

Independence

Chris Piech, CS109

Today, start with a cool program

G_1

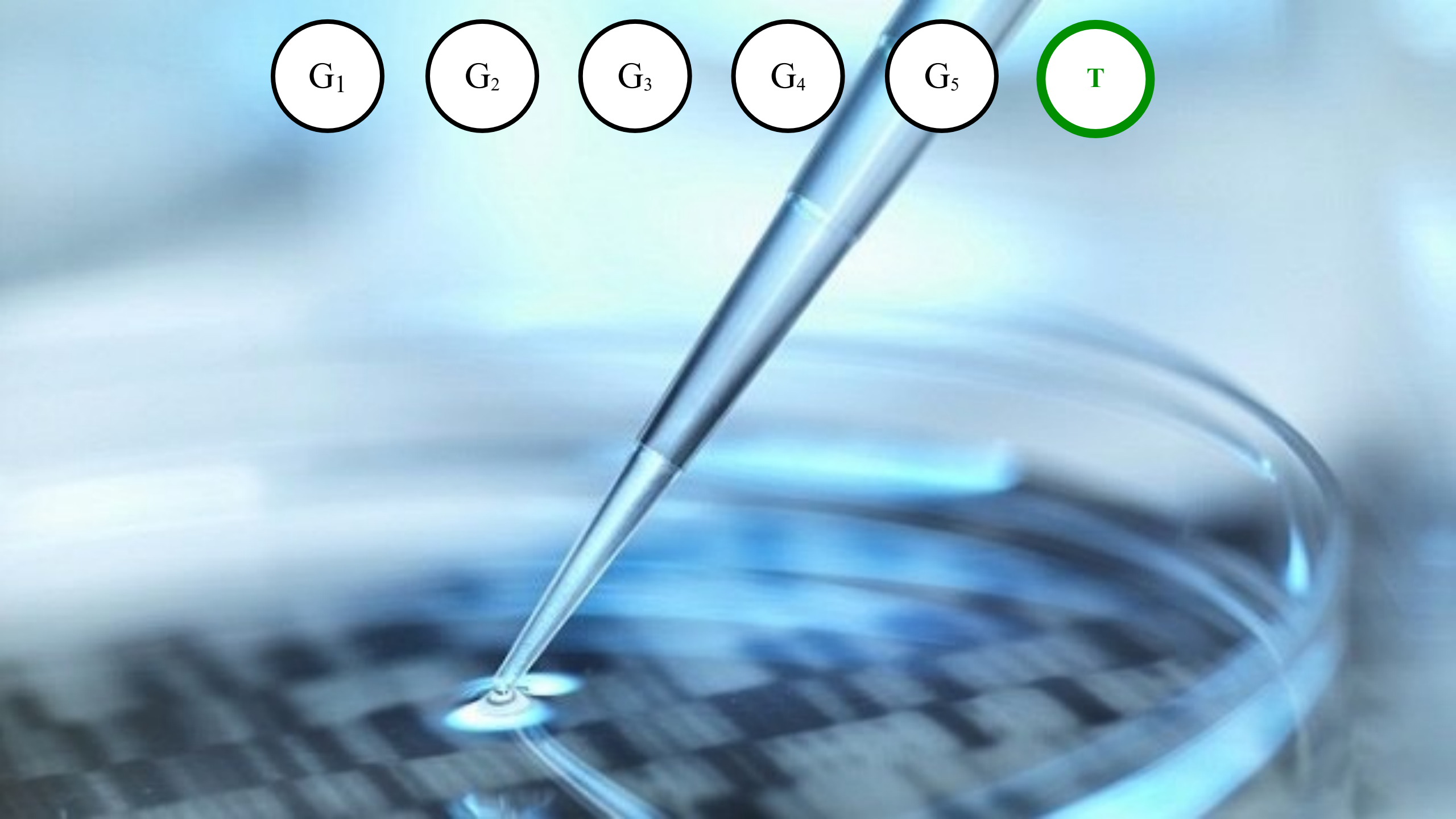
G_2

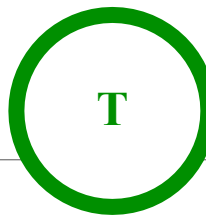
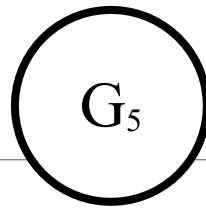
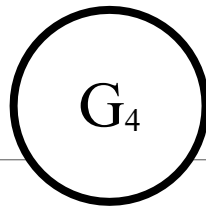
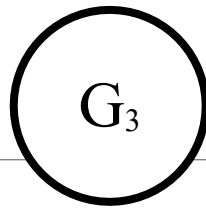
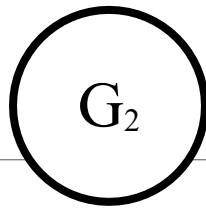
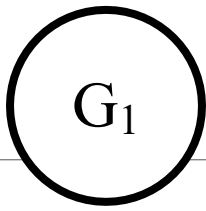
G_3

G_4

G_5

T





```

dna.txt — dna
dna.txt
1 False, True, False, False, True, False
2 True, True, False, True, True, False
3 True, True, False, True, True, True
4 False, True, False, True, True, False
5 False, True, False, False, True, False
6 True, True, False, True, True, True
7 False, False, True, False, False, False
8 False, False, True, False, True, False
9 True, False, False, True, False, False
10 False, True, False, True, True, False
11 True, False, False, True, False, False
12 True, False, True, True, False, False
13 False, True, False, False, True, False
14 False, False, True, True, False, False
15 True, True, False, False, True, True
16 True, False, True, True, False, False
17 True, True, True, True, True, True |
18 True, False, True, False, False, True
19 False, True, False, True, True, True
20 False, False, True, False, False, False
21 False, False, False, True, True, False
22 False, True, False, False, True, False
23 True, True, False, True, True, True
24 False, True, False, True, True, False
25 True, False, False, False, False, True
26 False, False, True, True, False, True
27 False, False, False, True, False, False
28 False, True, True, False, False, True
29 False, True, False, False, True, True
30 False, False, False, False, False, True
31 False, True, False, True, True, False
32 True, False, False, True, False, False
33 True, True, False, True, True, True
34 True, True, False, False, True, True
35 True, True, False, True, True, True
36 False, False, False, True, False, False
--

```

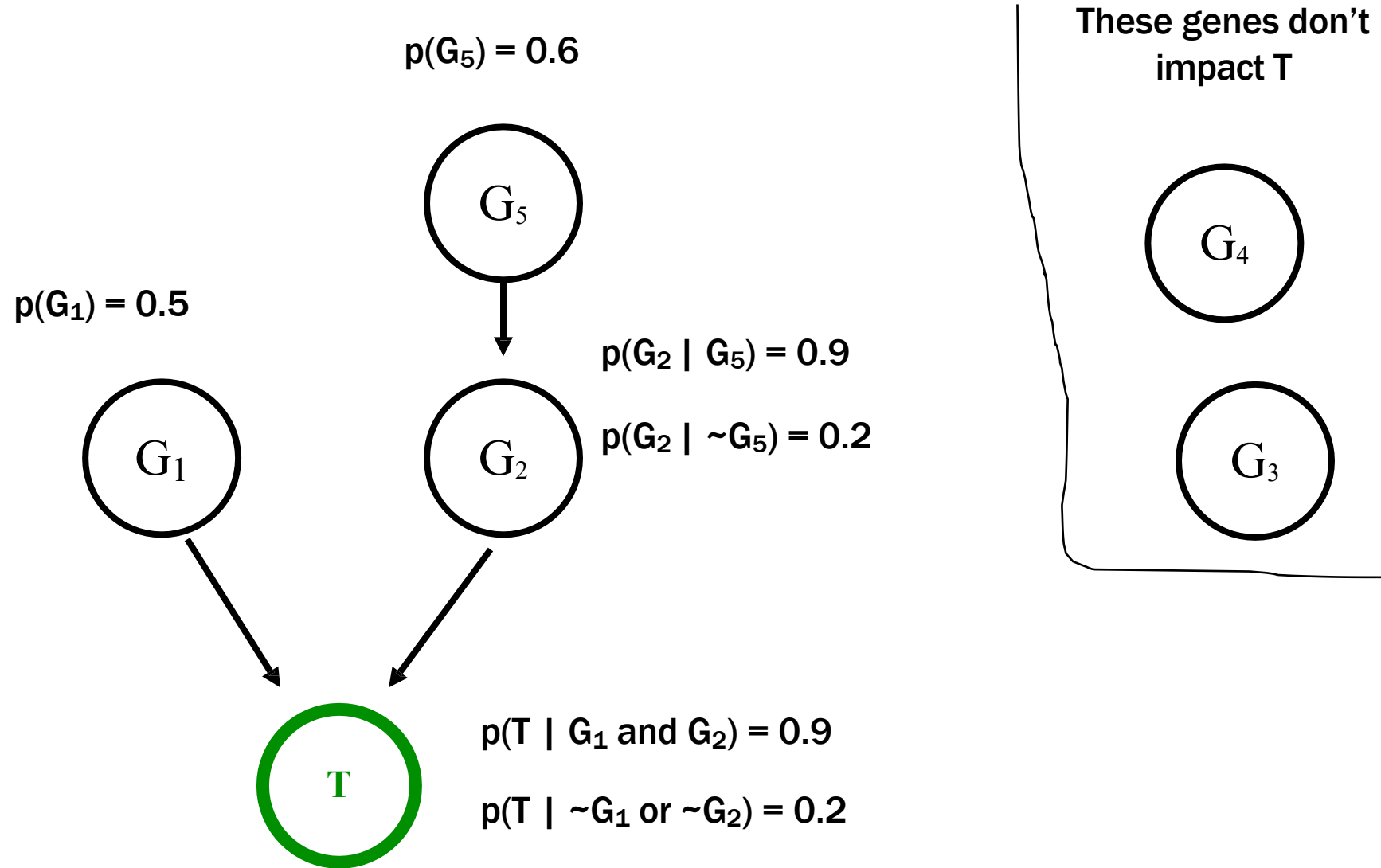


100,000 samples

6 observations per sample



Discovered Hypothesis



We've gotten ahead of ourselves



Source: The Hobbit

Start at the beginning



Source: The Hobbit

Review

Review: Conditional Probability

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

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

Review: Chain Rule

Definition of conditional probability:

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

The Chain Rule:

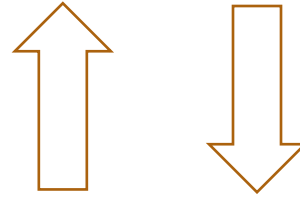
$$P(EF) = P(E|F)P(F)$$

Relationship Between Probabilities



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

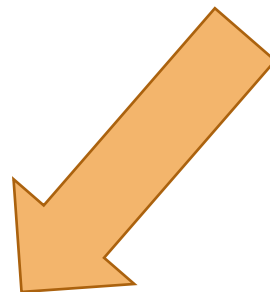
Chain rule
(Product rule)



