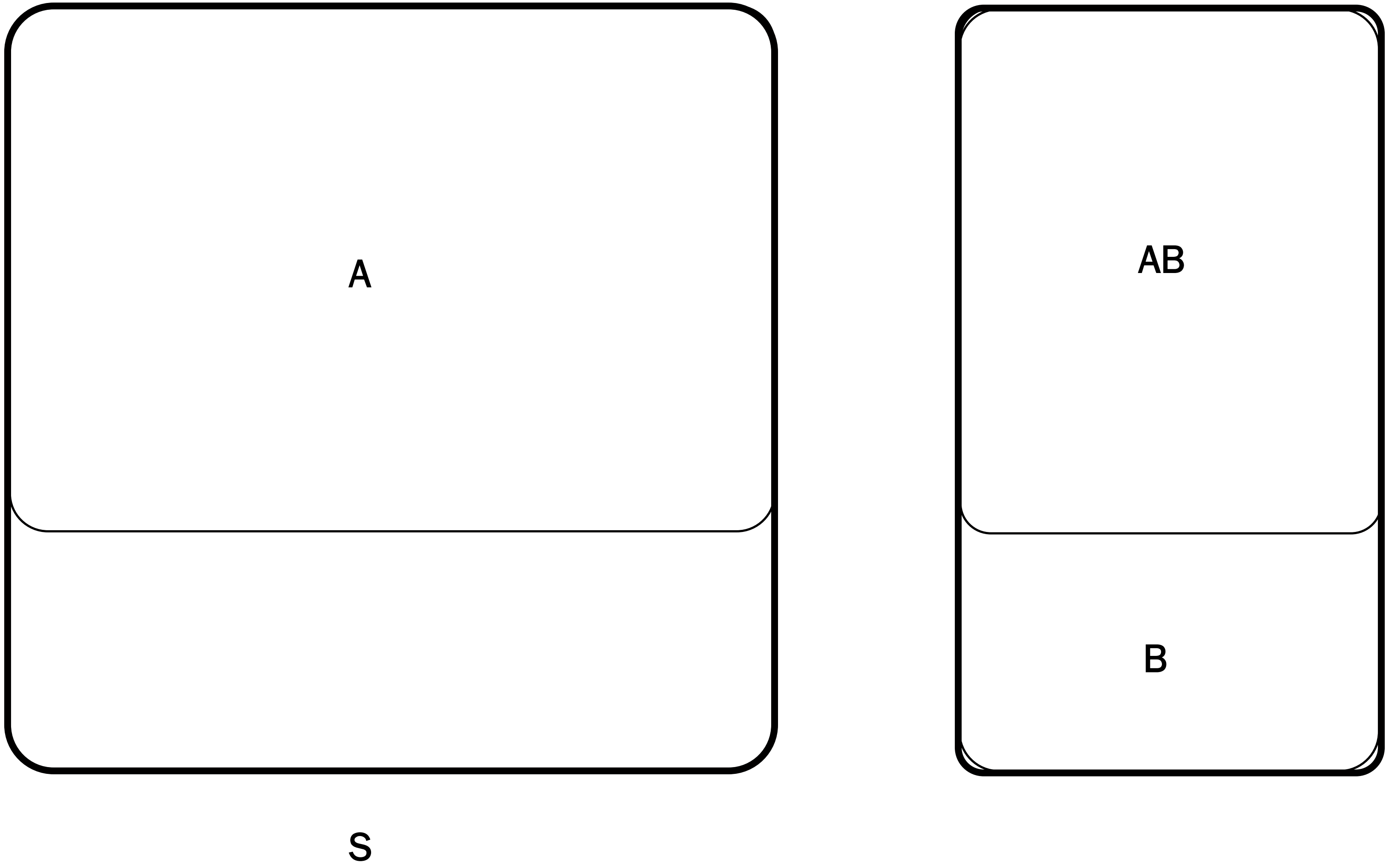
$$P(A|B) = P(A)$$

$$\frac{|AB|}{|B|} = \frac{|A|}{|S|}$$

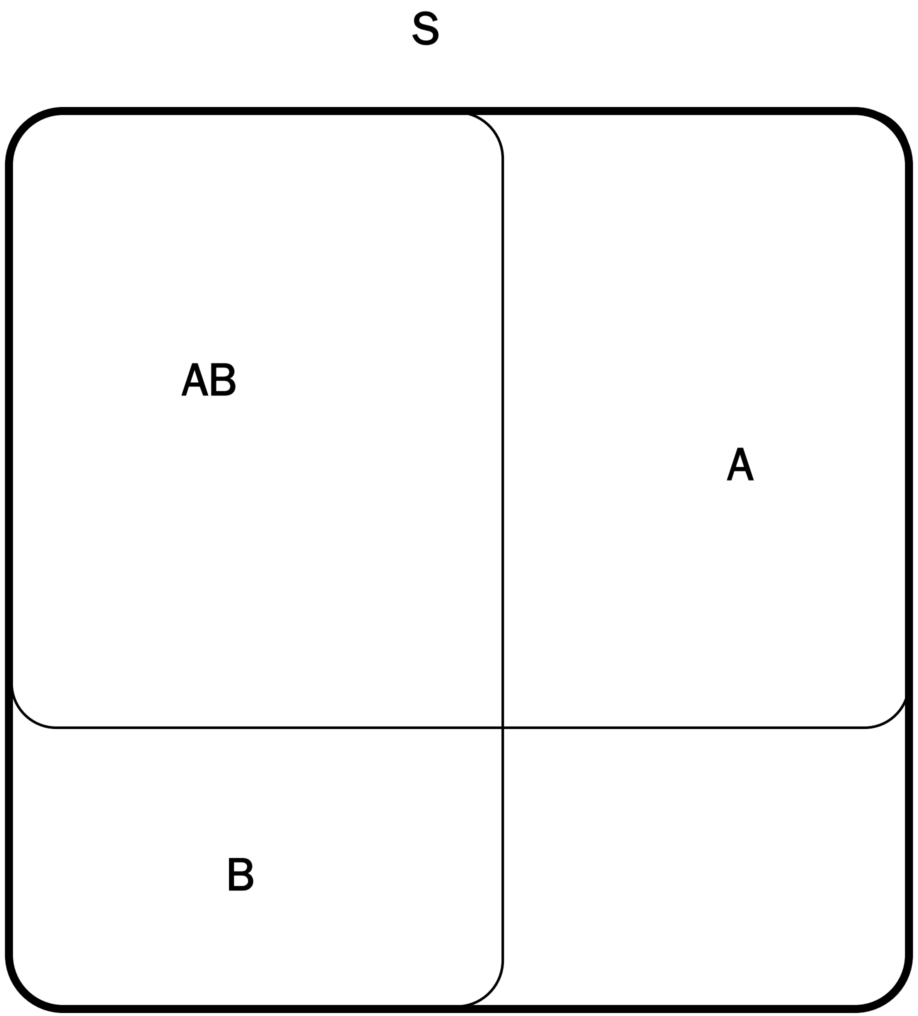
Independence

This ratio, P(A)…

… is the same as this one, P(A|B)



Independence



Independence Definition 1:

$$P(AB) = P(A)P(B)$$

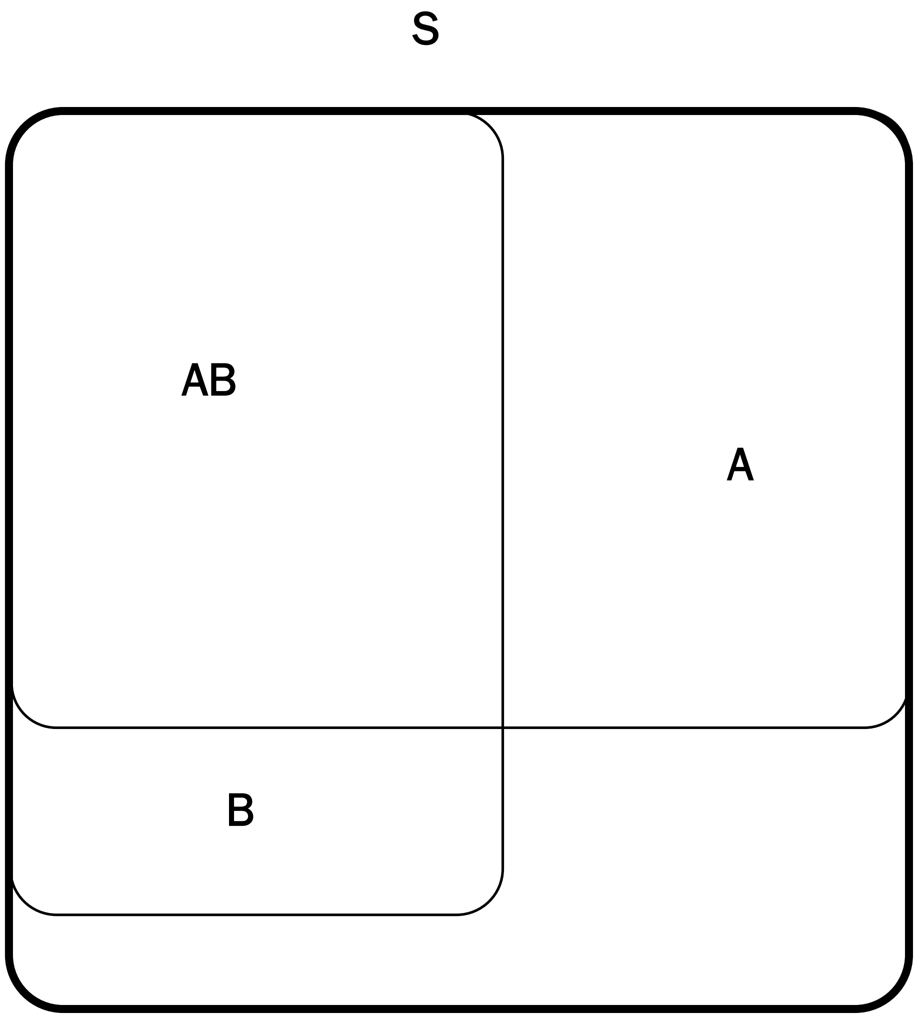
$$\frac{|AB|}{|S|} = \frac{|A|}{|S|} \nrightarrow \frac{|B|}{|S|}$$