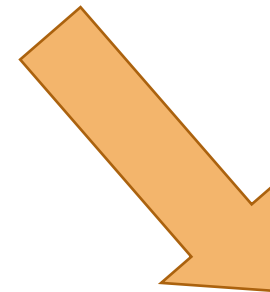
Definition of
conditional probability

$$P(E|F)$$

Law of Total
Probability

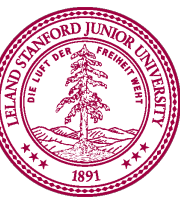


Bayes'
Theorem



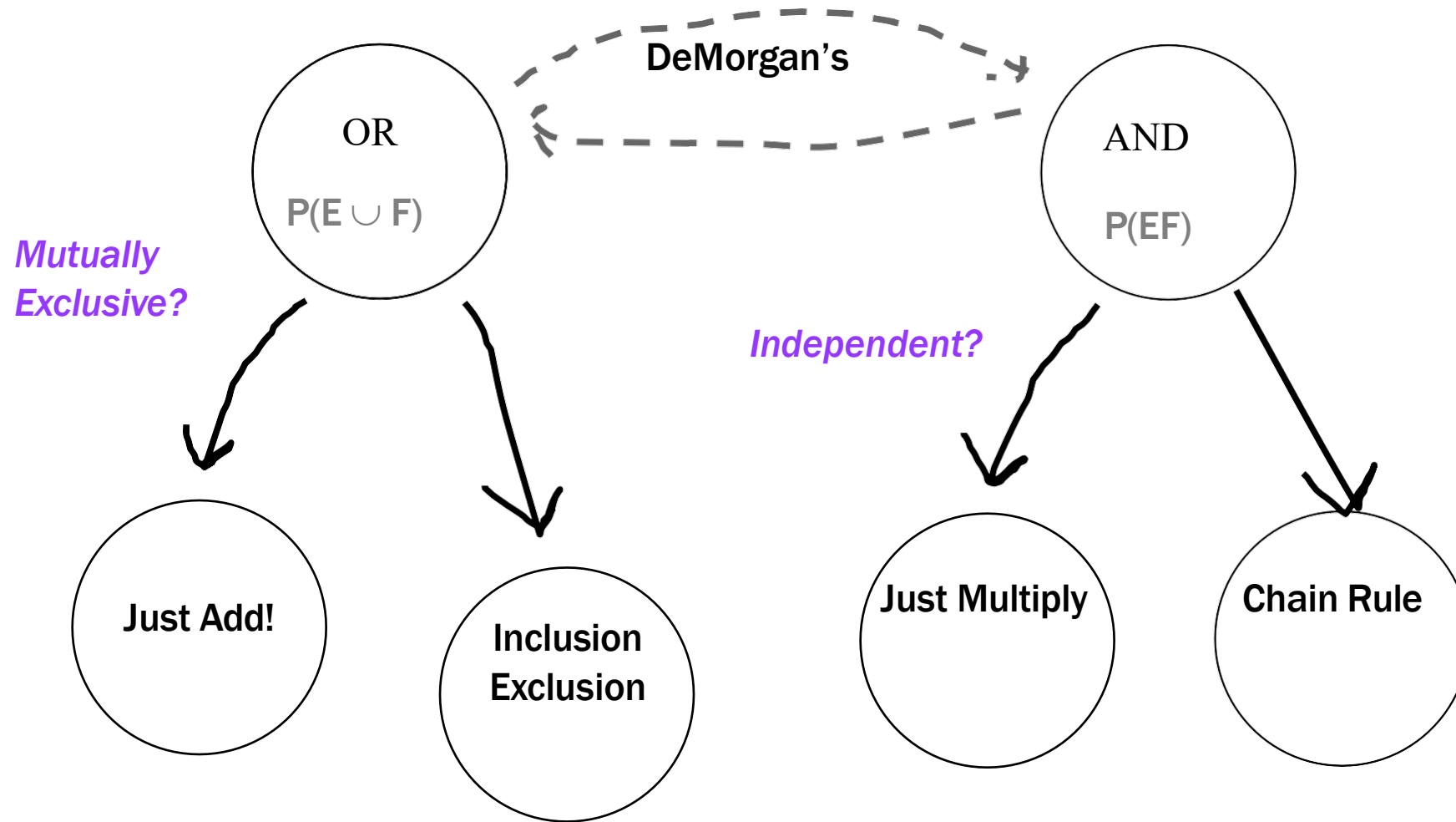
$$P(E)$$

$$P(F|E)$$

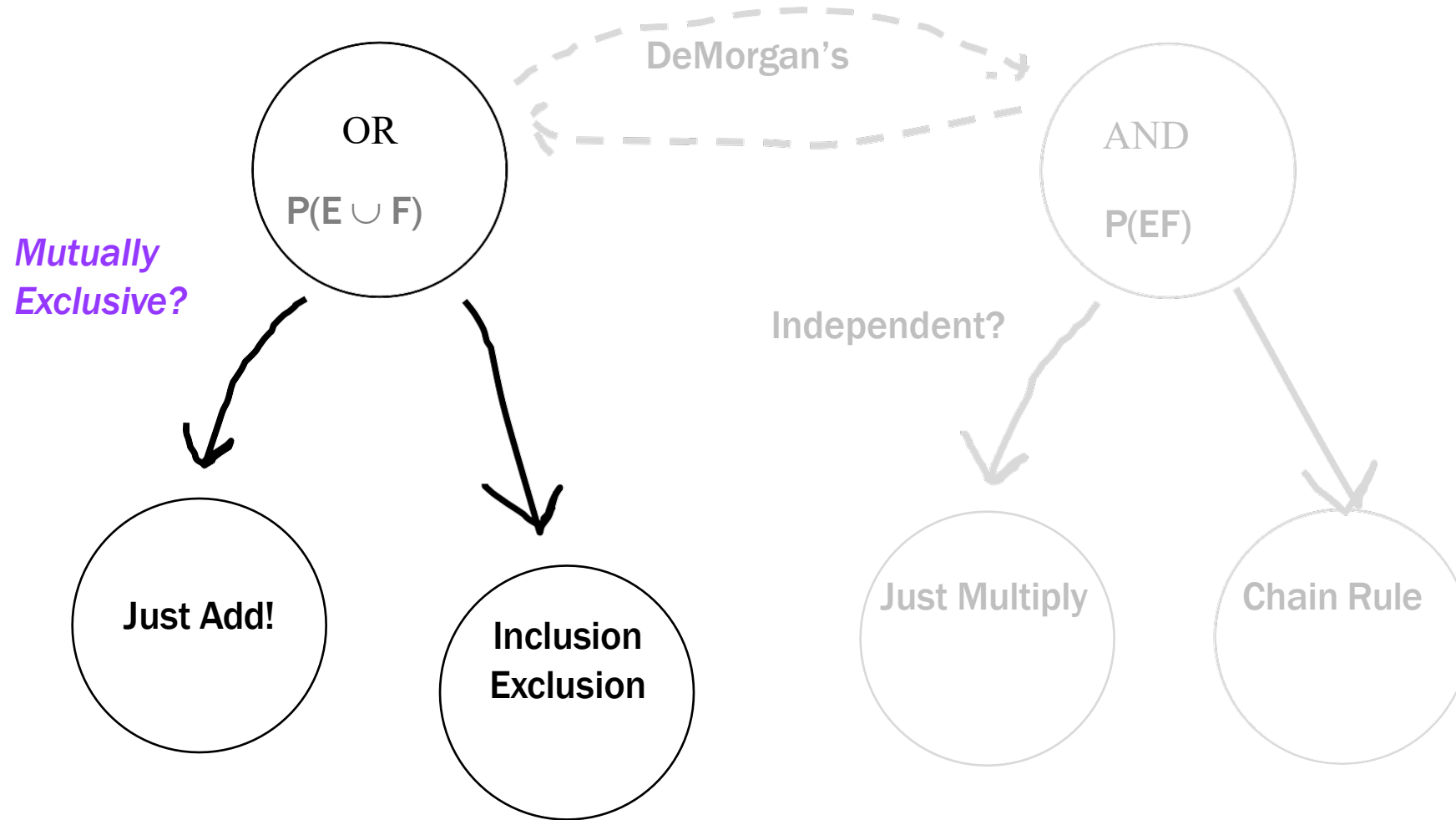


End Review

Today

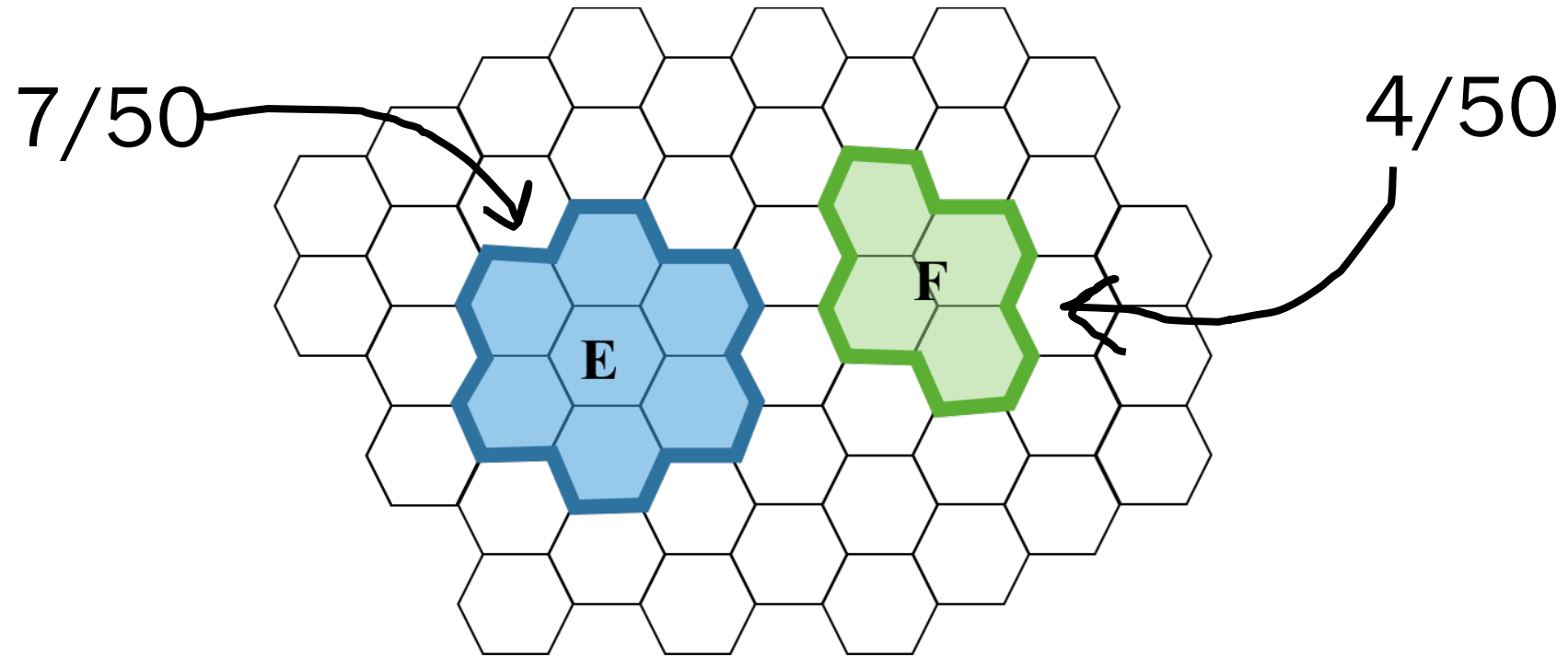


Today



Probability of “OR”

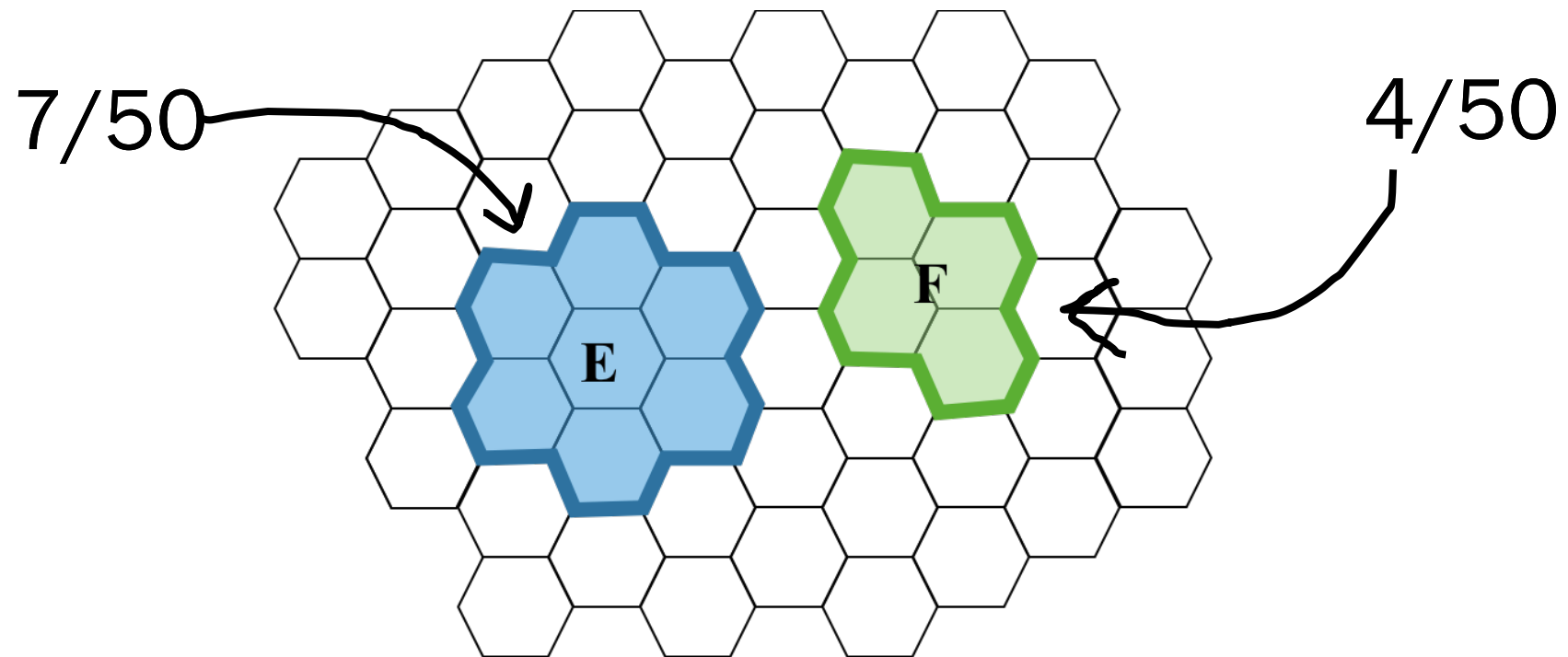
Review: OR with Mutually Exclusive Events



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

$$P(E \cup F) = P(E) + P(F)$$

Review: OR with Mutually Exclusive Events

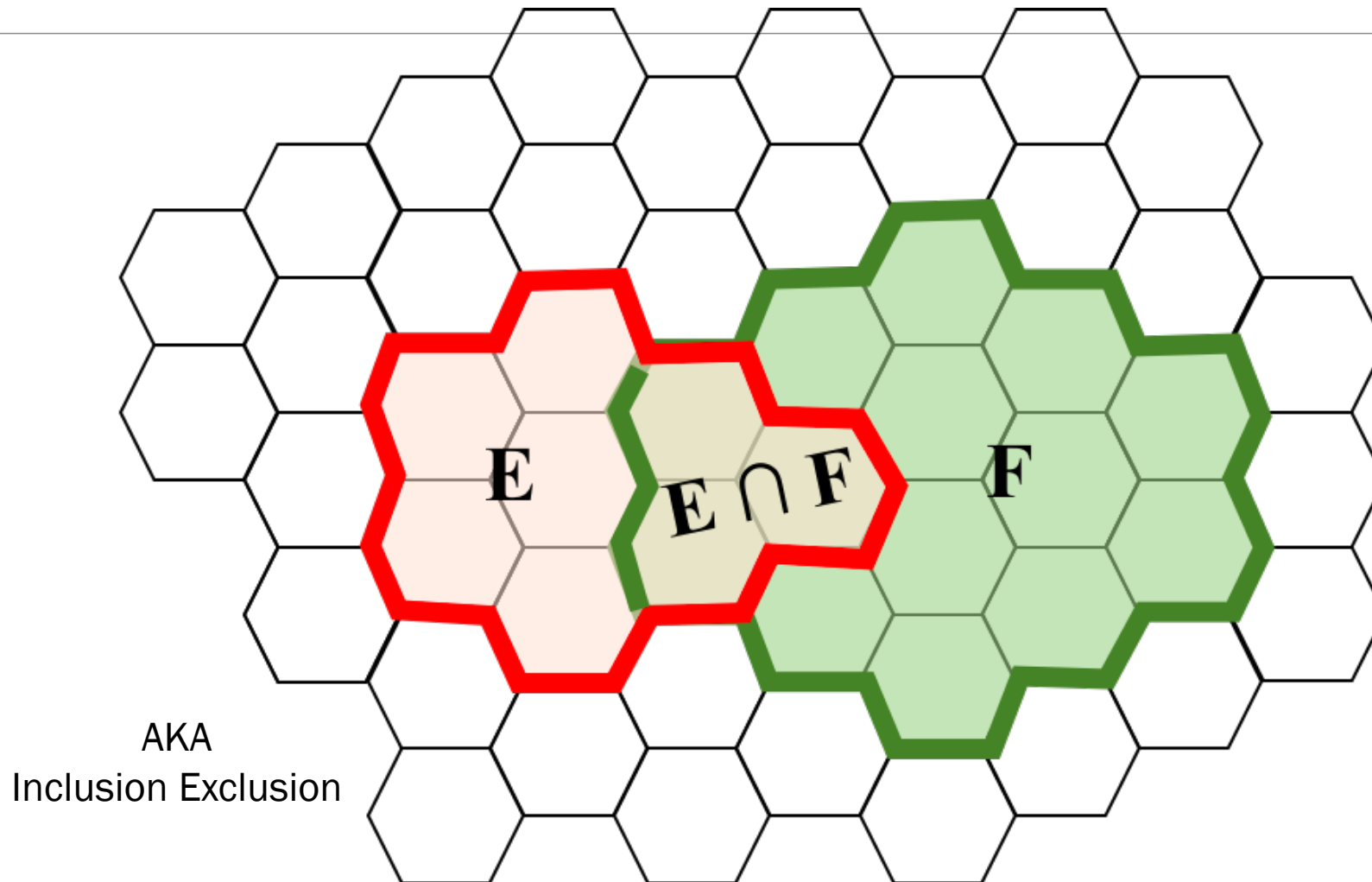


If events are mutually exclusive, probability of OR is simple:

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

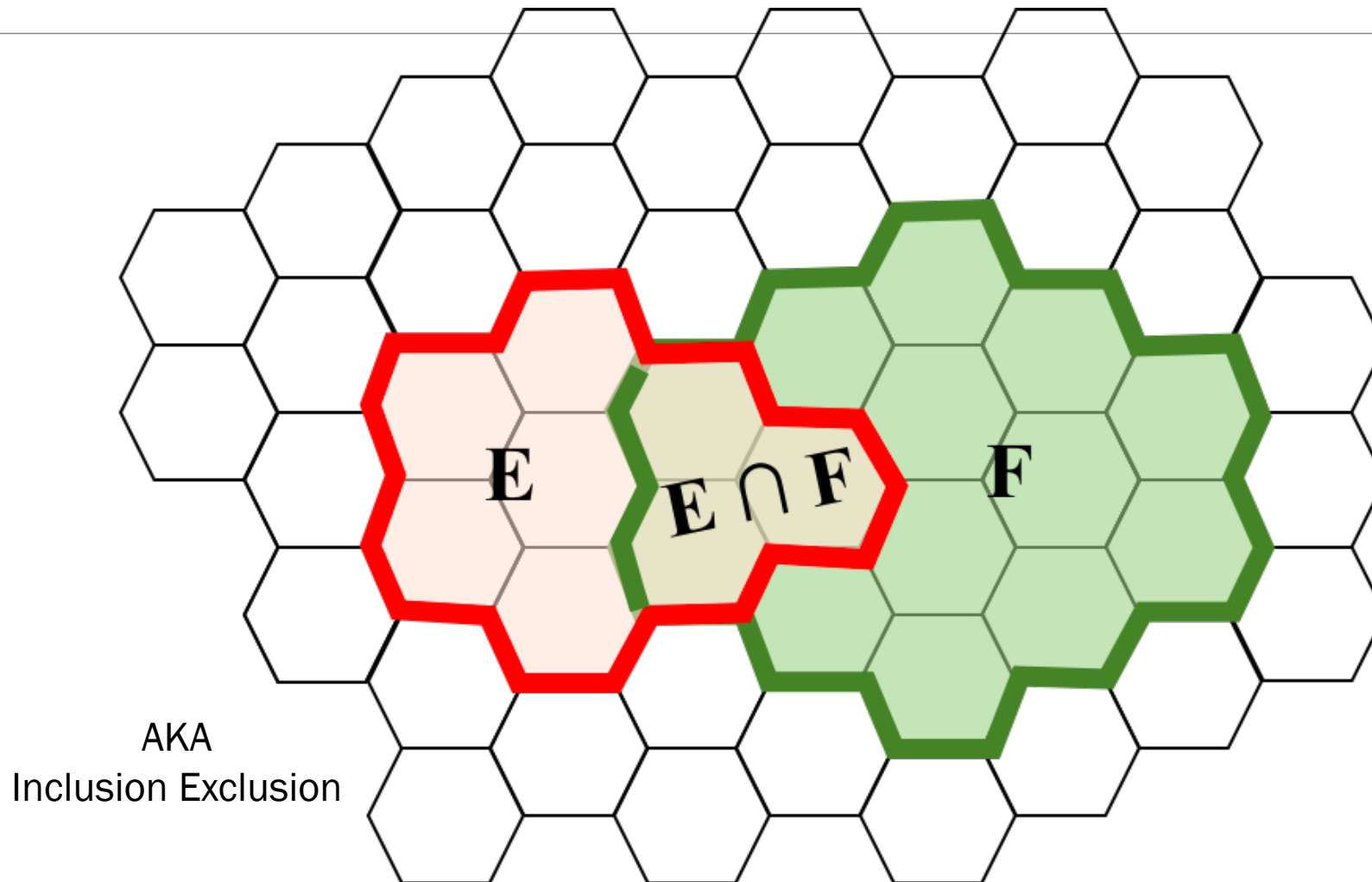
What about when they are not
Mutually exclusive?

OR *without* Mutually Exclusive Events



$$P(E \cup F) = P(E) + P(F) - P(EF)$$

OR *without* Mutually Exclusive Events

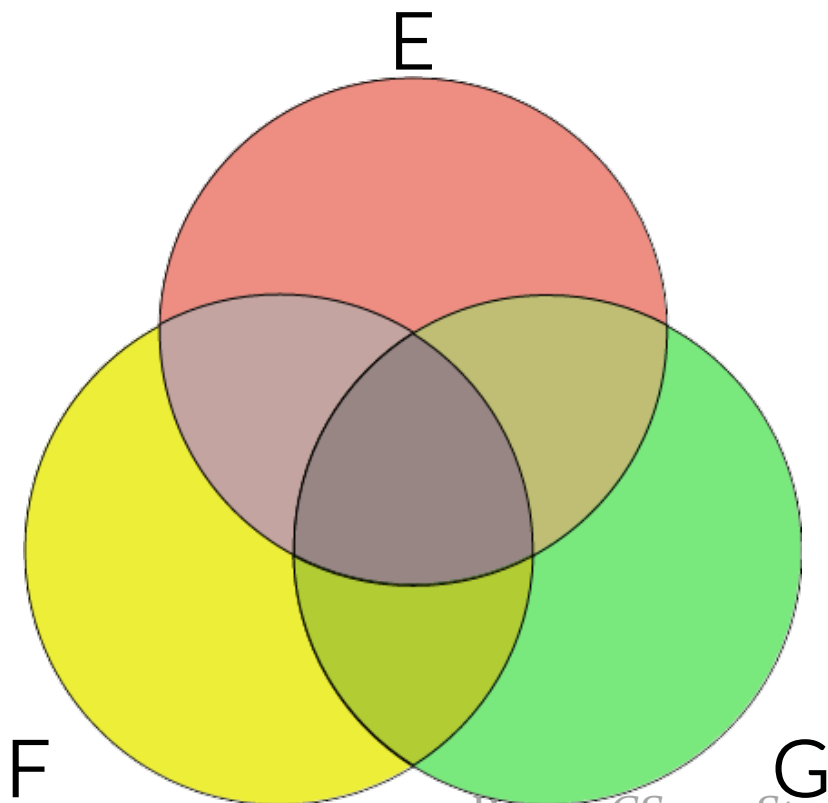


$$P(E \cup F) = \frac{8}{50} + \frac{14}{50} - \frac{3}{50} = \frac{19}{50}$$

More than two sets?

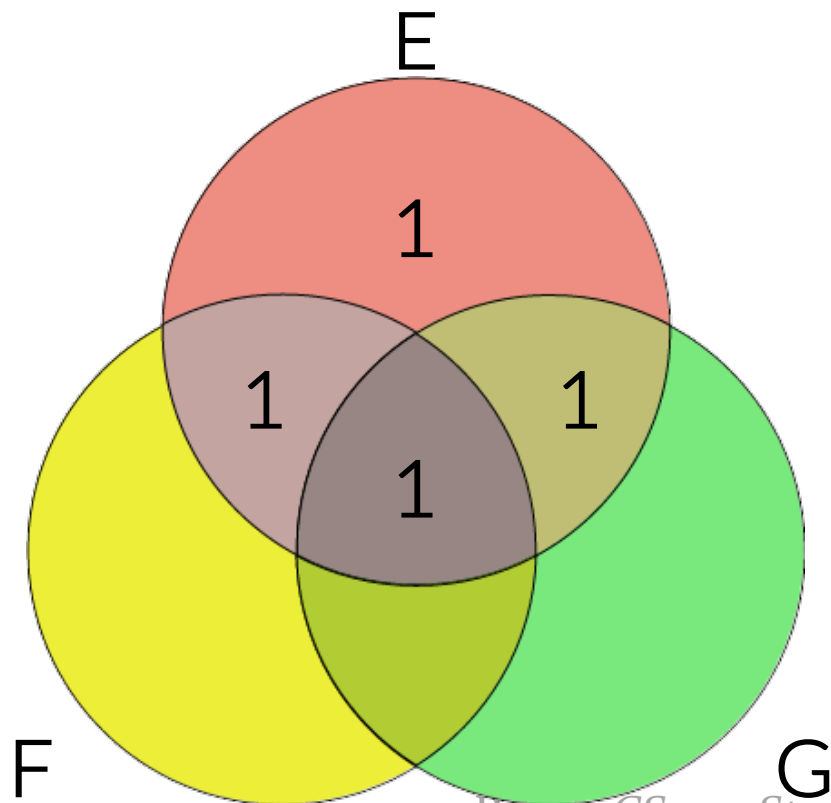
Inclusion / Exclusion with Three Events

$$P(E \cup F \cup G) =$$



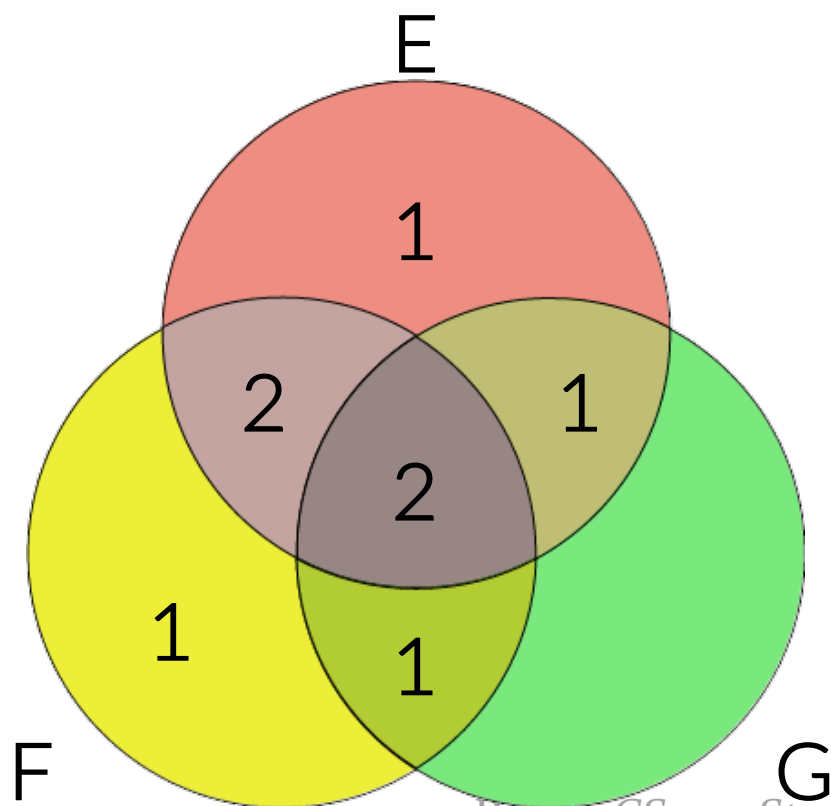
Inclusion / Exclusion with Three Events

$$P(E \cup F \cup G) = P(E)$$



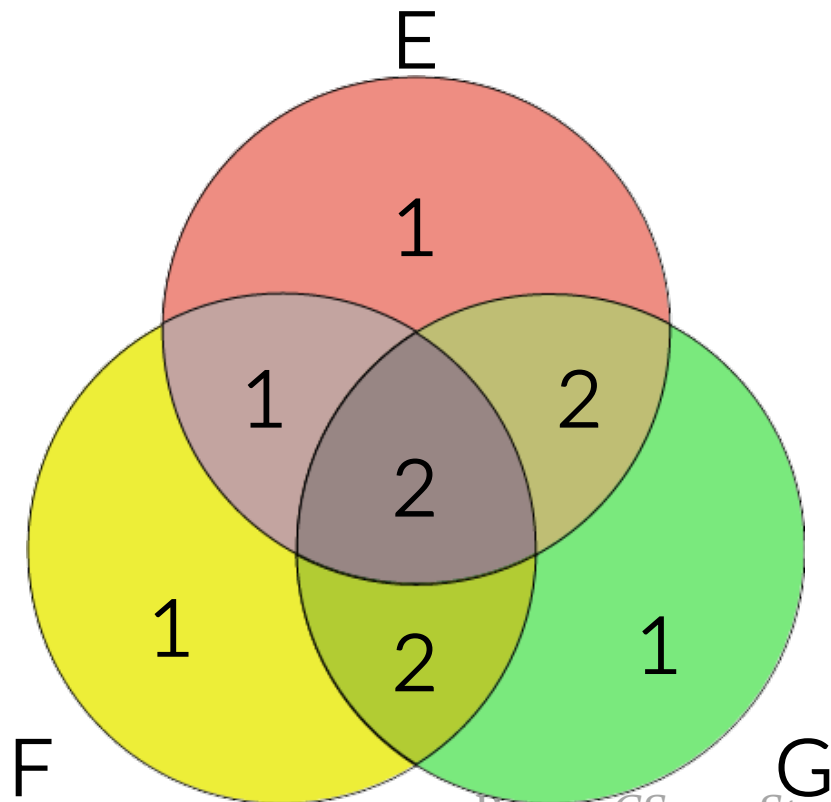
Inclusion / Exclusion with Three Events

$$P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(E \cap F) - P(E \cap G) - P(F \cap G) + P(E \cap F \cap G)$$



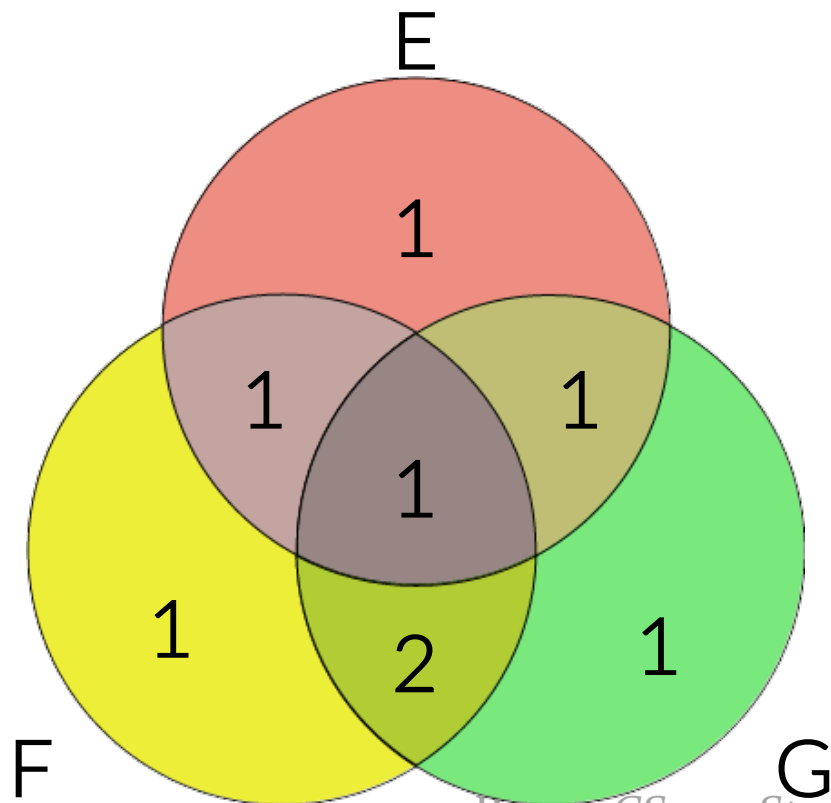
Inclusion / Exclusion with Three Events

$$P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(EF)$$



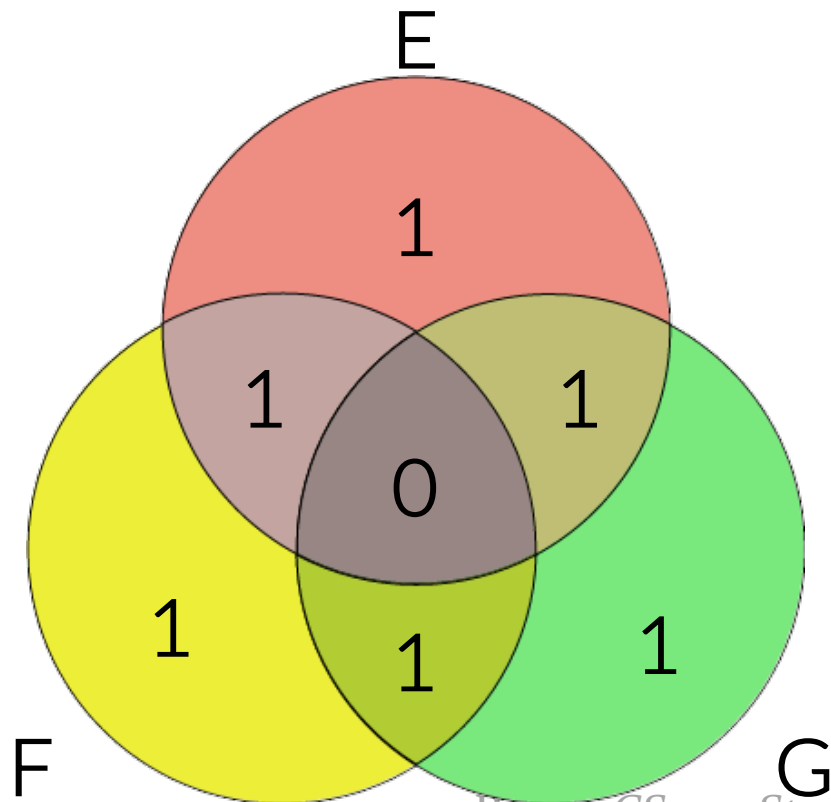
Inclusion / Exclusion with Three Events

$$P(E \cup F \cup G) = P(E) + P(F) + P(G) \\ - P(EF) - P(EG)$$



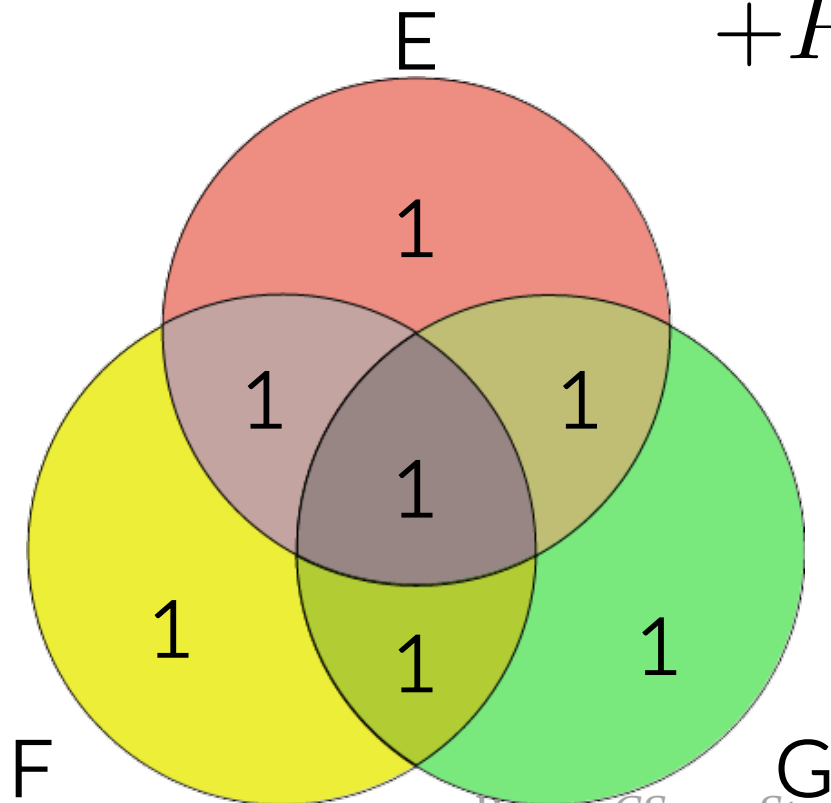
Inclusion / Exclusion with Three Events

$$P(E \cup F \cup G) = P(E) + P(F) + P(G) \\ - P(EF) - P(EG) - P(FG)$$



Inclusion / Exclusion with Three Events

$$\begin{aligned} P(E \cup F \cup G) &= P(E) + P(F) + P(G) \\ &\quad - P(EF) - P(EG) - P(FG) \\ &\quad + P(EFG) \end{aligned}$$



General Inclusion / Exclusion

$$P(E_1 \cup E_2 \cup \dots \cup E_n) = \sum_{r=1}^n (-1)^{r+1} Y_r$$

Y_1 = Sum of all events on their own

$$\sum_i P(E_i)$$

Y_2 = Sum of all pairs of events

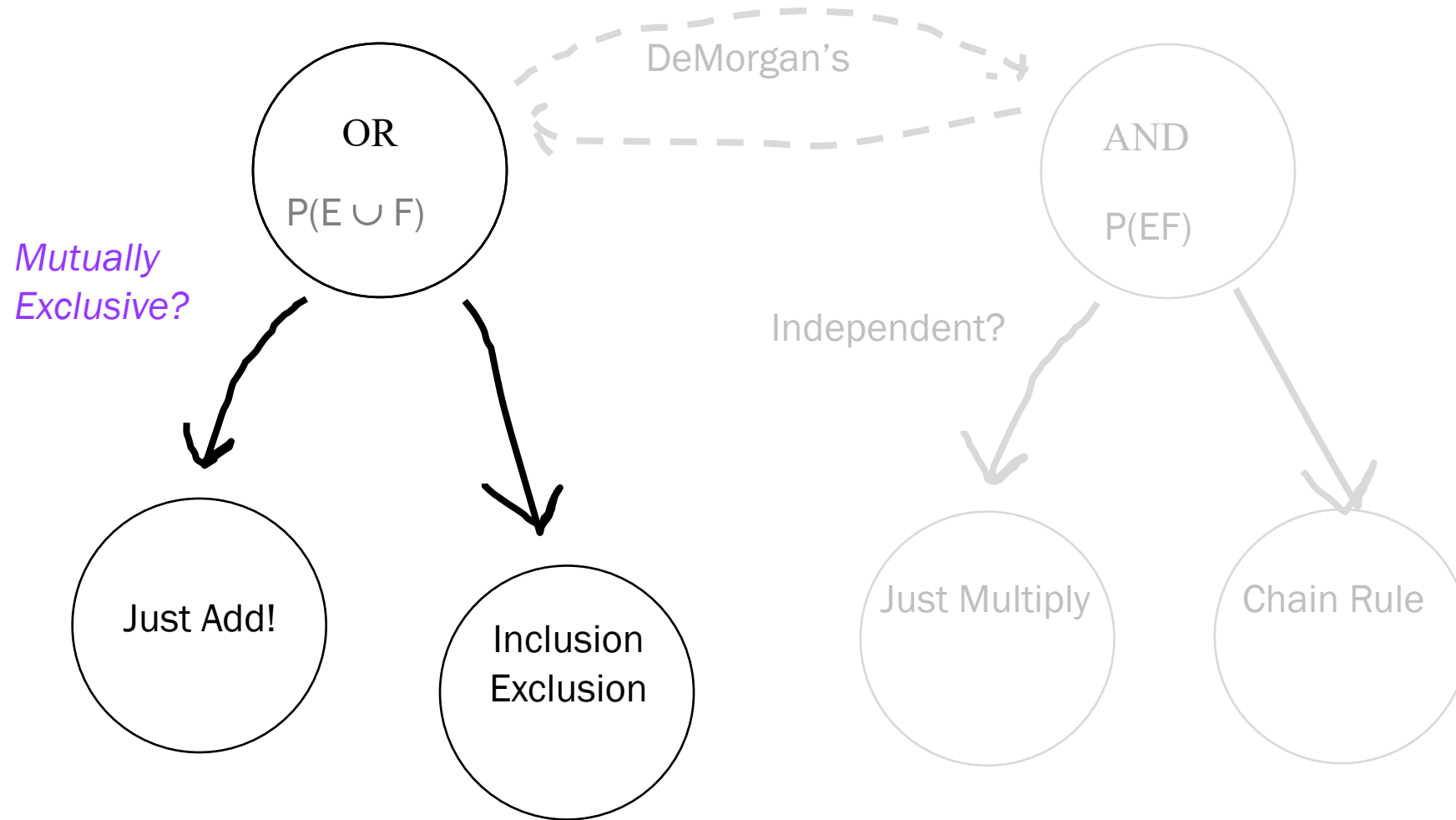
$$\sum_{i,j \text{ s.t. } i \neq j} P(E_i \cap E_j)$$

Y_3 = Sum of all triples of events

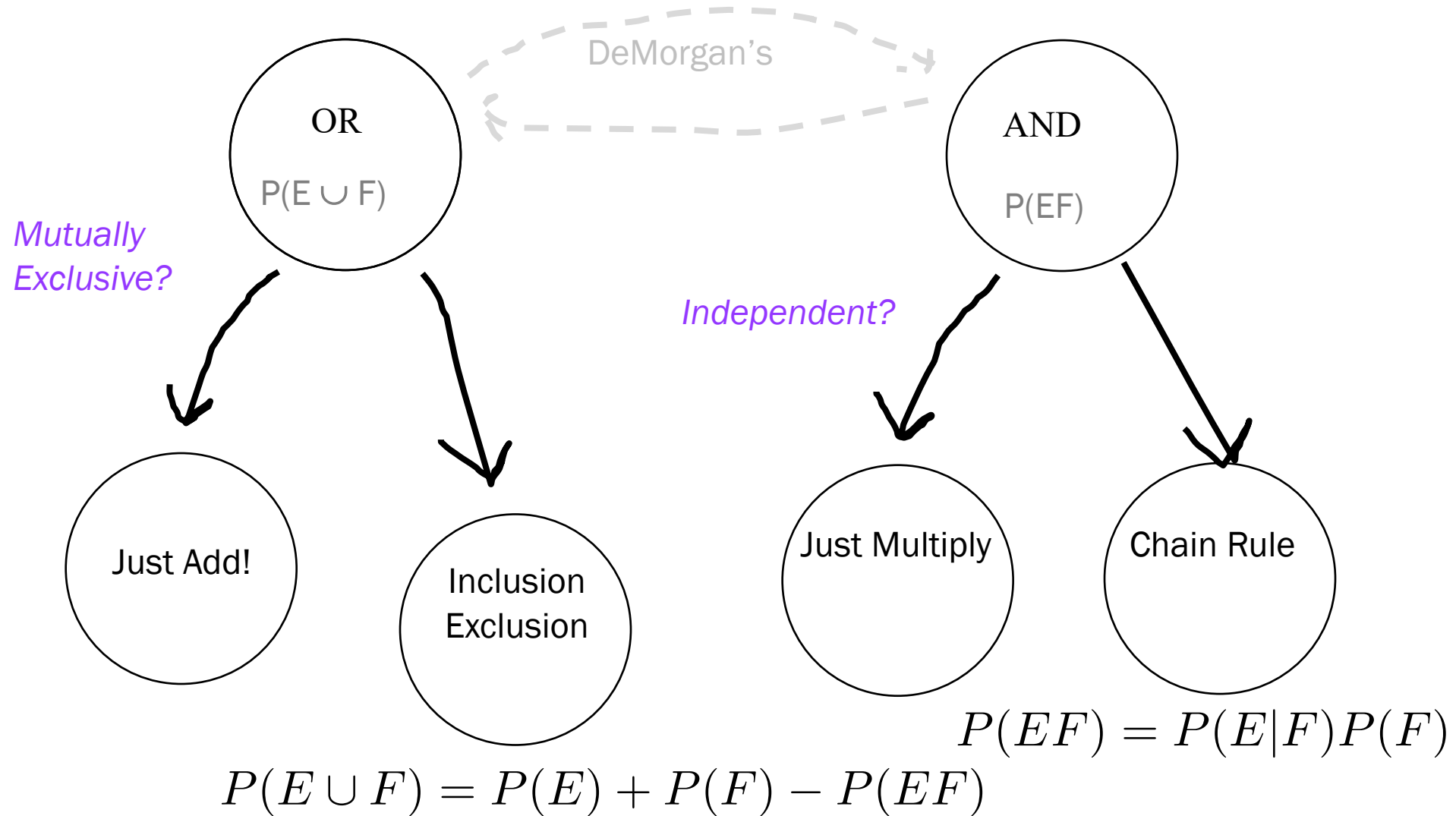
$$\sum_{i,j,k \text{ s.t. } i \neq j, j \neq k, i \neq k} P(E_i \cap E_j \cap E_k)$$

* Where Y_r is the sum, for all combinations of r events, of the probability of the union those events.