Independence Definition 2:

$$P(A|B) = P(A)$$

$$\frac{|AB|}{|B|} = \frac{|A|}{|S|}$$

Dependence



Independence Definition 1:

$$P(AB) = P(A)P(B)$$

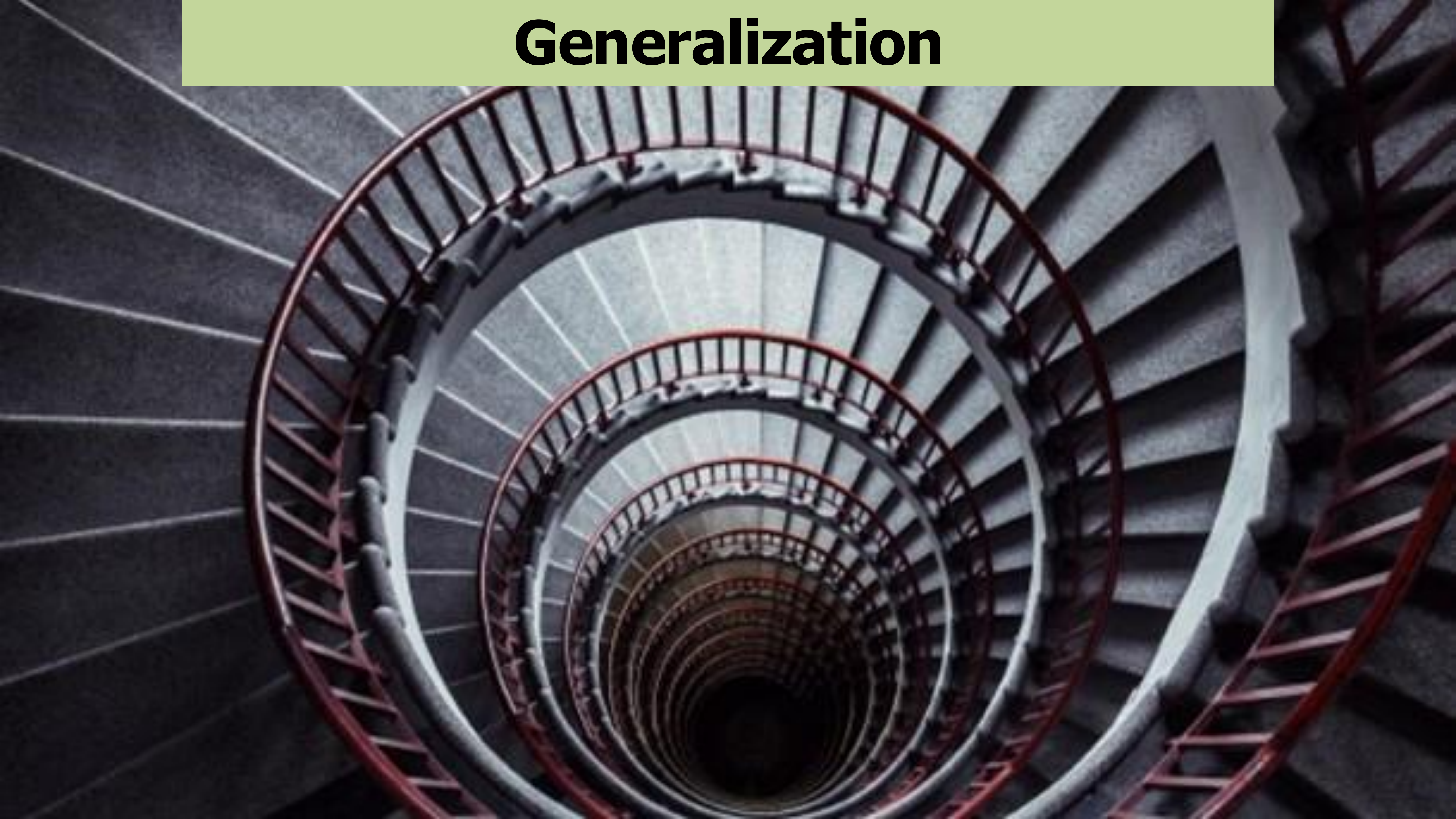
$$\frac{|AB|}{|S|} = \frac{|A|}{|S|} \not\rightarrow \frac{|B|}{|S|}$$

Independence Definition 2:

$$P(A|B) = P(A)$$

$$\frac{|AB|}{|B|} = \frac{|A|}{|S|}$$

Generalization



Generalized Independence

General definition of Independence:

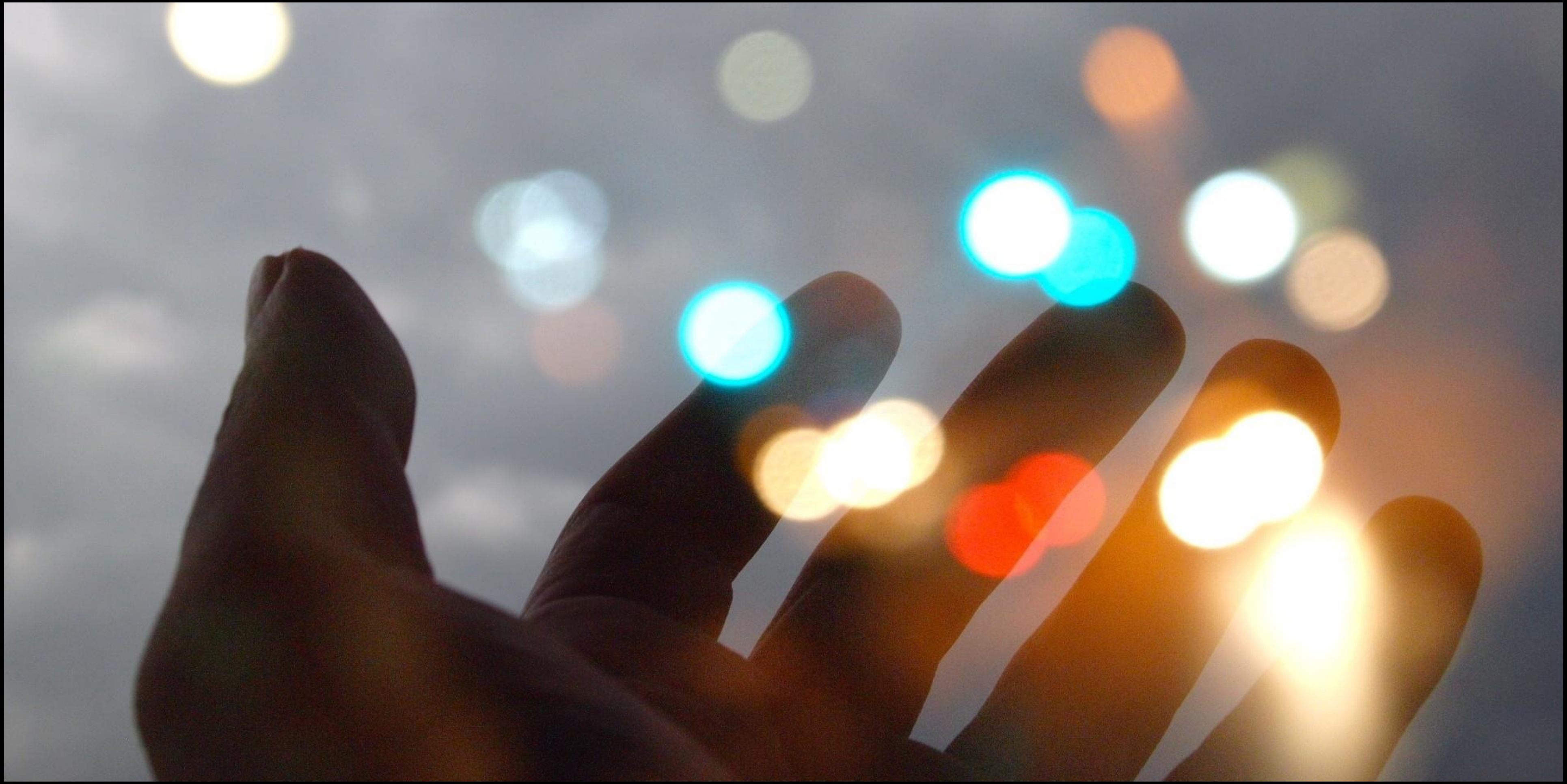
Events E$_1$, E$_2$, ..., E$_n$ are independent if **for every subset** with r elements (where *r* ≤ *n*) it holds that:

$$P(E_{1'}E_{2'}E_{3'}\ldots E_{r'}) = P(E_{1'})P(E_{2'})P(E_{3'})\ldots P(E_{r'})$$

<u>Example</u>: outcomes of *n* separate flips of a coin are all independent of one another

- Each flip in this case is called a “trial” of the experiment

Math > Intuition



Two Dice

Roll two 6-sided dice, yielding values D_1 and D_2

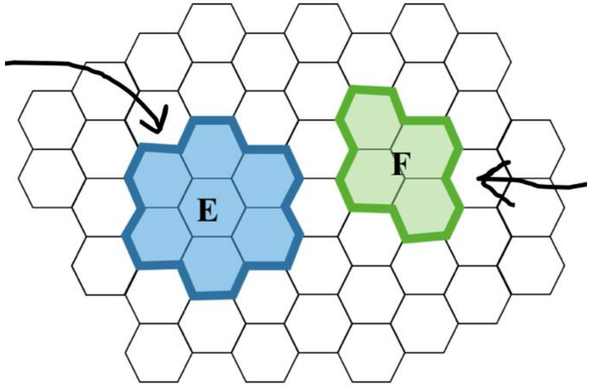
- Let E be event: $D_1 = 1$
- Let F be event: $D_2 = 6$
- Are E and F independent? Yes!

Let G be event: $D_1 + D_2 = 7$

- Are E and G independent? Yes!
- P(E) = 1/6, P(G) = 1/6, P(E G) = 1/36 [roll (1, 6)]
- Are F and G independent? Yes!
- P(F) = 1/6, P(G) = 1/6, P(F G) = 1/36 [roll (1, 6)]
- Are E, F and G independent? No!
- $P(EFG) = 1/36 \neq 1/216 = (1/6)(1/6)(1/6)$

New Ability

Properties of Pairs of Events



Mutually Exclusive

$$P(A \text{ and } B) = 0$$

also:

$$P(A \text{ or } B) = P(A) + P(B)$$

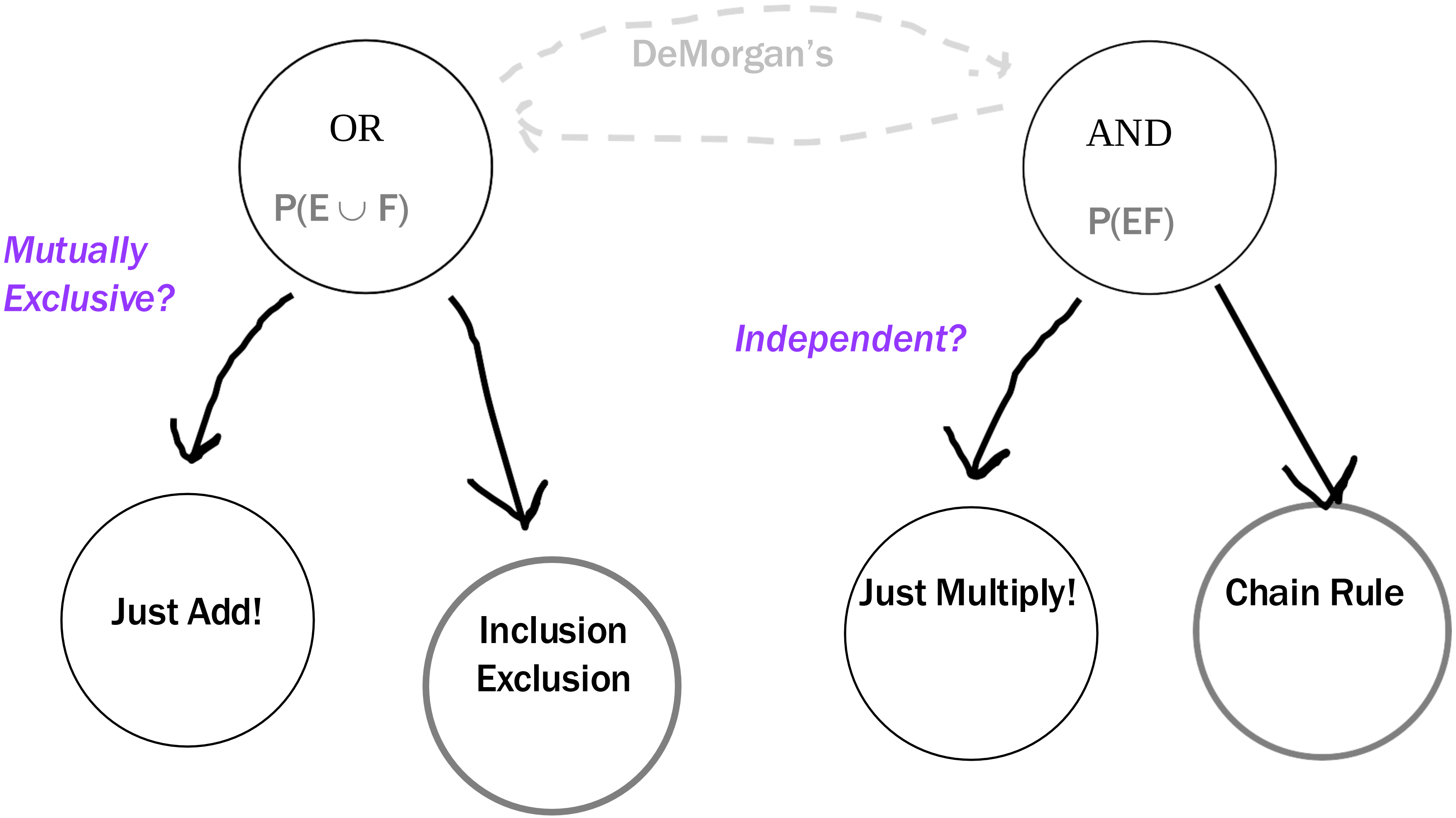
Independent

$$P(A) = P(A|B)$$

also:

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

Today


DeMorgan's
OR
P(E ∪ F)
AND
P(EF)
Mutually Exclusive?
Independent?
Just Add!
Inclusion Exclusion
Just Multiply!
Chain Rule

Think of the children as independent trials

Independence:
$P(A \text{ and } B) = P(A) \cdot P(B)$

Two parents both have an (A, a) gene pair.

- Each parent will pass on one of their genes (each gene equally likely) to a child.
- The probability of **any single child** having curly hair (the recessive trait) is 0.25, **independent** of other siblings.
- There are three children.

What is the probability that all three children have curly hair?

Let E_1, E_2, E_3 be the events that child 1, 2, and 3 have curly hair, respectively.

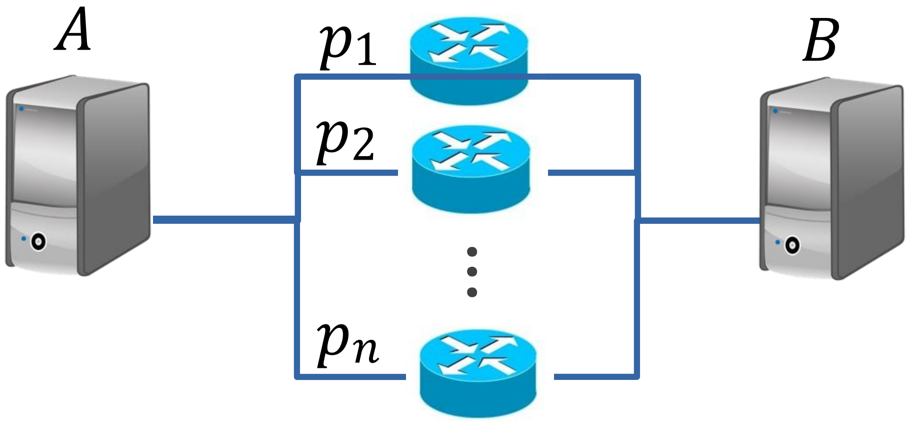
$$P(E_1E_2E_3) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)$$
$$= P(E_1)P(E_2)P(E_3)$$

Network reliability

Consider the following parallel network:

- n independent routers, each with probability p_{i} of functioning (where $1 \leq i \leq n$)
- E = functional path from A to B exists.

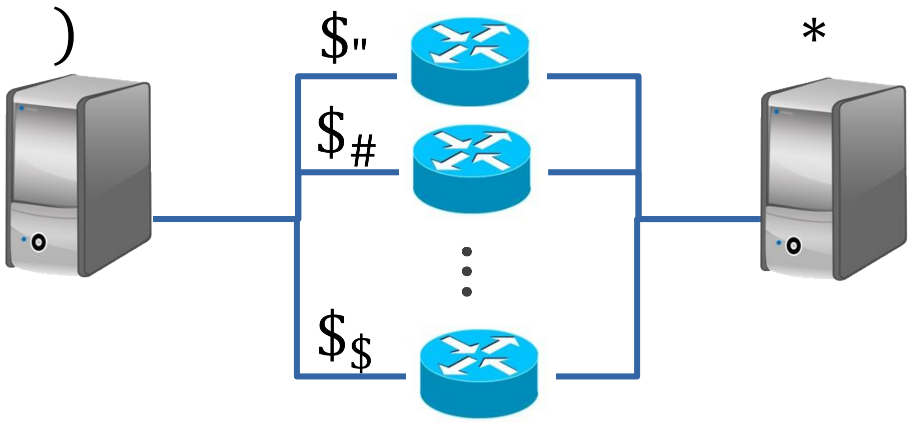
What is $P(E)$?



Network reliability

Consider the following parallel network:

- # independent routers, each with probability $\$_!$ of functioning (where 1 ≤ ' ≤ #)
- " = functional path from A to B exists.

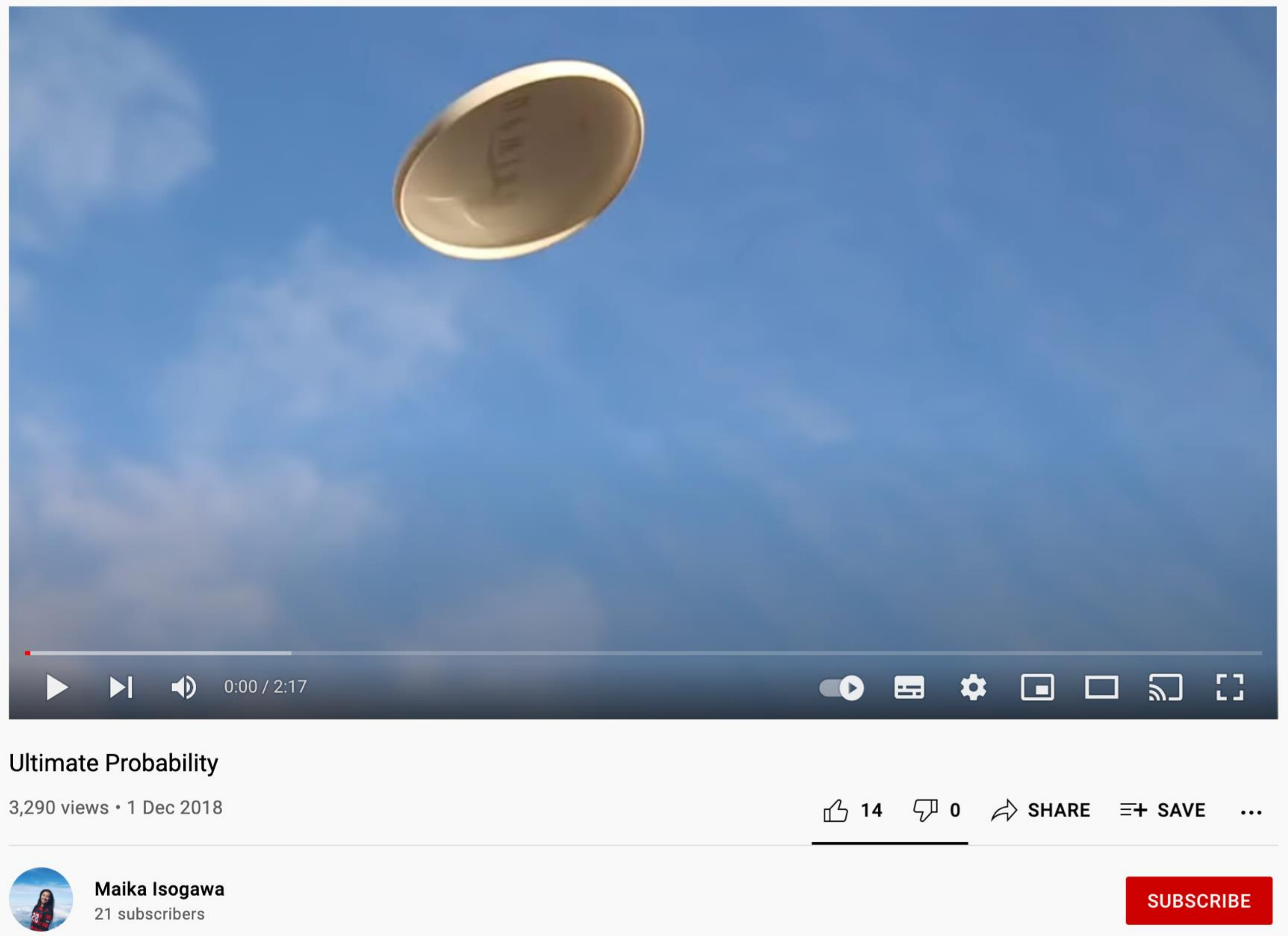


What is $P(E)$?

$$
\begin{aligned}
P(E) &= P(\geq 1 \text{ one router works})\\
&= 1 - P(\text{all routers fail})\\
&= 1 - (1 - p_!)(1 - p_")\cdots(1 - p_\%)\\
&= 1 - \prod_{\& \ !}^{\%}(1 - p_\&)
\end{aligned}
$$

≥ 1 with independent trials: take complement

Story: Ultimate Probability



Ultimate Probability

3,290 views • 1 Dec 2018

Maika Isogawa

21 subscribers

https://www.maikaisogawa.com/ultimate-frisbee-probability/

https://www.youtube.com/watch?v=H2IfTwGisOg

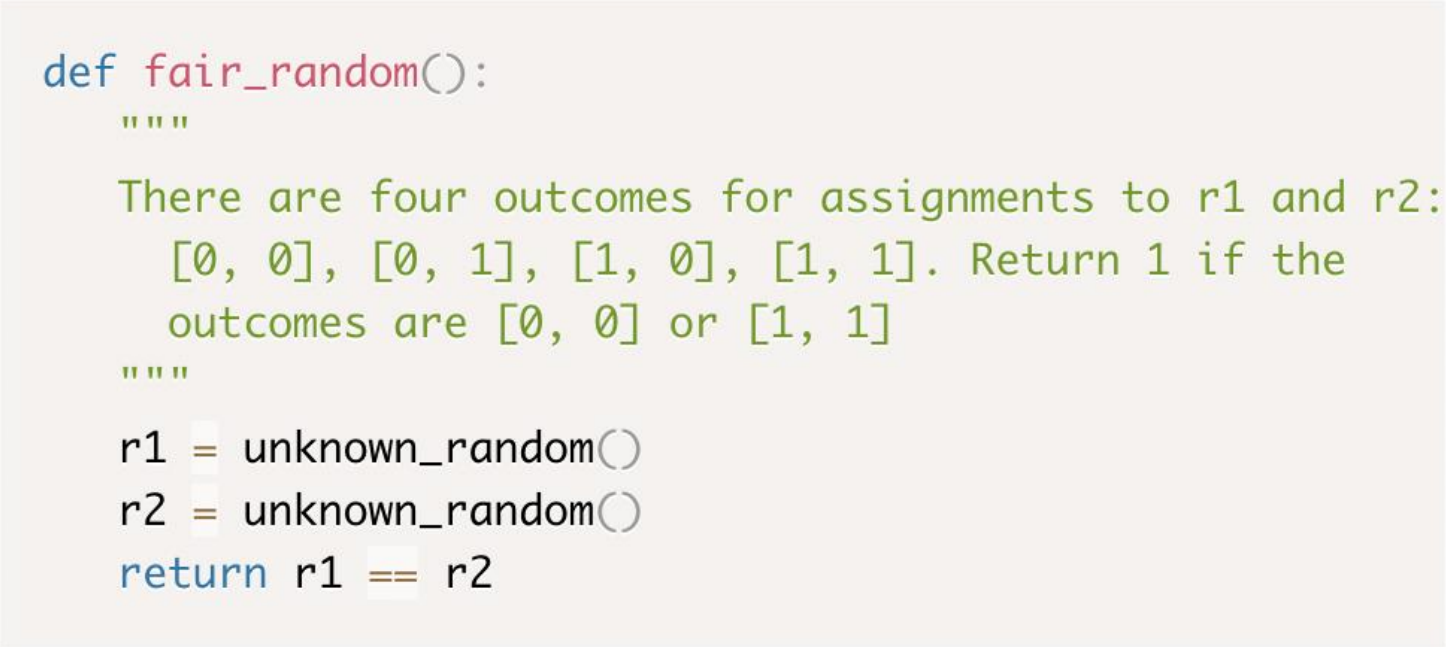
Practice: Lets do the Frisbee Problem.