Today



Today



Probability of “AND”

WE THE PEOPLE
insure domestic Tranquility, provide for the common defence,
and our Posterity, do ordain and establish this Constitution

Article I. All legislative Powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and House of Representatives. The House of Representatives shall be composed of Members chosen every second Year by the People of the several States, and the Electors in each State shall have the Qualifications requisite for Electors of the most numerous Branch of the State Legislature. No Representative or Member of the House shall be, when elected, seven Years old, seven Years a Citizen of the United States, and seven Years a Citizen of that State in which he shall be chosen. And no Person shall be Representative of any State unless he shall, when elected, have seven Years residence in that State. The Electors in each State shall have the Qualifications requisite for Electors of the most numerous Branch of the State Legislature. No Senator or Representative shall, when elected, be less than thirty Years old and seven Years a Citizen of the United States, and three Years a Citizen of that State in which he shall be chosen. And no Person shall be Senator unless he shall, when elected, have fourteen Years residence in that State. The Electors in each State shall have the Qualifications requisite for Electors of the most numerous Branch of the State Legislature. The Senate shall be composed of two Senators from each State, chosen by the Electors in that State for six Years, and each Senator shall have the Qualifications requisite for Electors of the most numerous Branch of the State Legislature. The Senate shall, when first convened, not exceed one Year, and shall meet at such Time and Place as they shall determine. The Senate shall choose their Speaker and other Officers, and shall have the sole Power to impeach and try all Officers, civil and military, and Judges of the supreme and inferior Courts; and the Oath of Office shall be sworn to by each Senator. Two thirds of the Senate shall be a Quorum to do Business, and a Majority of that Quorum shall be necessary to pass any Bill. The Senate shall have the sole Power to confirm and reject all Appointments, and all Commissions shall be in the Name of the President of the United States, by and with the Advice and Consent of the Senate. The Senate shall have the sole Power to ratify and reject all Treaties, made by the President of the United States, which shall be valid, only when ratified by two thirds of the Senate. The Senate shall have the sole Power to confirm and reject all Appointments, and all Commissions shall be in the Name of the President of the United States, by and with the Advice and Consent of the Senate. The Senate shall have the sole Power to ratify and reject all Treaties, made by the President of the United States, which shall be valid, only when ratified by two thirds of the Senate.

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

Independence is reciprocal

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

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

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

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

$$= P(B)$$

Alternative Definition of Independence

$$\begin{aligned}P(A, B) &= P(A) \cdot P(B|A) \\ &= P(A) \cdot P(B)\end{aligned}$$

Probability of and

Since B is independent of A

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
Piech, CS109, Stanford University

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 $P(EF)$?

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

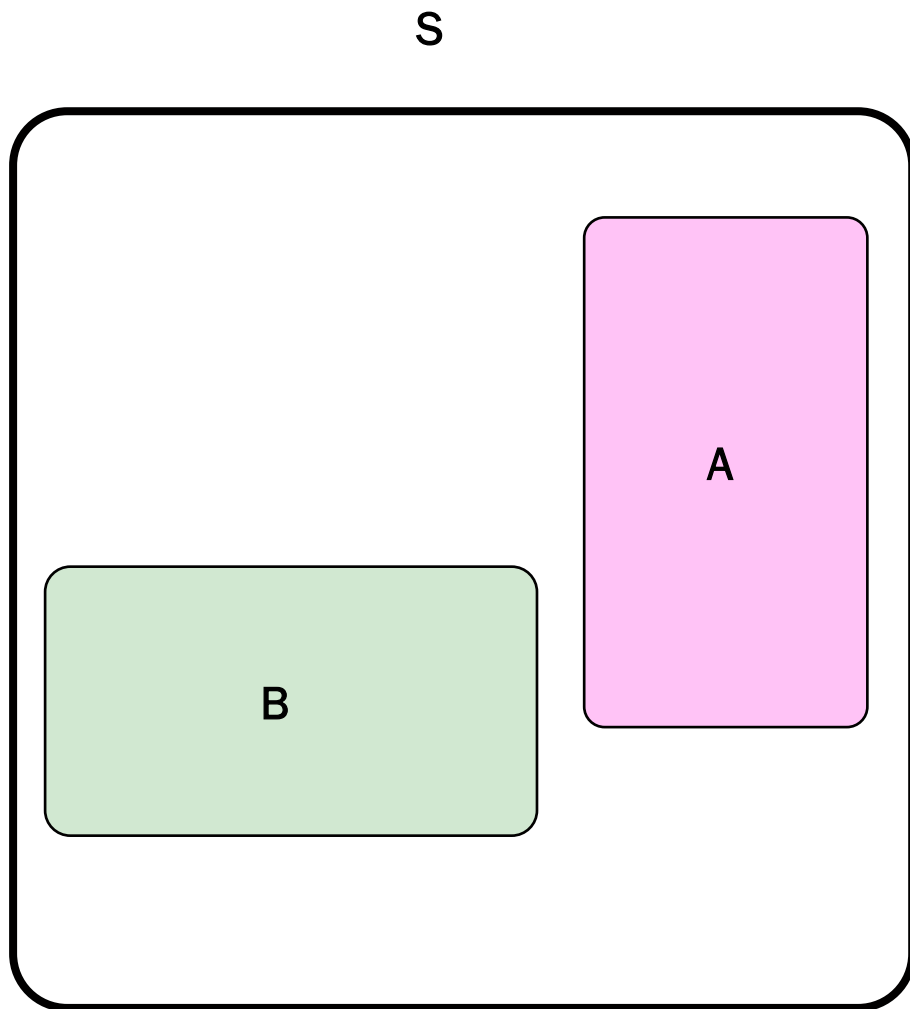
Let G be event: $D_1 + D_2 = 5$ $\{(1, 4), (2, 3), (3, 2), (4, 1)\}$

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

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

What does independence look like?

Independence

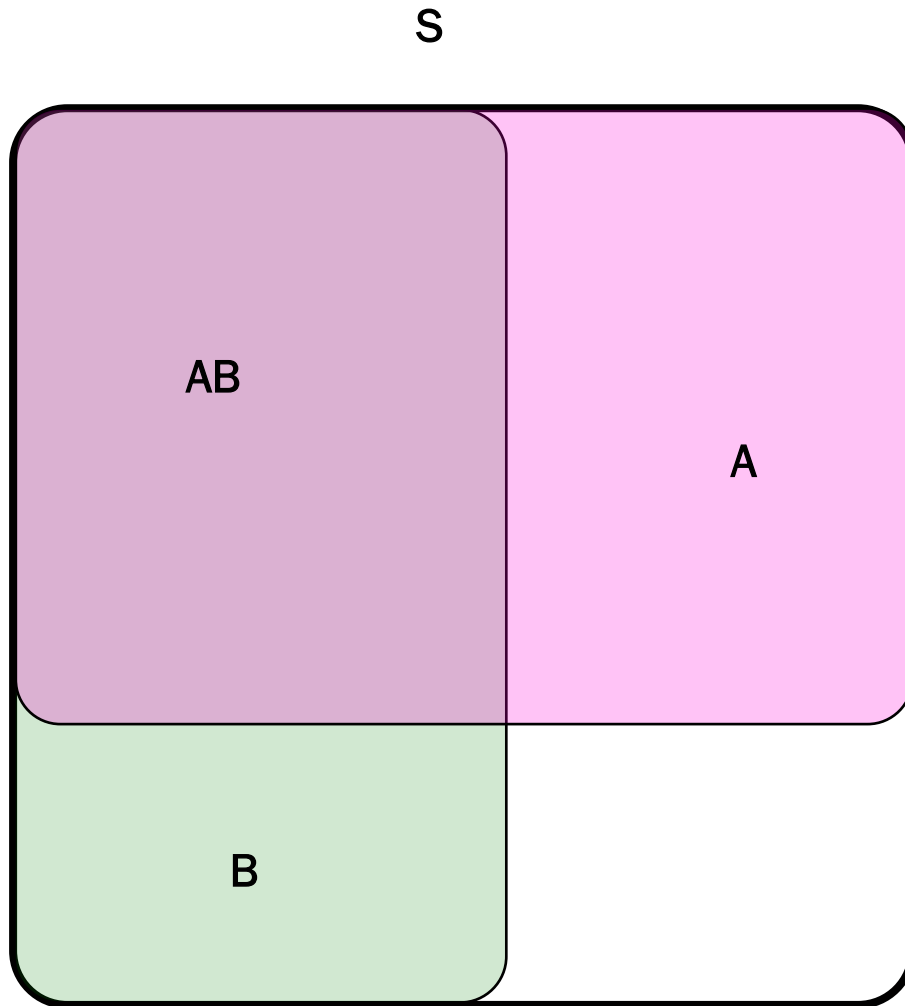


Independence Definition 1:

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

An arrow points from the $|AB|$ term in the second equation to the 0 in the first equation, indicating that $|AB| = 0$ because the sets A and B are disjoint.

Independence



Independence Definition 1:

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

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

Independence Definition 2:

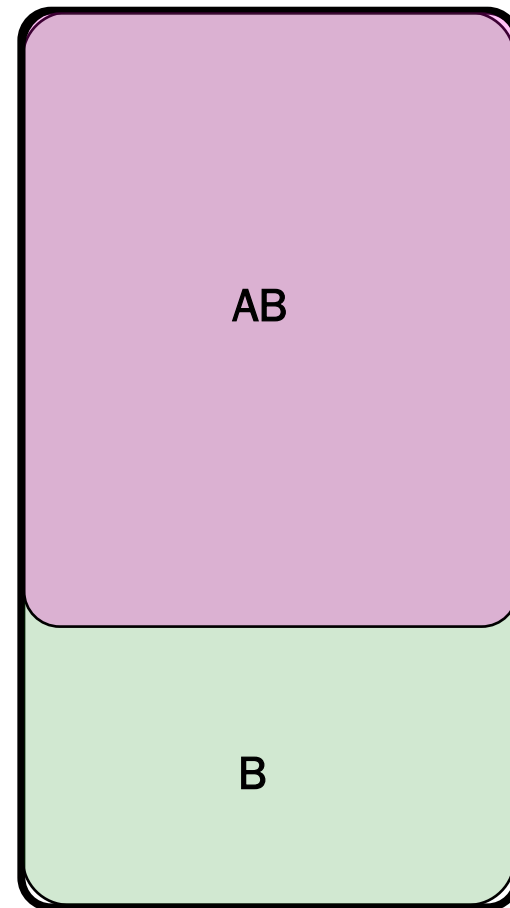
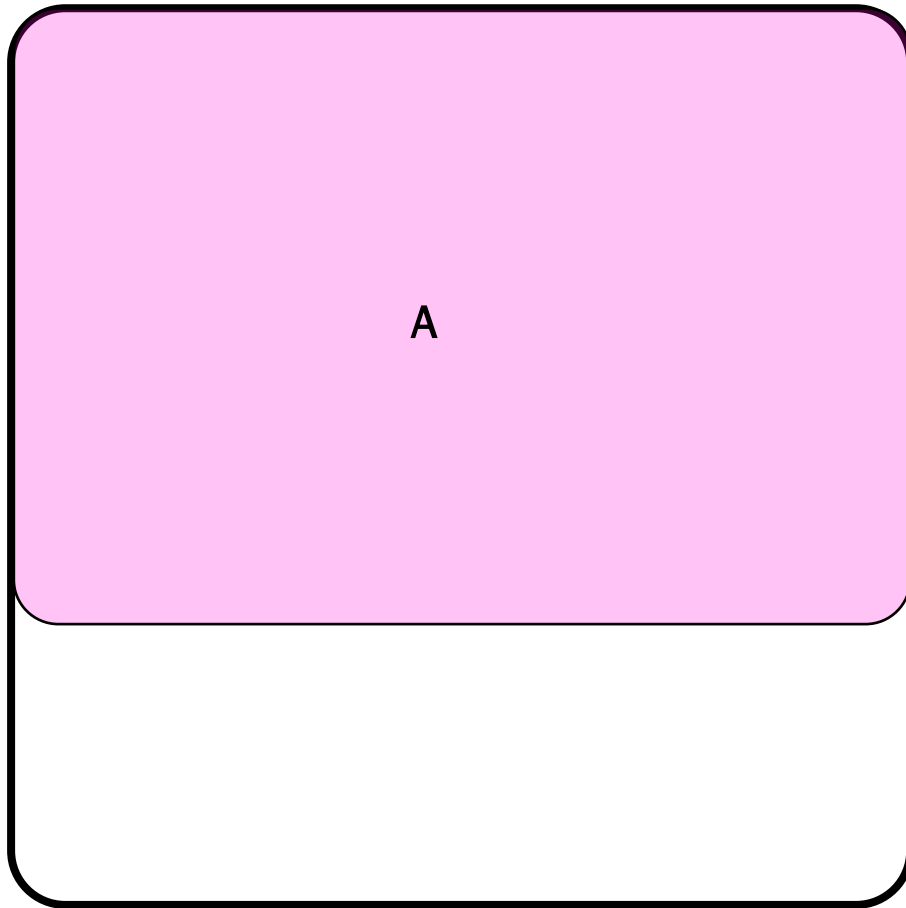
$$P(A|B) = P(A)$$

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

Independence

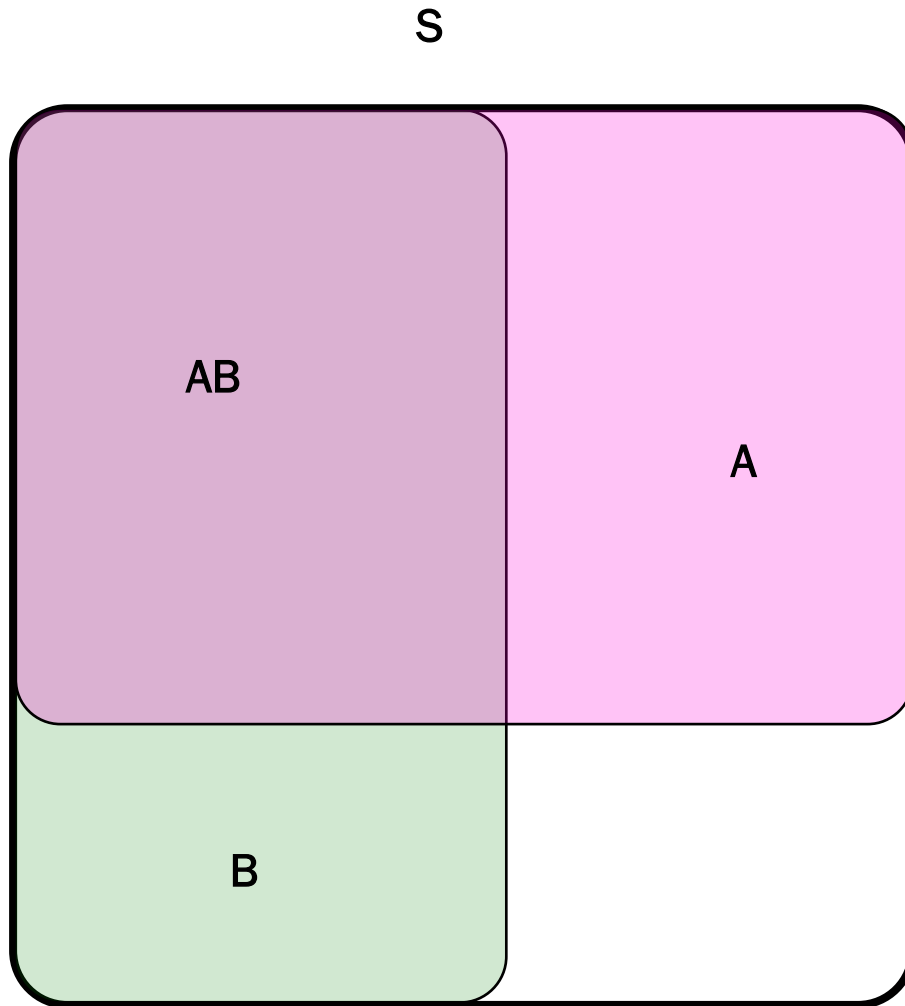
This ratio, $P(A)$...