You flip **two frisbees**. For each frisbee, the probability that it lands "**heads**" is **0.6**. The two frisbees are considered "**even**" if both frisbees are heads **or** both frisbees are tails.

What is the probability that the frisbees are even?

You have started PSet2!

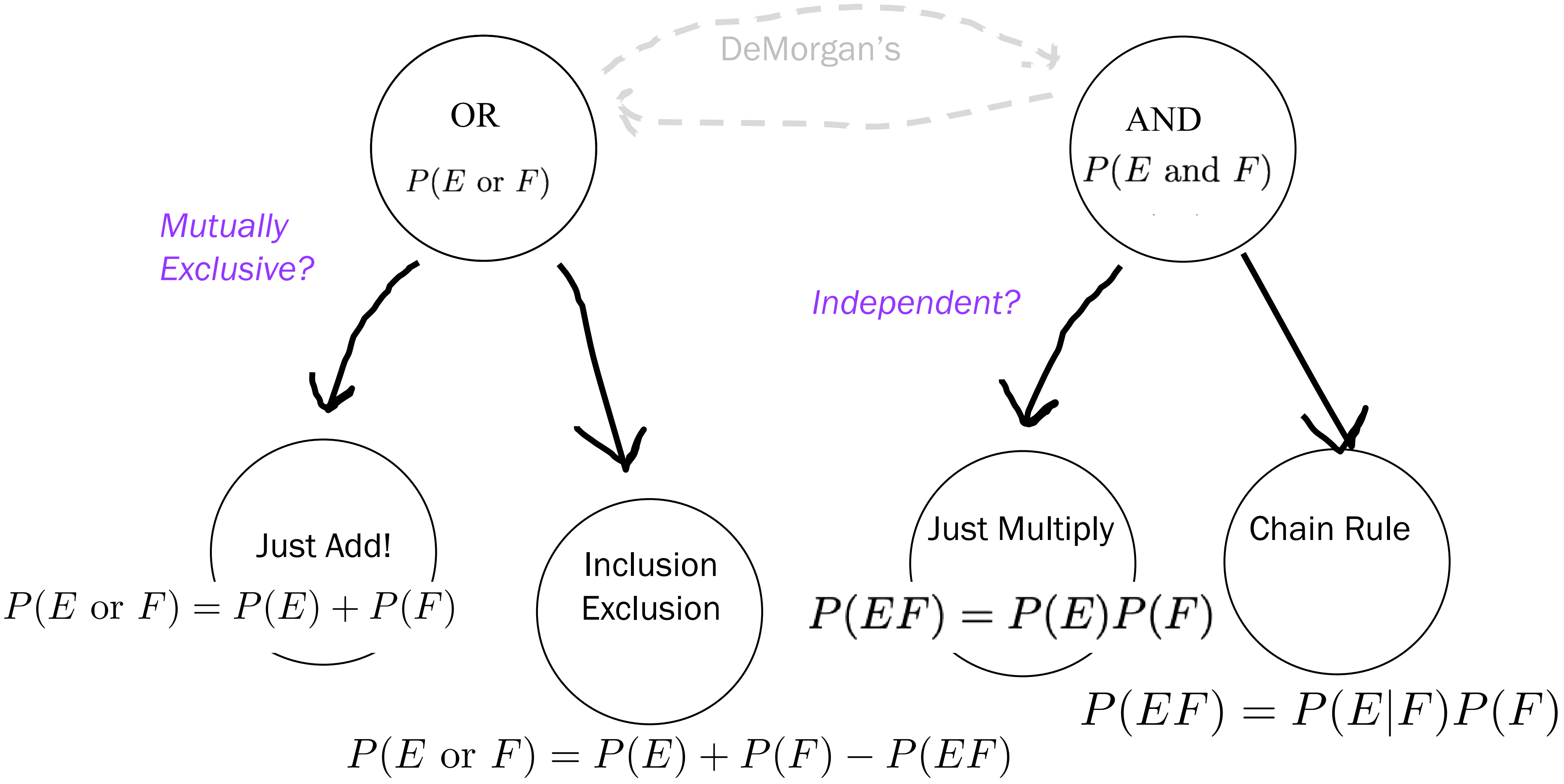
Let *p* be the probability of a 1 from unknown_random.
What is the probability of a True from fair_random if p =.45?



```
def fair_random():
    """
    There are four outcomes for assignments to r1 and r2:
      [0, 0], [0, 1], [1, 0], [1, 1]. Return 1 if the
      outcomes are [0, 0] or [1, 1]
    """
    r1 = unknown_random()
    r2 = unknown_random()
    return r1 == r2
```

https://cs109psets.netlify.app/win25/pset2/fairRandom

Today



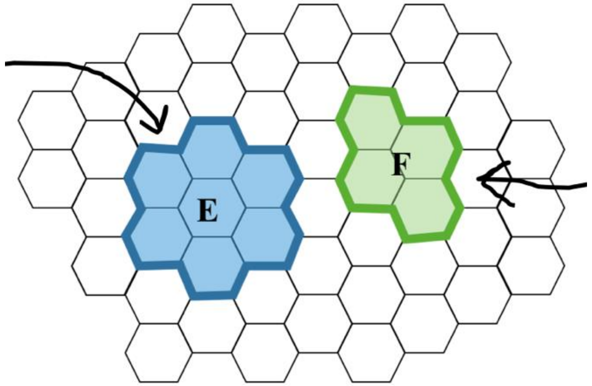
$$P(E \text{ or } F) = P(E) + P(F)$$

$$P(EF) = P(E)P(F)$$

$$P(EF) = P(E|F)P(F)$$

$$P(E \text{ or } F) = P(E) + P(F) - P(EF)$$

Pedagogical Pause



Mutually Exclusive

$$P(A \text{ and } B) = 0$$

Makes ***OR*** easy:

$$P(A \text{ or } B) = P(A) + P(B)$$

Independent

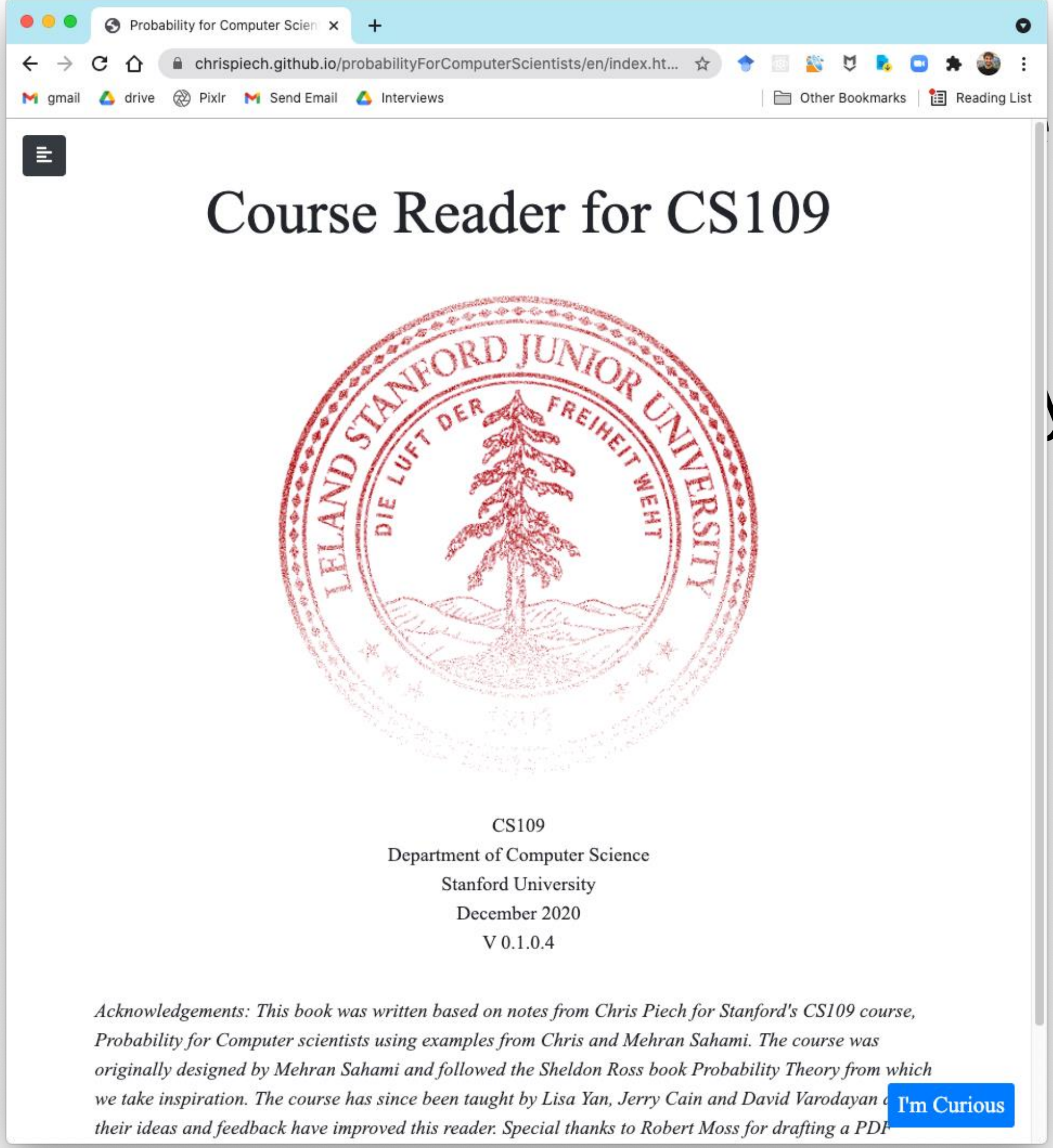
$$P(A) = P(A|B)$$

Makes ***AND*** easy:

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

Say a coin comes up heads with probability 0.6. Flip the coin 10 times. Each coin flip is an **independent** trial. What is the probability of exactly 4 heads?

The Most Important Core Probability Question


Probability for Computer Scien
chrispiech.github.io/probabilityForComputerScientists/en/index.ht...
gmail
drive
Pixlr
Send Email
Interviews
Other Bookmarks
Reading List
I'm Curious