... is the same as this one, $P(A|B)$



S

Independence



Independence Definition 1:

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

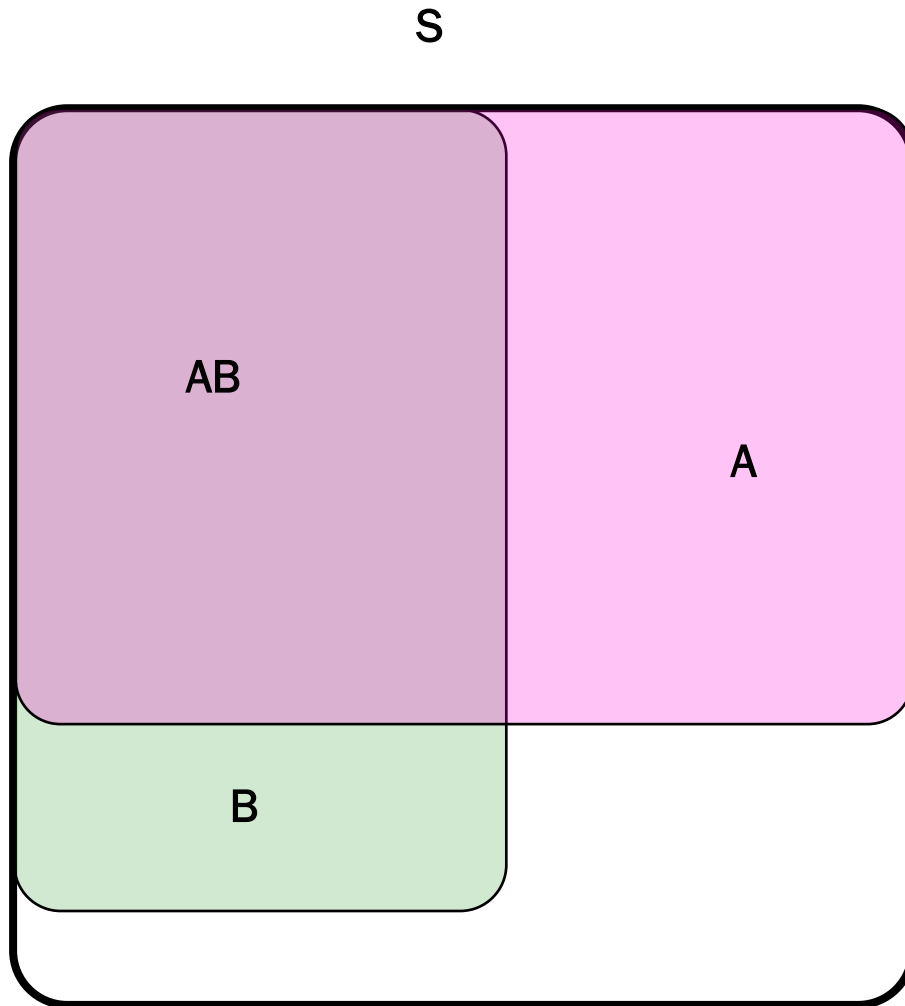
$$\frac{|AB|}{|S|} = \frac{|A|}{|S|} \times \frac{|B|}{|S|}$$

Independence Definition 2:

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

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

Dependence



Independence Definition 1:

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

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

Independence Definition 2:

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

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

More Intuition through proofs:

Independence

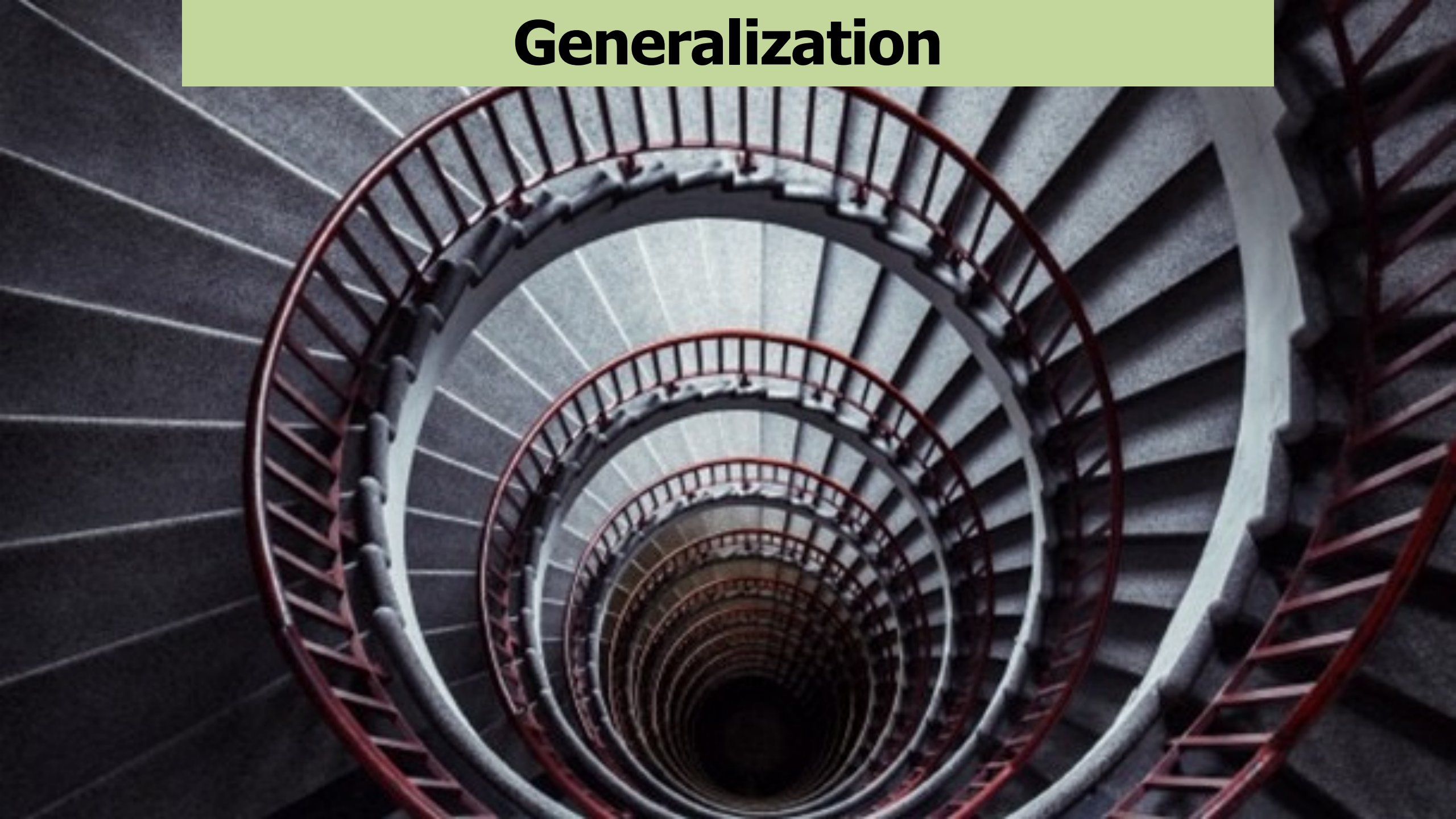
Given independent events A and B , prove that A and B^C are independent

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

Generalization



Generalized Independence

General definition of Independence:

Events E_1, E_2, \dots, E_n are independent if **for every subset** with r elements (where $r \leq n$) it holds that:

$$P(E_1, E_2, E_3, \dots, E_r) = P(E_1)P(E_2)P(E_3) \dots P(E_r)$$

Example: 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!**

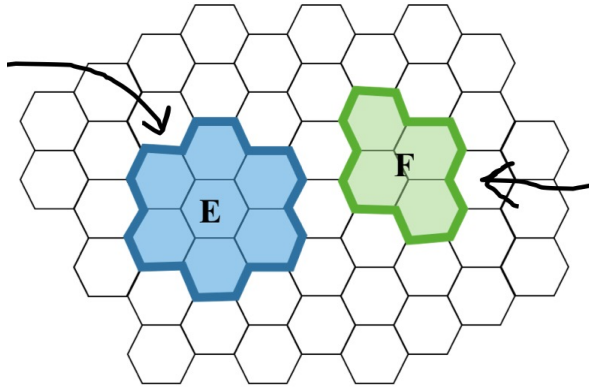
Let G be event: $D_1 + D_2 = 7$

- Are E and G independent? **Yes!**
- $P(E) = 1/6$, $P(G) = 1/6$, $P(E \cap G) = 1/36$ [roll (1, 6)]
- Are F and G independent? **Yes!**
- $P(F) = 1/6$, $P(G) = 1/6$, $P(F \cap G) = 1/36$ [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:

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



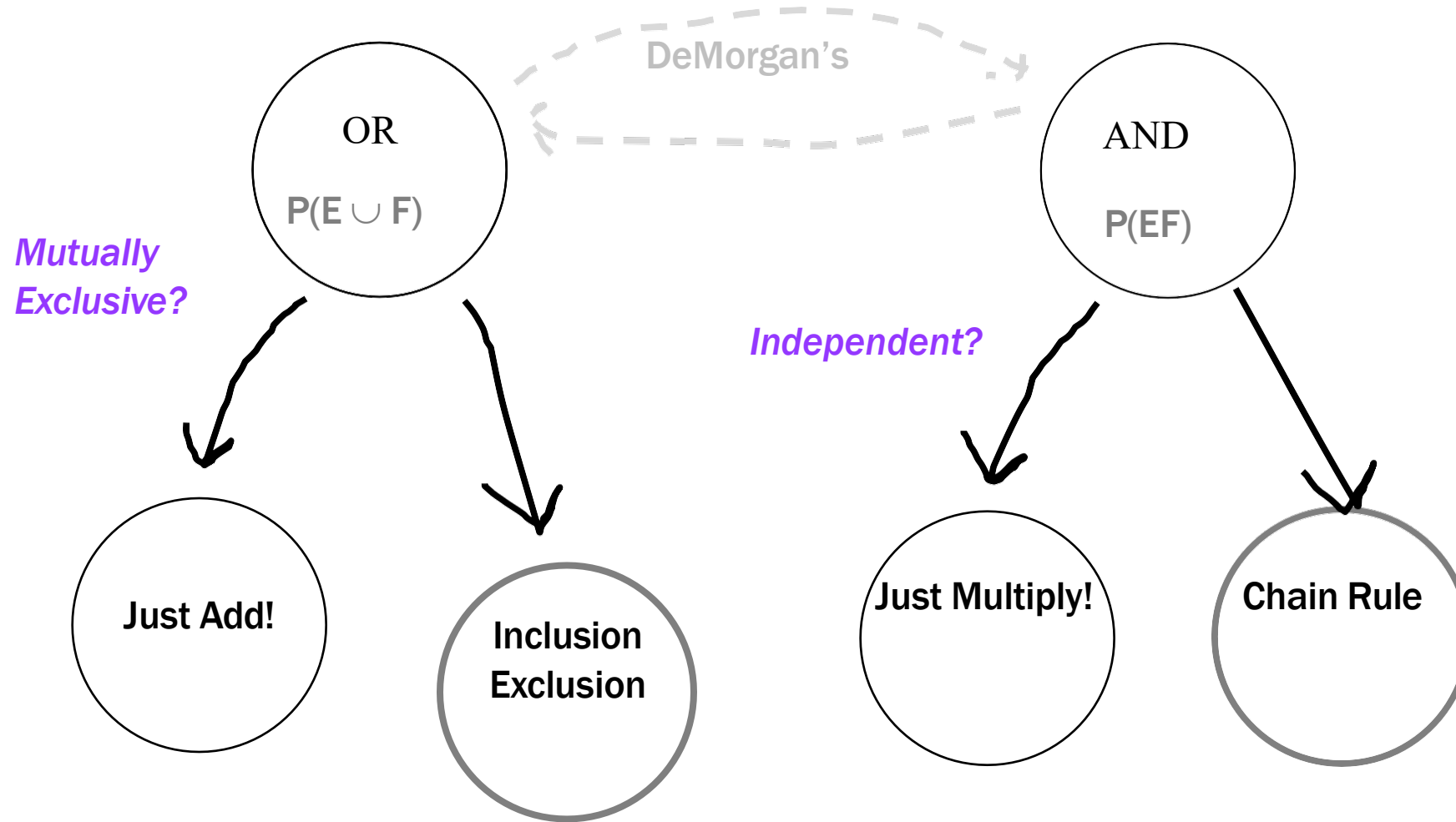
Independent

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

also:

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

Today



Think of the children as independent trials

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.

$$\begin{aligned} P(E_1 E_2 E_3) &= P(E_1) P(E_2 | E_1) P(E_3 | E_1 E_2) \\ &= P(E_1) P(E_2) P(E_3) \end{aligned}$$

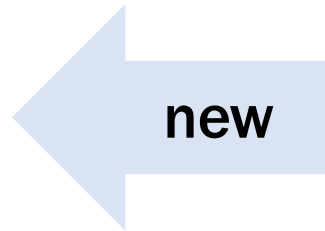
Independence

Two events E and F are defined as independent if:

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

For independent events E and F ,

- $P(E|F) = P(E)$
- E and F^C are independent.



Independence of complements

Statement:

If E and F are independent, then E and F^C are independent.