Course Reader for CS109

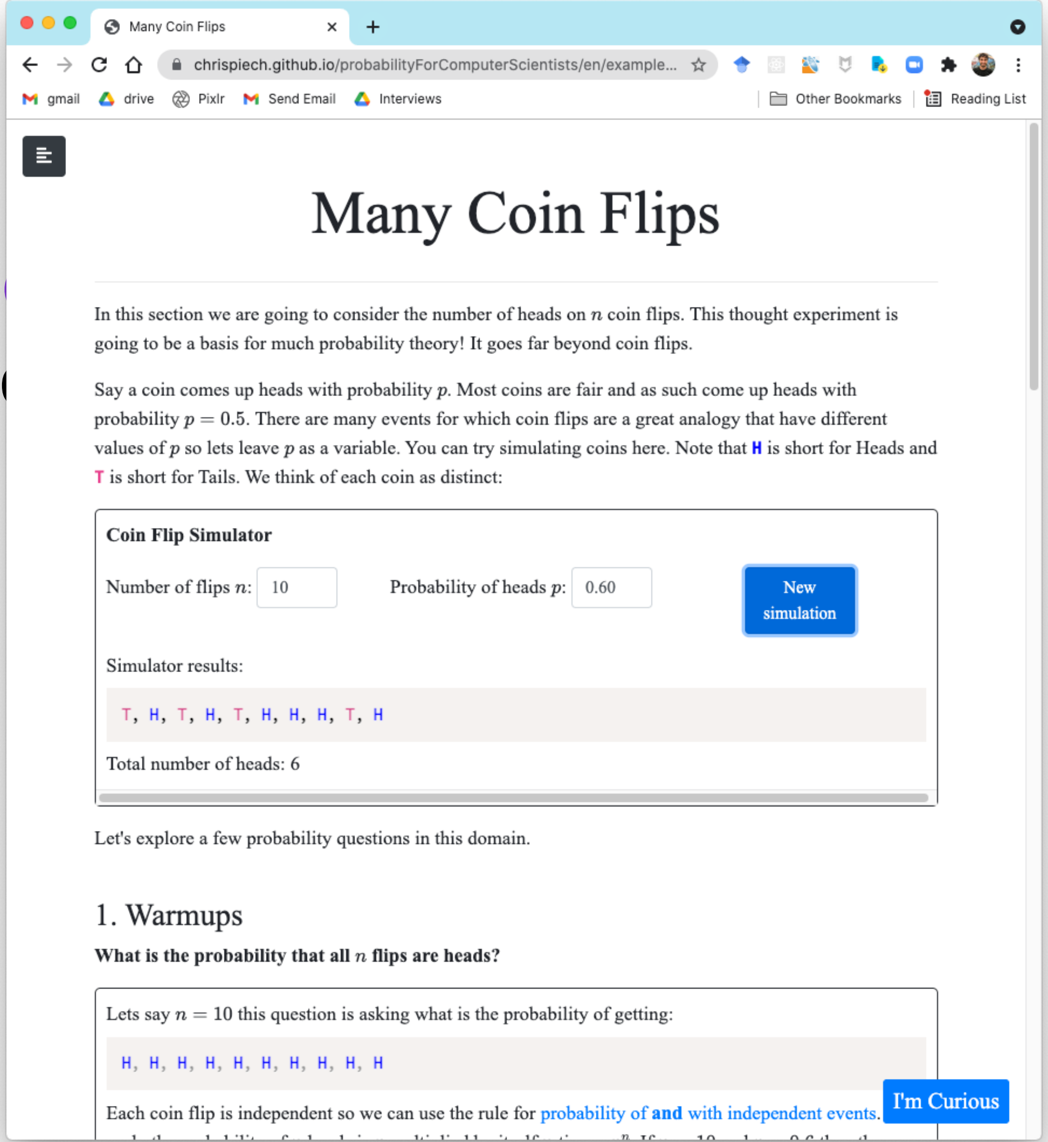
CS109
Department of Computer Science
Stanford University
December 2020
V 0.1.0.4

Acknowledgements: This book was written based on notes from Chris Piech for Stanford's CS109 course, Probability for Computer scientists using examples from Chris and Mehran Sahami. The course was originally designed by Mehran Sahami and followed the Sheldon Ross book Probability Theory from which we take inspiration. The course has since been taught by Lisa Yan, Jerry Cain and David Varodayan a their ideas and feedback have improved this reader. Special thanks to Robert Moss for drafting a PDF

ads

dep

y of (



Many Coin Flips

In this section we are going to consider the number of heads on n coin flips. This thought experiment is going to be a basis for much probability theory! It goes far beyond coin flips.

Say a coin comes up heads with probability p. Most coins are fair and as such come up heads with probability $p = 0.5$. There are many events for which coin flips are a great analogy that have different values of p so lets leave p as a variable. You can try simulating coins here. Note that H is short for Heads and T is short for Tails. We think of each coin as distinct:

Coin Flip Simulator

Number of flips n: 10 Probability of heads p: 0.60

Simulator results:

```
T, H, T, H, T, H, H, H, T, H
```

Total number of heads: 6

Let's explore a few probability questions in this domain.

1. Warmups

What is the probability that all n flips are heads?

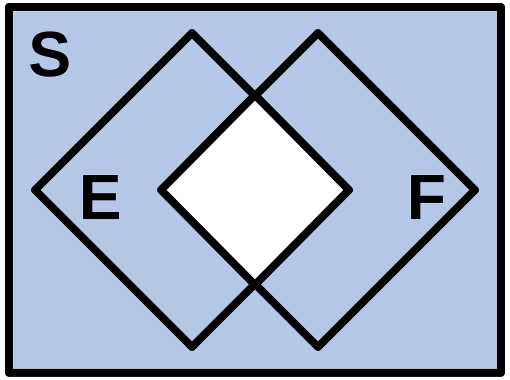
Lets say $n = 10$ this question is asking what is the probability of getting:

```
H, H, H, H, H, H, H, H, H, H
```

Each coin flip is independent so we can use the rule for probability of **and** with independent events.

De Morgan's Laws

De Morgan's Law lets you alternate between AND and OR.

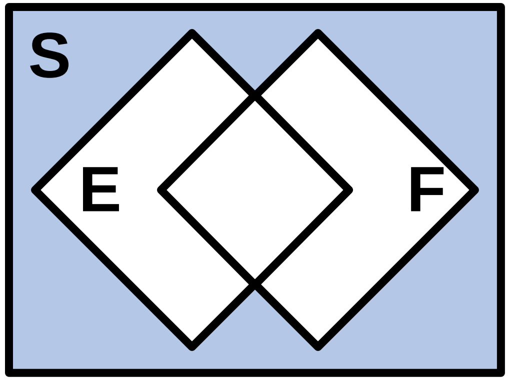


$$(E \text{ and } F)^C = E^C \text{ or } F^C$$

In probability:

$$P(E_1 E_2 \cdots E_n)$$
$$= 1 - P((E_1 E_2 \cdots E_n)^C)$$
$$= 1 - P(E_1^C \text{ or } E_2^C \text{ or } \cdots E_n^C)$$

Great if E^C mutually exclusive!



$$(E \text{ or } F)^C = E^C \text{ and } F^C$$

In probability:

$$P(E_1 \text{ or } E_2 \text{ or } \cdots E_n)$$
$$= 1 - P((E_1 \text{ or } E_2 \text{ or } \cdots E_n)^C)$$
$$= 1 - P(E_1^C E_2^C \cdots E_n^C)$$

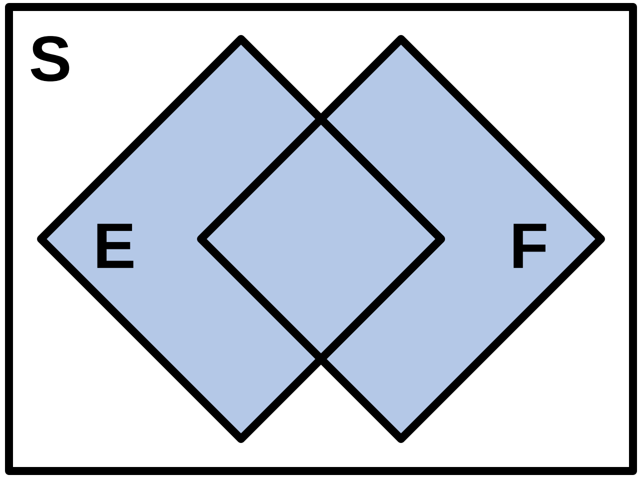
Great if E_i independent!

Sets Review

Say E and F are events in S

Event that is in E or F

E ∪ F



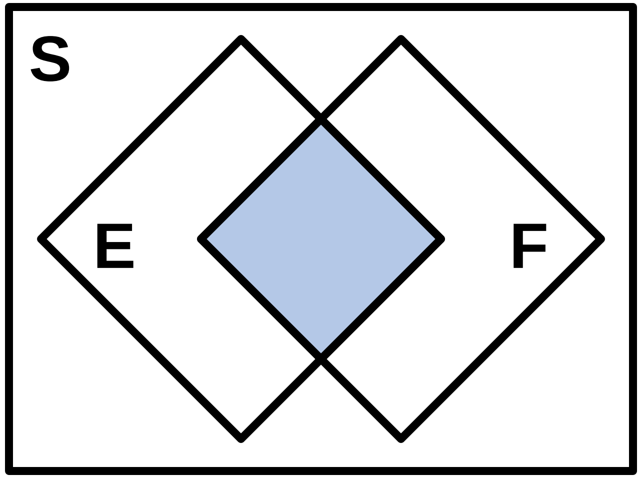
- S = {1, 2, 3, 4, 5, 6} die roll outcome
- E = {1, 2} F = {2, 3} E ∪ F = {1, 2, 3}

Sets Review

Say E and F are events in S

Event that is in E and F

$\mathbf{E \cap F}$

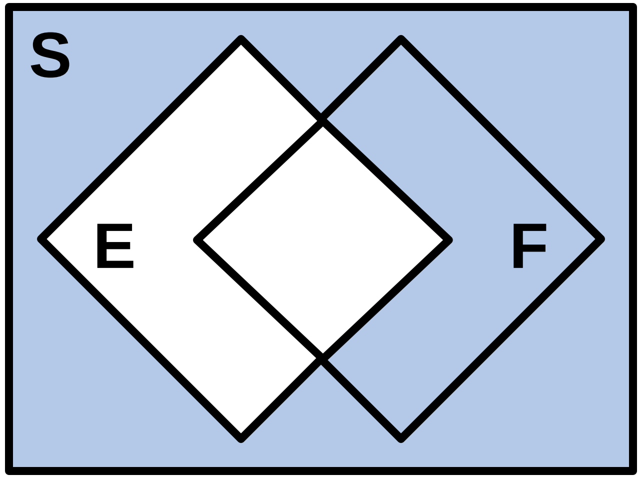


Sets Review

Say E and F are events in S

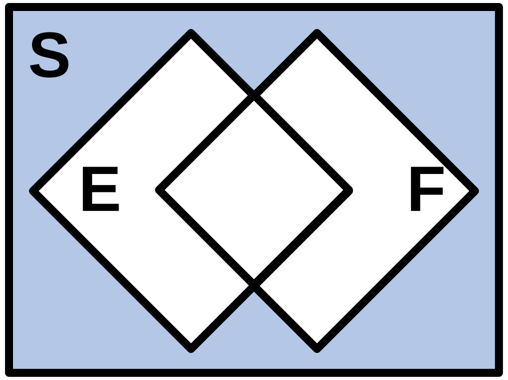
Event that is not in E (called complement of E)

$\mathbf{E^c}$



Sets Review

Say E and F are subsets of S



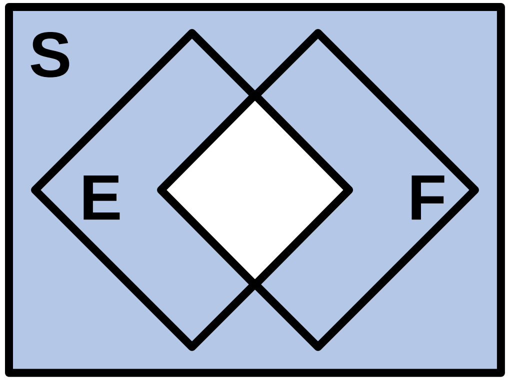
Which of these two is it?

a) $(E \text{ or } F)^C$

b) $(E^C \text{ and } F^C)$

Sets Review

Say E and F are subsets of S



Which of these two is it?