Proof:

$$\begin{aligned}P(EF^C) &= P(E) - P(EF) \\ &= P(E) - P(E)P(F) \\ &= P(E)[1 - P(F)] \\ &= P(E)P(F^C)\end{aligned}$$

E and F^C are independent

Intersection

Independence of E and F

Factoring

Complement

Definition of independence

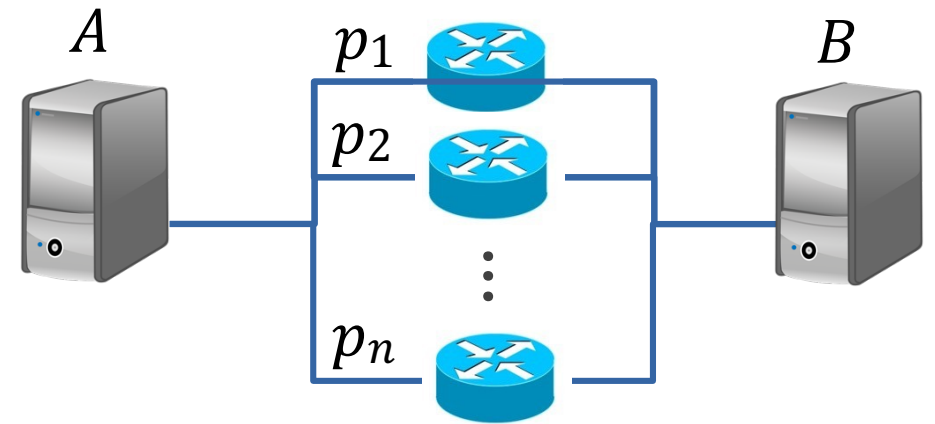
Knowing whether F does or doesn't happen doesn't change our belief about E happening.

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.

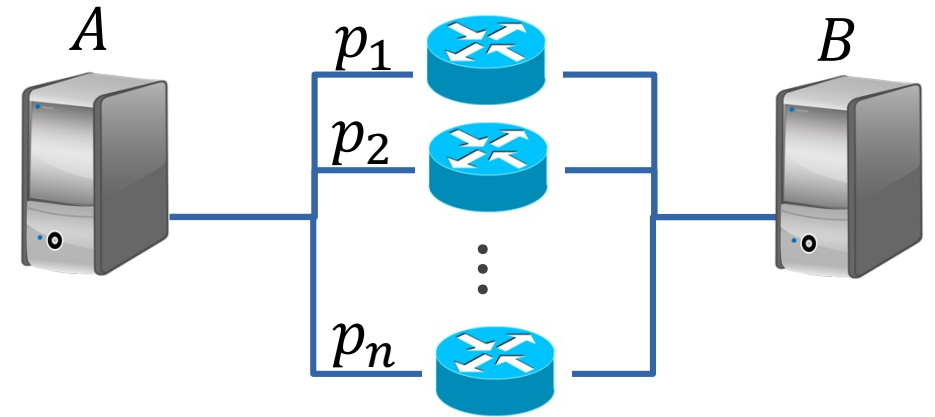
What is $P(E)$?



Network reliability

Consider the following parallel network:

- 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 $P(E)$?

$$\begin{aligned} P(E) &= P(\geq 1 \text{ one router works}) \\ &= 1 - P(\text{all routers fail}) \\ &= 1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n) \\ &= 1 - \prod_{i=1}^n (1 - p_i) \end{aligned}$$

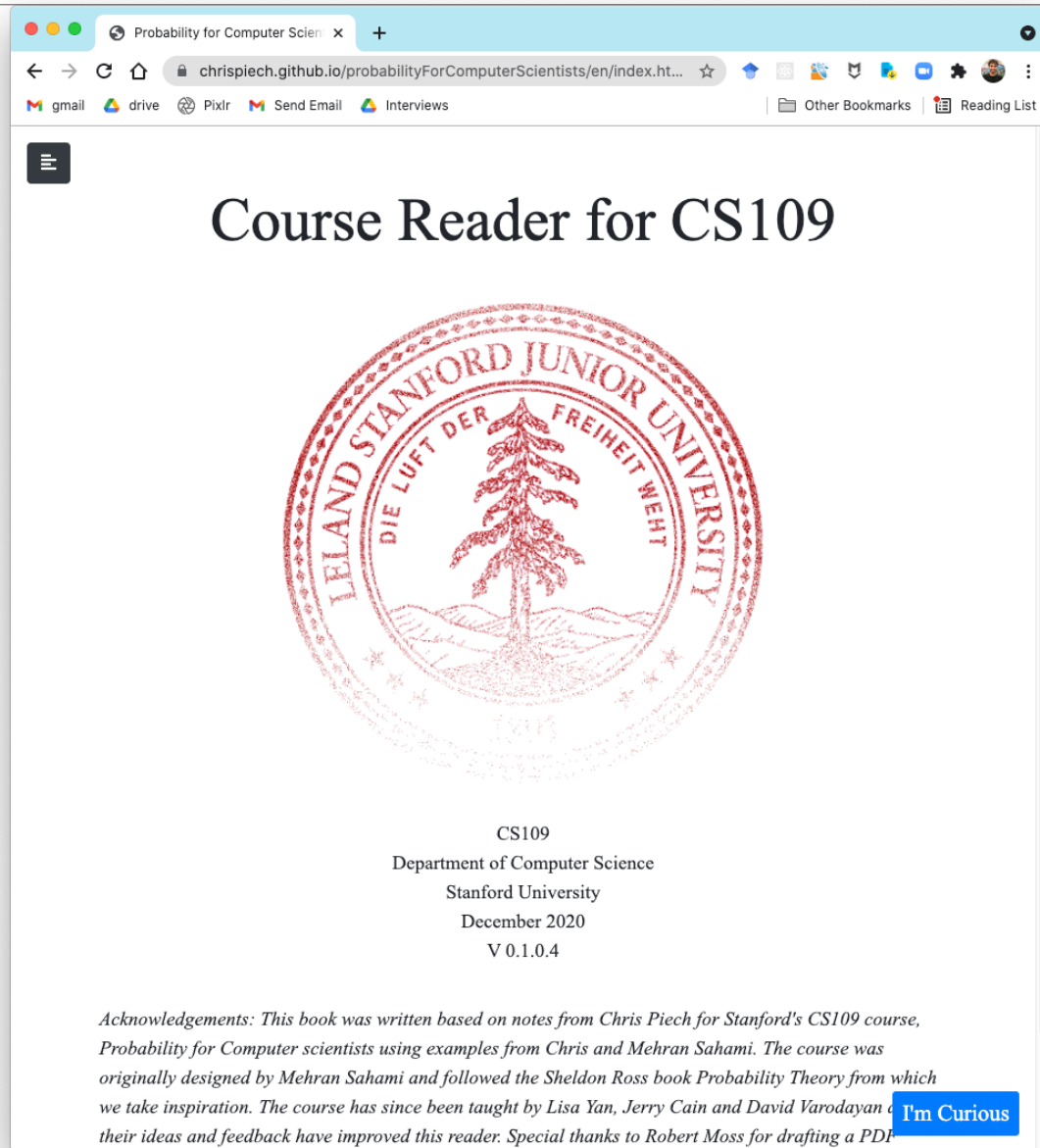
≥ 1 with independent trials:
take complement

The Most Important Core Probability Question

Say a coin comes up heads with probability p

- Flip the coin n times
- Each coin flip is an **independent** trial
- What is the probability of exactly k heads?


The Most Important Core Probability Question



Probability for Computer Science

chrispiech.github.io/probabilityForComputerScientists/en/index.ht...

Course Reader for CS109



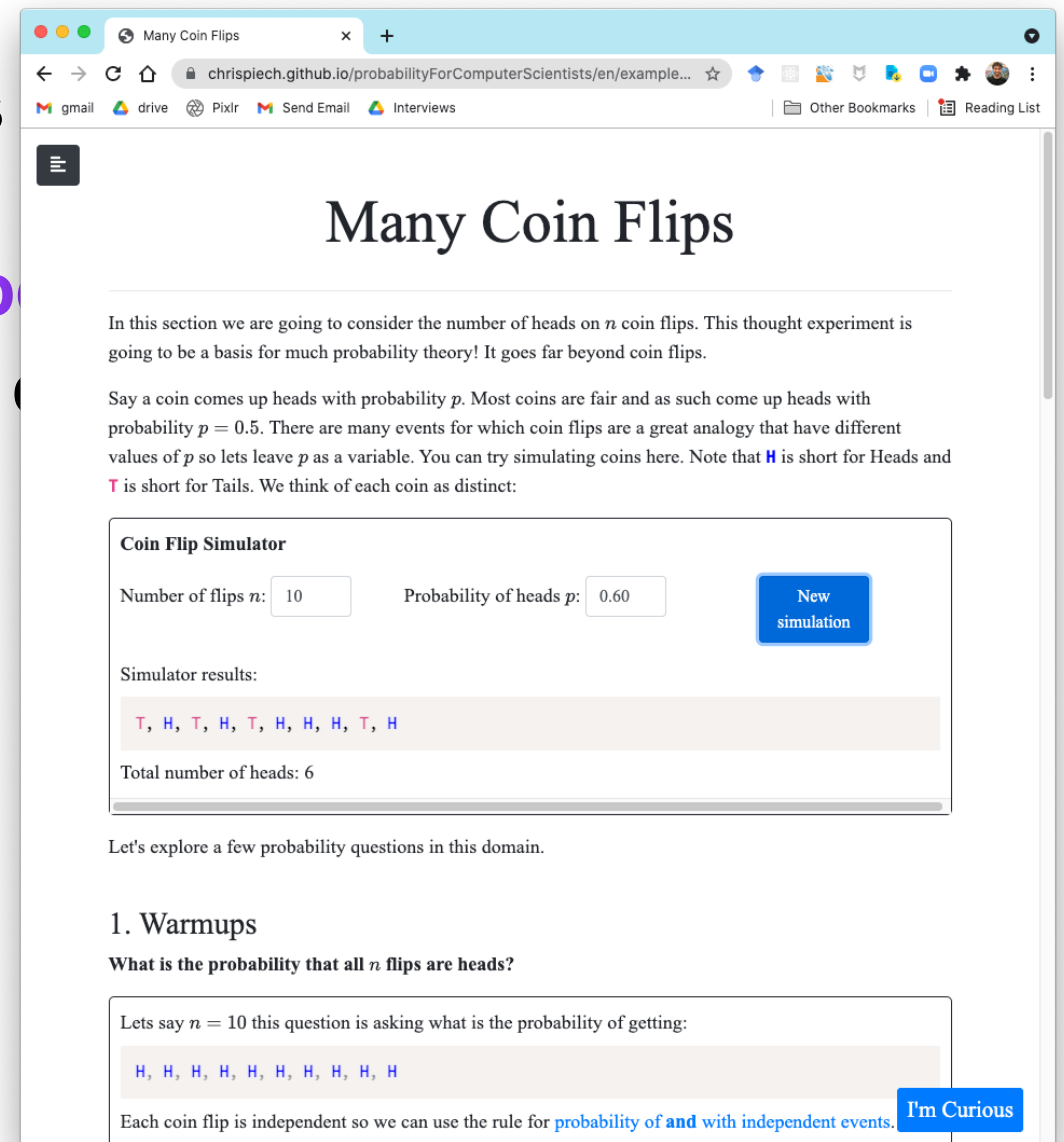
CS109
Department of Computer Science
Stanford University
December 2020
V 0.1.0.4

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 and their ideas and feedback have improved this reader. Special thanks to Robert Moss for drafting a PDF

I'm Curious

ads

dep
y of c



Many Coin Flips

chrispiech.github.io/probabilityForComputerScientists/en/example...

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 : Probability of heads p :

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?

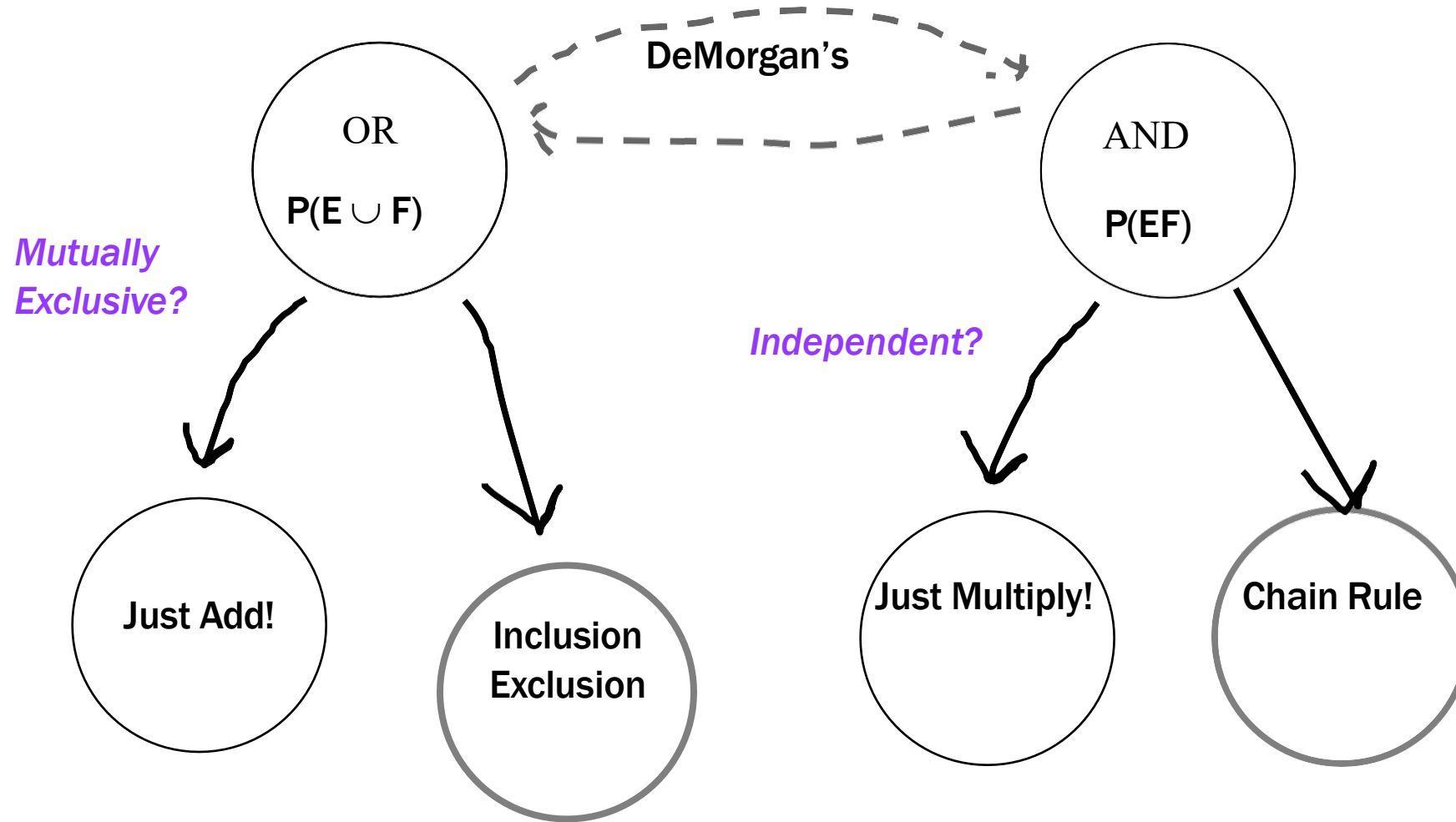
Lets say $n = 10$ this question is asking what is the probability of getting:

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.