a) $(E \text{ and } F)^C$

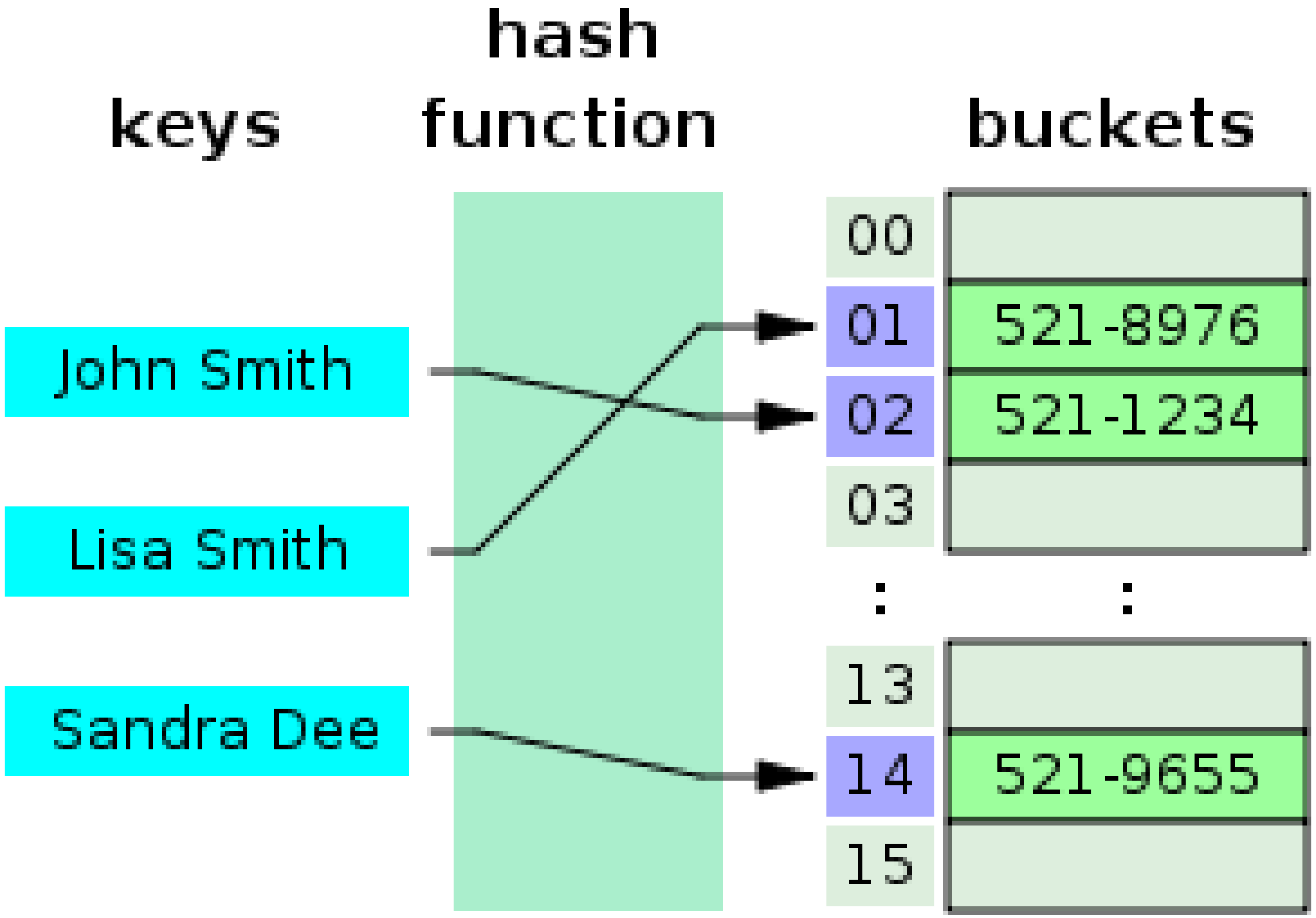
b) $(E^C \text{ or } F^C)$

Augustin Demorgan

Jason Alexander

- British Mathematician who wrote the book “Formal Logic”in 1847
- Celebrity lookalike is Jason Alexander from Seinfeld.

Hash Tables


keys
hash function
buckets
John Smith
Lisa Smith
Sandra Dee
00
01
02
03
:
13
14
15
521-8976
521-1234
:
521-9655

Hash table fun

- m strings are hashed (not uniformly) into a hash table with n buckets.
- Each string hash is an **independent trial** w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

E = bucket 1 has ≥ 1 string hashed into it?

Hash table fun

- m strings are hashed (not uniformly) into a hash table with n buckets.
- Each string hash is an **independent trial** w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if
E = bucket 1 has $\geq$ 1 string hashed into it?

Def: S_i = string i hashes to bucket 1
S_i^C = string i doesn't hash to bucket 1

$$P(S_i) = p_1$$
$$P(S_i^C) = 1 - p_1$$

Hash table fun

- m strings are hashed (not uniformly) into a hash table with n buckets.
- Each string hash is an **independent trial** w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if
E = bucket 1 has ≥ 1 string hashed into it?

Def: S_i = string i hashes to bucket 1
S_i^C = string i doesn't hash to bucket 1

$$P(S_i) = p_1$$
$$P(S_i^C) = 1 - p_1$$

WTF (not-real acronym for Want To Find):

$$P(E) = P(S_1 \text{ or } S_2 \text{ or } \cdots \text{ or } S_m)$$

Hash table fun

- m strings are hashed (not uniformly) into a hash table with n buckets.
- Each string hash is an **independent trial** w.p. p_{i} of getting hashed into bucket i.

What is $P(E)$ if
E = bucket 1 has ≥ 1 string hashed into it?

Def: S_{i} = string i hashes to bucket 1
S^C_{i} = string i doesn't hash to bucket 1

$$P(S_{\text{i}}) = p_1$$
$$P(S^C_{\text{i}}) = 1 - p_1$$

<u>WTF</u> (not-real acronym for Want To Find):

$$\begin{aligned} P(E) &= P(S_1 \text{ or } S_2 \text{ or } \cdots \text{ or } S_m) \\ &= 1 - P\left((S_1 \text{ or } S_2 \text{ or } \cdots \cup S_m)^c\right) \\ &= 1 - P(S_1^c S_2^c \cdots S_m^c) \end{aligned}$$

DeMorgans Law

Hash table fun

- m strings are hashed (not uniformly) into a hash table with n buckets.
- Each string hash is an **independent trial** w.p. p_{i} of getting hashed into bucket i.

What is $P(E)$ if
E = bucket 1 has ≥ 1 string hashed into it?

Def: S_{i} = string i hashes to bucket 1
S^{C}_{i} = string i doesn't hash to bucket 1

$P(S_{\mathrm{i}}) = p_1$
$P(S^{C}_{\mathrm{i}}) = 1 - p_1$

WTF (not-real acronym for Want To Find):

$$P(E) = P(S_1 \text{ or } S_2 \text{ or } \cdots \text{ or } S_m)$$
$$= 1 - P\left((S_1 \text{ or } S_2 \text{ or } \cdots \cup S_m)^c\right)$$
$$= 1 - P(S_1^c S_2^c \cdots S_m^c)$$

DeMorgans Law

$$= 1 - P(S_1^c) P(S_2^c) \cdots P(S_m^c)$$

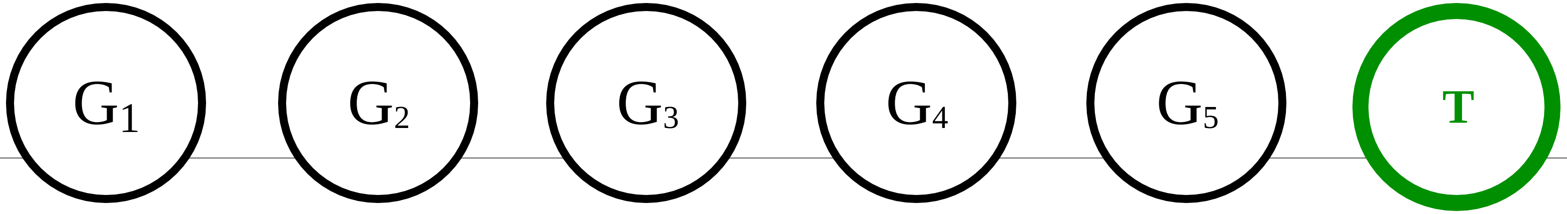
S_i independent trials

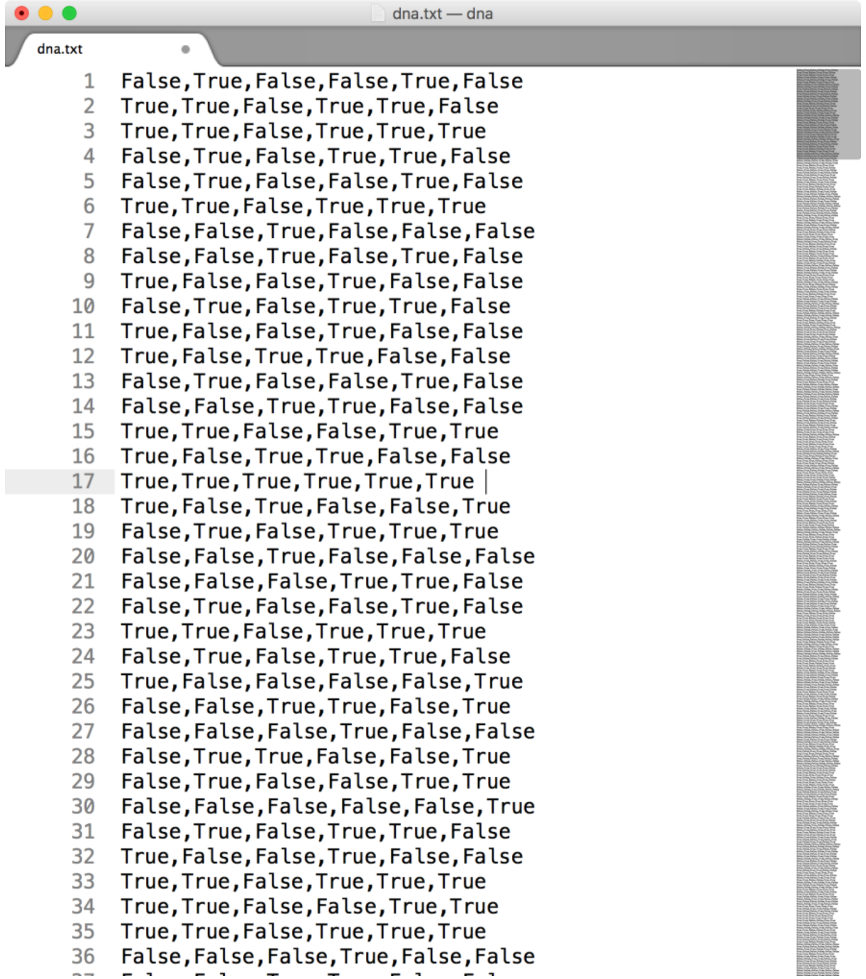
$$= 1 - (P(S_1^c))^m$$
$$= 1 - (1 - p_1)^m$$

Here we are

Source: The Hobbit

G_1 G_2 G_3 G_4 G_5 T



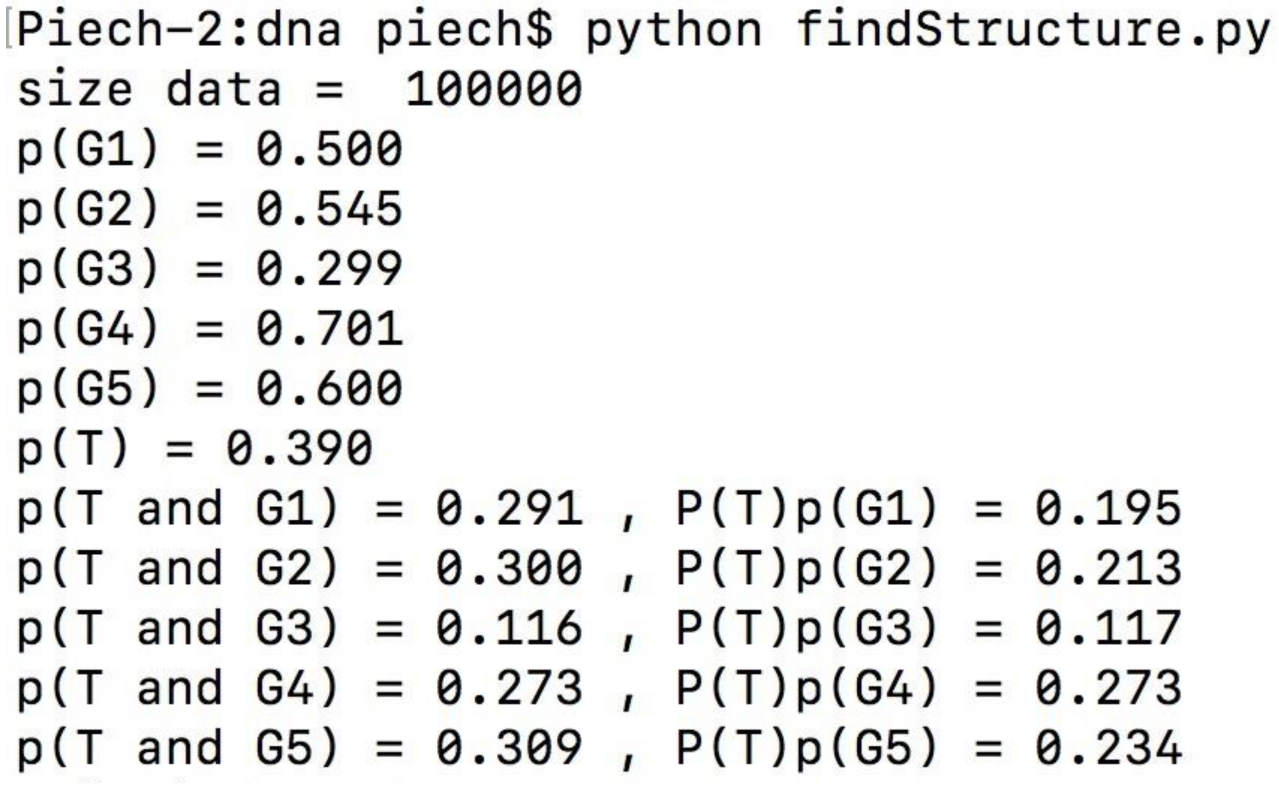


```
False,True,False,False,True,False
True,True,False,True,True,False
True,True,False,True,True,True
False,True,False,True,True,False
False,True,False,False,True,False
True,True,False,True,True,True
False,False,True,False,False,False
False,False,True,False,True,False
True,False,False,True,False,False
False,True,False,True,True,False
True,False,False,True,False,False
True,False,True,True,False,False
False,True,False,False,True,False
False,False,True,True,False,False
True,True,False,False,True,True
True,False,True,True,False,False
True,True,True,True,True,True
True,False,True,False,False,True
False,True,False,True,True,True
False,False,True,False,False,False
False,False,False,True,True,False
False,True,False,False,True,False
True,True,False,True,True,True
False,True,False,True,True,False
True,False,False,False,False,True
False,False,True,True,False,True
False,False,False,True,False,False
False,True,True,False,False,True
False,True,False,False,True,True
False,False,False,False,False,True
False,True,False,True,True,False
True,False,False,True,False,False
True,True,False,True,True,True
True,True,False,False,True,True
True,True,False,True,True,True
False,False,False,True,False,False
```

100,000 samples

6 observations per sample

Discovered Pattern

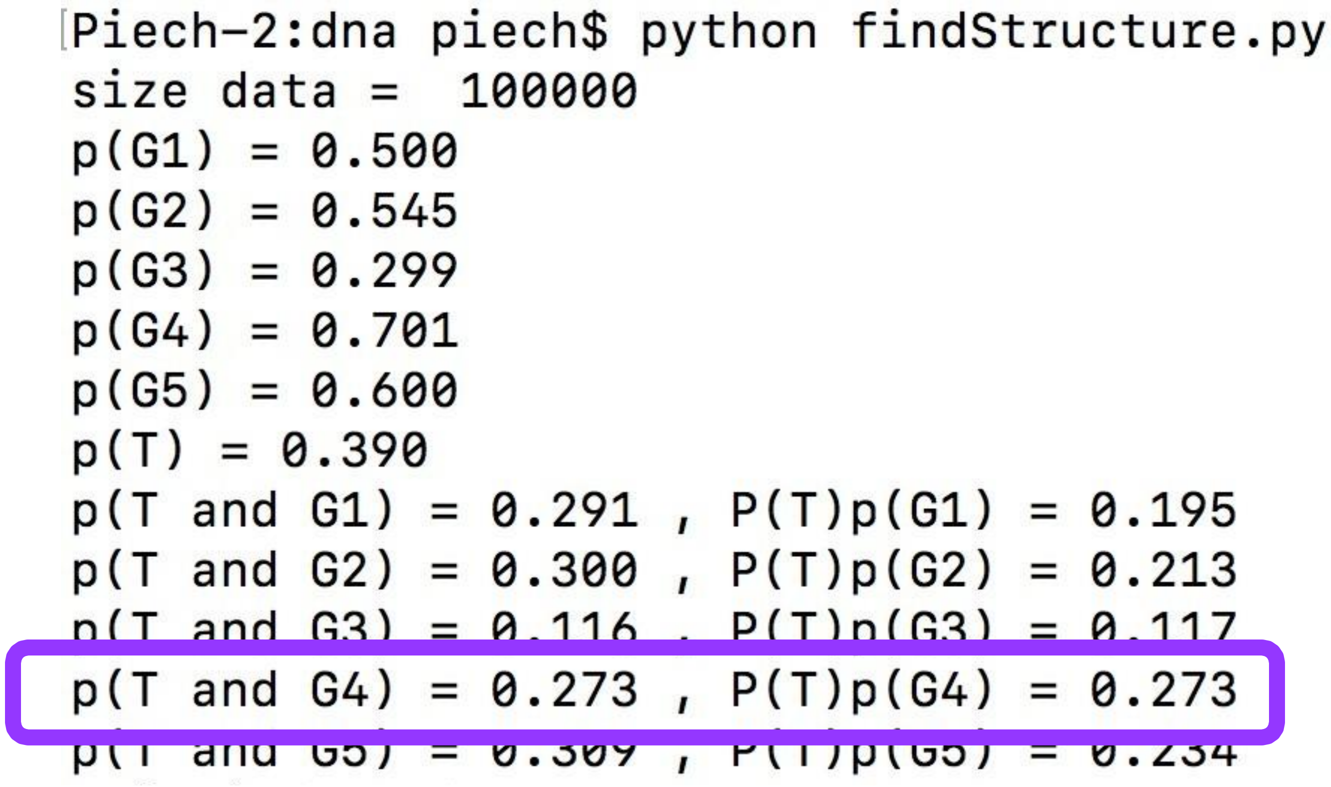


```
[Piech-2:dna piech$ python findStructure.py
size data =  100000
p(G1) = 0.500
p(G2) = 0.545
p(G3) = 0.299
p(G4) = 0.701
p(G5) = 0.600
p(T) = 0.390
p(T and G1) = 0.291 , P(T)p(G1) = 0.195
p(T and G2) = 0.300 , P(T)p(G2) = 0.213
p(T and G3) = 0.116 , P(T)p(G3) = 0.117
p(T and G4) = 0.273 , P(T)p(G4) = 0.273
p(T and G5) = 0.309 , P(T)p(G5) = 0.234
```

...

```
p(T and G5 | G2) = 0.450
p(T | G2)p(G5 | G2) = 0.450
```

Discovered Pattern

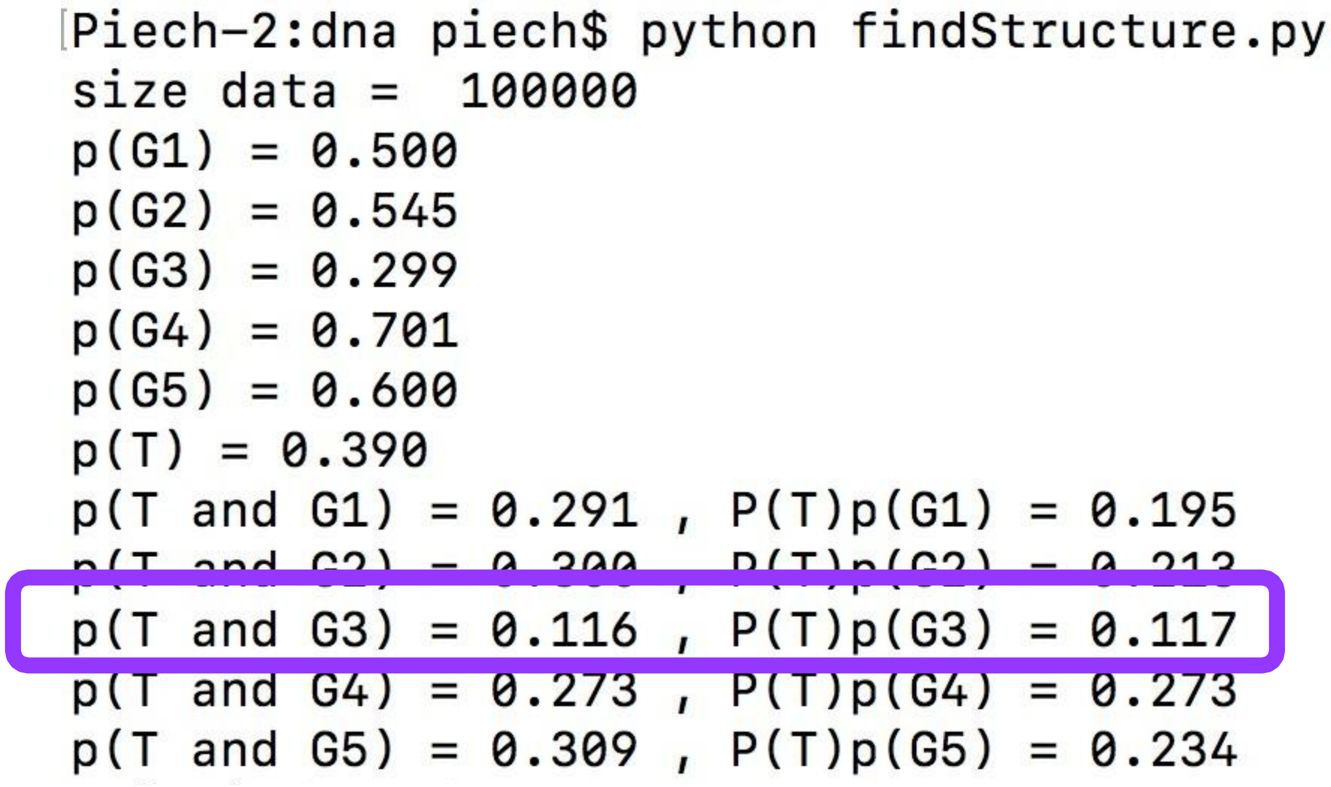


```
Piech-2:dna piech$ python findStructure.py
size data =  100000
p(G1) = 0.500
p(G2) = 0.545
p(G3) = 0.299
p(G4) = 0.701
p(G5) = 0.600
p(T) = 0.390
p(T and G1) = 0.291 , P(T)p(G1) = 0.195
p(T and G2) = 0.300 , P(T)p(G2) = 0.213
p(T and G3) = 0.116 , P(T)p(G3) = 0.117
p(T and G4) = 0.273 , P(T)p(G4) = 0.273
p(T and G5) = 0.309 , P(T)p(G5) = 0.234
```

...

```
p(T and G5 | G2) = 0.450
p(T | G2)p(G5 | G2) = 0.450
```

Discovered Pattern

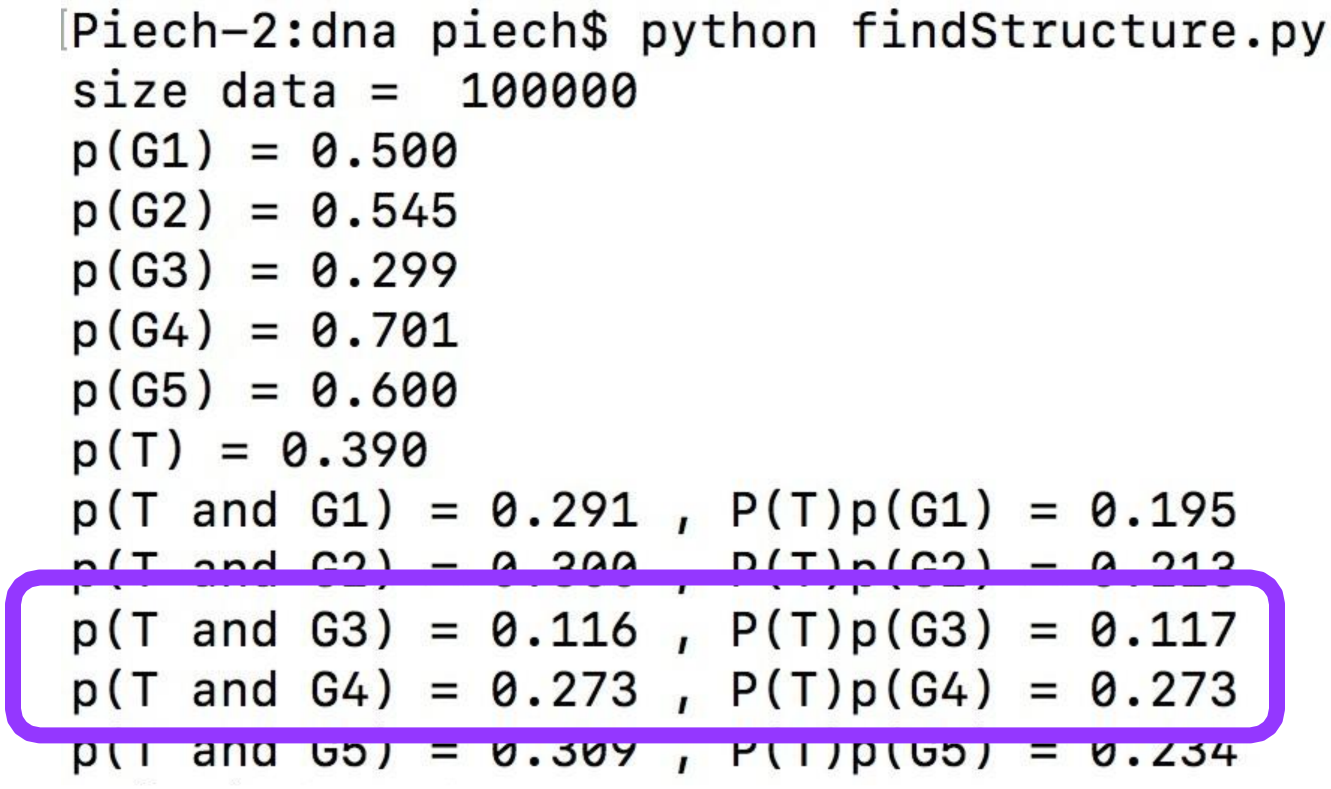


```
Piech-2:dna piech$ python findStructure.py
size data =  100000
p(G1) = 0.500
p(G2) = 0.545
p(G3) = 0.299
p(G4) = 0.701
p(G5) = 0.600
p(T) = 0.390
p(T and G1) = 0.291 , P(T)p(G1) = 0.195
p(T and G2) = 0.300 , P(T)p(G2) = 0.213
p(T and G3) = 0.116 , P(T)p(G3) = 0.117
p(T and G4) = 0.273 , P(T)p(G4) = 0.273
p(T and G5) = 0.309 , P(T)p(G5) = 0.234
```

...

```
p(T and G5 | G2) = 0.450
p(T | G2)p(G5 | G2) = 0.450
```

Discovered Pattern

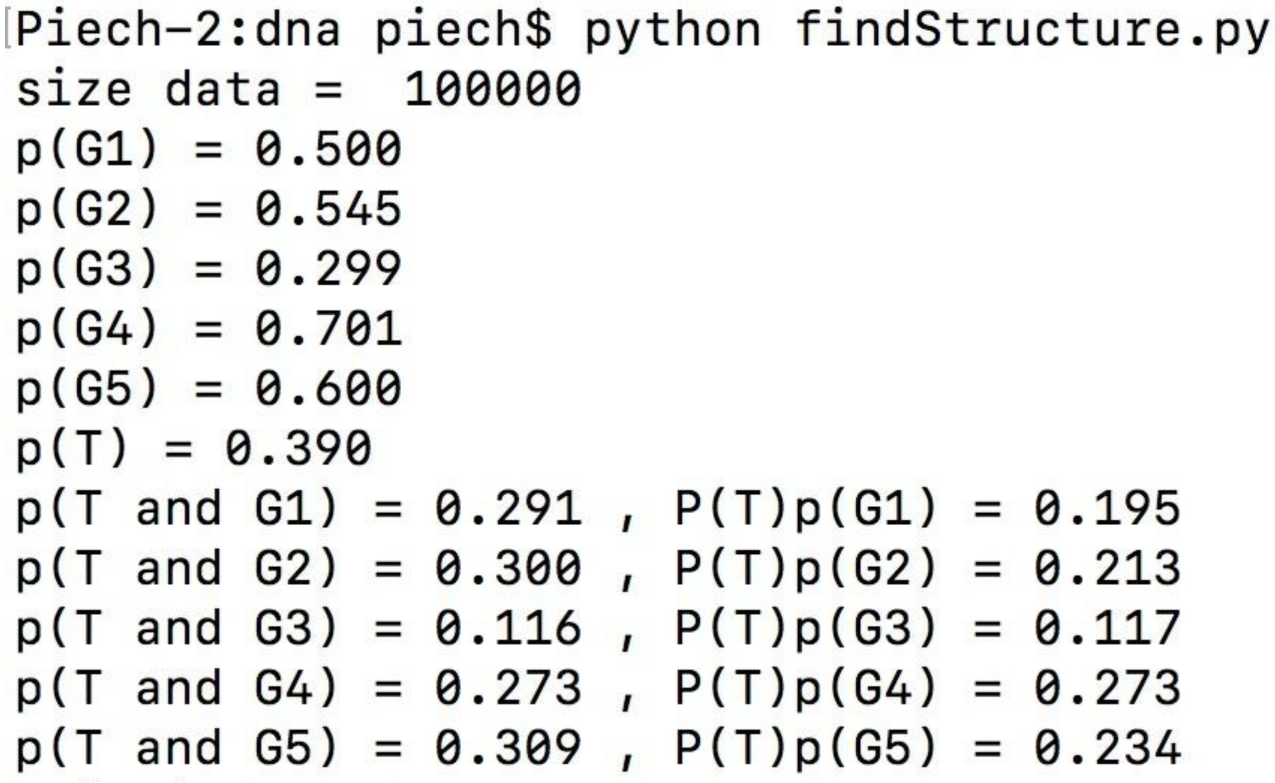


```
Piech-2:dna piech$ python findStructure.py
size data =  100000
p(G1) = 0.500
p(G2) = 0.545
p(G3) = 0.299
p(G4) = 0.701
p(G5) = 0.600
p(T) = 0.390
p(T and G1) = 0.291 , P(T)p(G1) = 0.195
p(T and G2) = 0.300 , P(T)p(G2) = 0.213
p(T and G3) = 0.116 , P(T)p(G3) = 0.117
p(T and G4) = 0.273 , P(T)p(G4) = 0.273
p(T and G5) = 0.309 , P(T)p(G5) = 0.234
```

...

```
p(T and G5 | G2) = 0.450
p(T | G2)p(G5 | G2) = 0.450
```

Discovered Pattern



```
Piech-2:dna piech$ python findStructure.py
size data =  100000
p(G1) = 0.500
p(G2) = 0.545
p(G3) = 0.299
p(G4) = 0.701
p(G5) = 0.600
p(T) = 0.390
p(T and G1) = 0.291 , P(T)p(G1) = 0.195
p(T and G2) = 0.300 , P(T)p(G2) = 0.213
p(T and G3) = 0.116 , P(T)p(G3) = 0.117
p(T and G4) = 0.273 , P(T)p(G4) = 0.273
p(T and G5) = 0.309 , P(T)p(G5) = 0.234
```

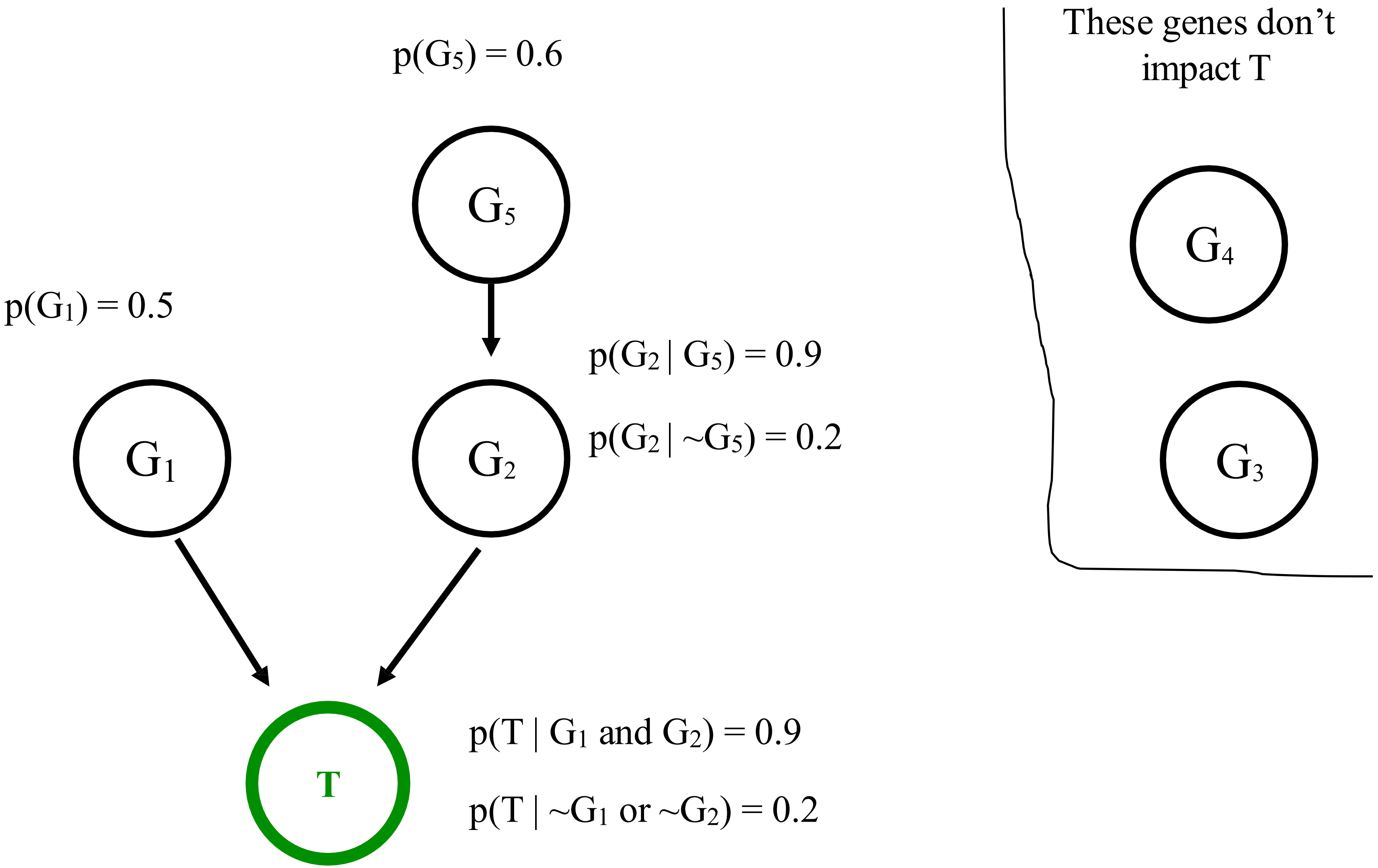
...

```
p(T and G5 | G2) = 0.450
p(T | G2)p(G5 | G2) = 0.450
```

Independence relationships can change with conditioning.

If E and F are independent, that does not mean they will still be independent given another event G.

Only Causal Structure That Fits



$p(G_5) = 0.6$

$p(G_1) = 0.5$

$p(G_2 \mid G_5) = 0.9$

$p(G_2 \mid \sim G_5) = 0.2$

$p(T \mid G_1 \text{ and } G_2) = 0.9$

$p(T \mid \sim G_1 \text{ or } \sim G_2) = 0.2$

These genes don’t impact T

See you Friday