I'm Curious

Pedagogical Pause



Sets Review

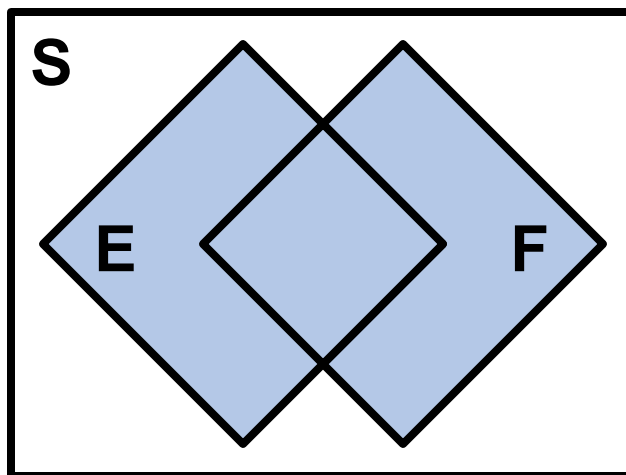


Sets Review

Say E and F are events in S

Event that is in E or F

$$E \cup F$$



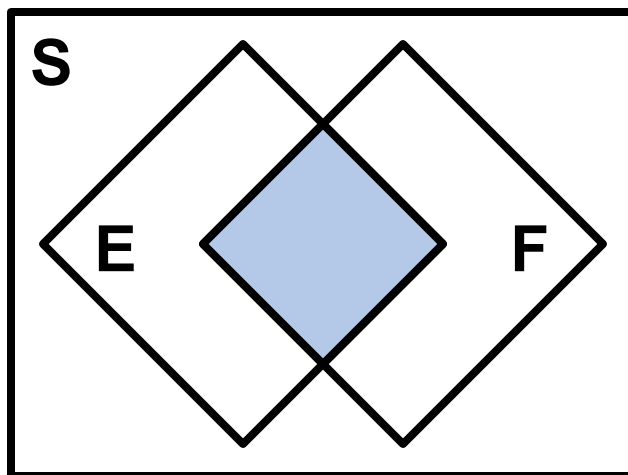
- $S = \{1, 2, 3, 4, 5, 6\}$ die roll outcome
- $E = \{1, 2\}$ $F = \{2, 3\}$ $E \cup F = \{1, 2, 3\}$

Sets Review

Say E and F are events in S

Event that is in E and F

$$E \cap F$$

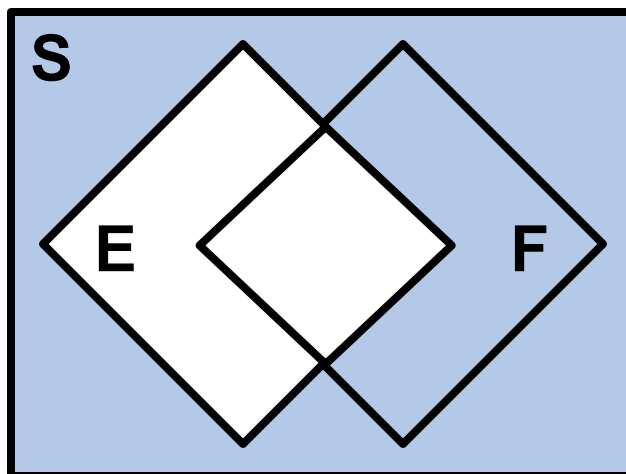


Sets Review

Say E and F are events in S

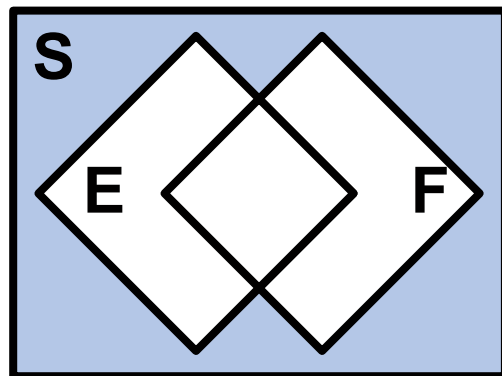
Event that is not in E (called complement of E)

E^c or $\sim E$



Sets Review

Say E and F are subsets of S



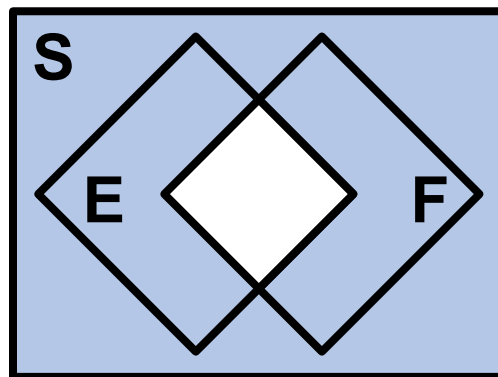
Which of these two is it?

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

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

Sets Review

Say E and F are subsets of S



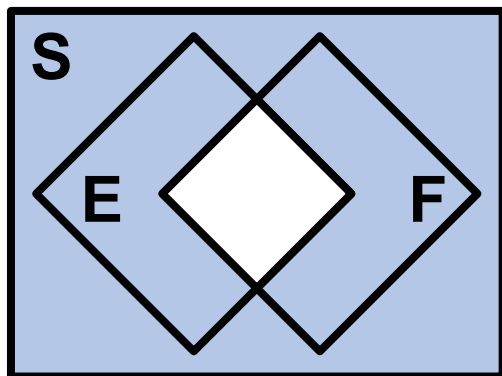
Which of these two is it?

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

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

De Morgan's Laws

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



$$(E \cap F)^C = E^C \cup F^C$$

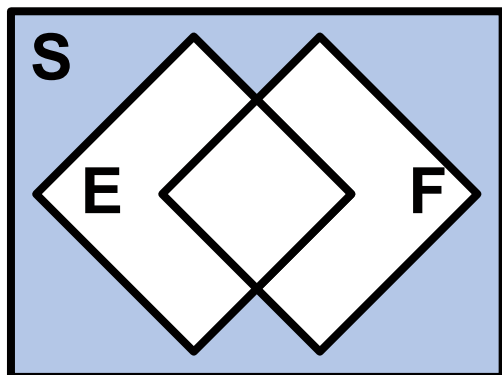
In probability:

$$P(E_1 E_2 \cdots E_n)$$

$$= 1 - P\left((E_1 E_2 \cdots E_n)^C\right)$$

$$= 1 - P\left(E_1^C \cup E_2^C \cup \cdots \cup E_n^C\right)$$

Great if E_i^C mutually exclusive!



$$(E \cup F)^C = E^C \cap F^C$$

In probability:

$$P(E_1 \cup E_2 \cup \cdots \cup E_n)$$

$$= 1 - P\left((E_1 \cup E_2 \cup \cdots \cup E_n)^C\right)$$

$$= 1 - P(E_1^C E_2^C \cdots E_n^C)$$

Great if E_i independent!

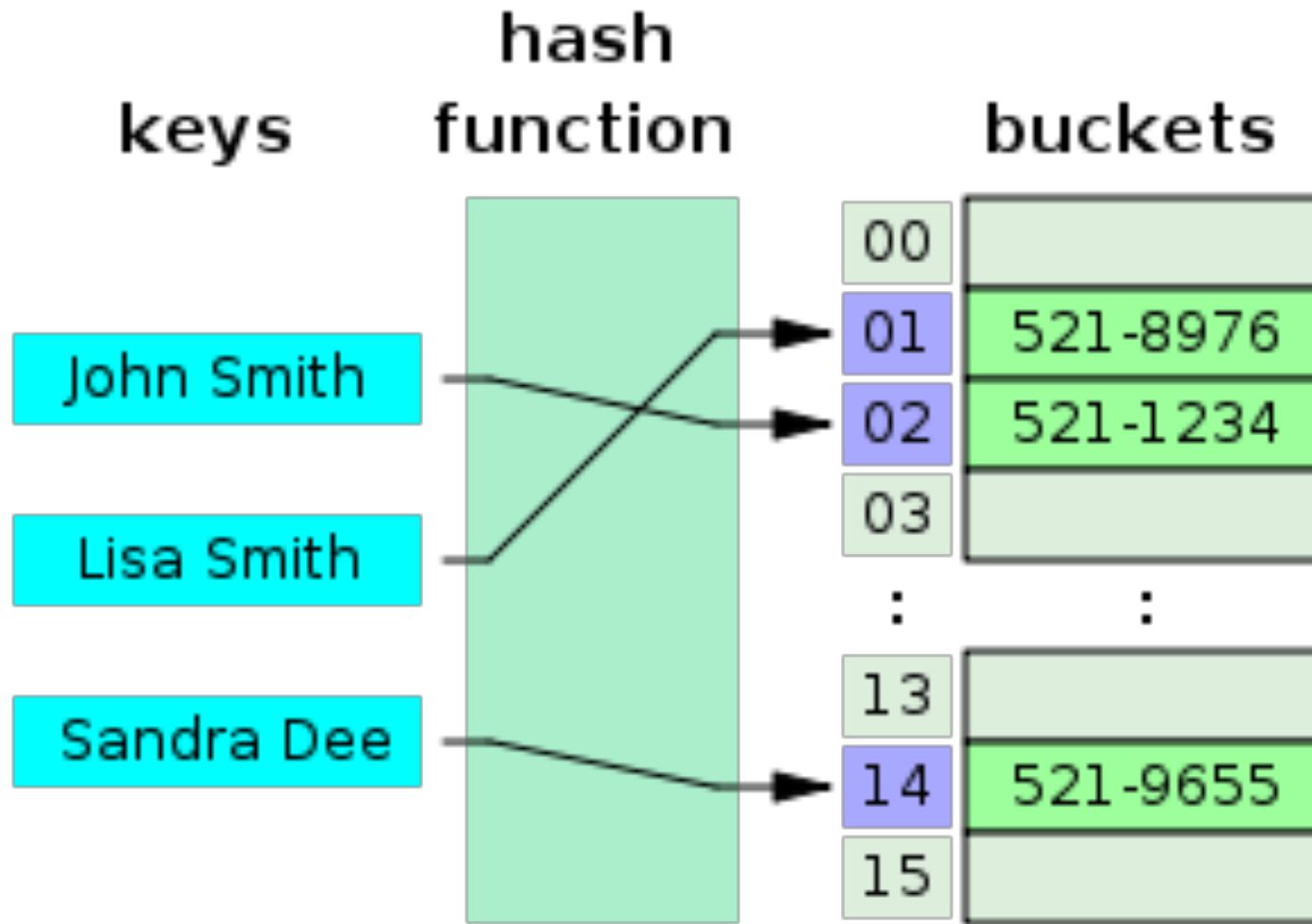
Augustin Demorgan



Jason Alexander

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

Hash Tables. Hardest Core Probability Question




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

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

Define: $S_i =$ 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

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

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

$$P(E) = P(S_1 \cup S_2 \cup \dots \cup S_m)$$

$$= 1 - P\left((S_1 \cup S_2 \cup \dots \cup S_m)^c\right)$$

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

$$= 1 - P(S_1^c)P(S_2^c) \dots P(S_m^c) = 1 - \left(P(S_1^c)\right)^m$$

$$= 1 - (1 - p_1)^m$$

Define: $S_i =$ string i hashes to bucket 1
 $S_i^c =$ string i doesn't hash to bucket 1

Complement

De Morgan's Law

S_i independent trials

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

More hash table fun: Possible approach?

- 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ **at least 1** of buckets 1 to k has ≥ 1 string hashed into it?

$P(E) =$

More hash table fun: Possible approach?

- 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ **at least 1** of buckets 1 to k has ≥ 1 string hashed into it?

$P(E) =$

Define $F_i =$ bucket i has at least one string in it

More hash table fun: Possible approach?

- 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ **at least 1** of buckets 1 to k has ≥ 1 string hashed into it?

$$P(E) = P(F_1 \cup F_2 \cup \dots \cup F_k)$$

Define $F_i =$ bucket i has at least one string in it

 F_i bucket events are *dependent!* So we cannot just add.

More hash table fun: Possible approach?

- 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ **at least 1** of buckets 1 to k has ≥ 1 string hashed into it?

$$\begin{aligned} P(E) &= P(F_1 \cup F_2 \cup \dots \cup F_k) \\ &= 1 - P\left((F_1 \cup F_2 \cup \dots \cup F_k)^C\right) \\ &= 1 - P\left(F_1^C F_2^C \dots F_k^C\right) \\ &= \end{aligned}$$

Define $F_i =$ bucket i has at least one string in it

 F_i bucket events are *dependent!* So we cannot just add.

More hash table fun: Possible approach?

- 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ **at least 1** of buckets 1 to k has ≥ 1 string hashed into it?

$$\begin{aligned} P(E) &= P(F_1 \cup F_2 \cup \dots \cup F_k) \\ &= 1 - P\left((F_1 \cup F_2 \cup \dots \cup F_k)^C\right) \\ &= 1 - P\left(F_1^C F_2^C \dots F_k^C\right) \\ &= \end{aligned}$$

Define $F_i =$ bucket i has at least one string in it

$$\begin{aligned} &= P(\text{buckets 1 to } k \text{ all denied strings}) \\ &= (P(\text{each string hashes to } k + 1 \text{ or higher})) \\ &= (1 - p_1 - p_2 \dots - p_k)^m \end{aligned}$$

 F_i bucket events are *dependent*! So we cannot just add.

More hash table fun: Possible approach?

- 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ **at least 1** of buckets 1 to k has ≥ 1 string hashed into it?

$$\begin{aligned}P(E) &= P(F_1 \cup F_2 \cup \dots \cup F_k) \\&= 1 - P\left((F_1 \cup F_2 \cup \dots \cup F_k)^C\right) \\&= 1 - P(F_1^C F_2^C \dots F_k^C) \\&= 1 - (1 - p_1 - p_2 \dots - p_k)^m\end{aligned}$$

Define $F_i =$ bucket i has at least one string in it

$$\begin{aligned}&= P(\text{buckets 1 to } k \text{ all denied strings}) \\&= (P(\text{each string hashes to } k + 1 \text{ or higher})) \\&= (1 - p_1 - p_2 \dots - p_k)^m\end{aligned}$$

 F_i bucket events are *dependent*! So we cannot just add.

The fun never stops with hash tables

- 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it? 
2. $E =$ at least 1 of buckets 1 to k has ≥ 1 string hashed into it? 

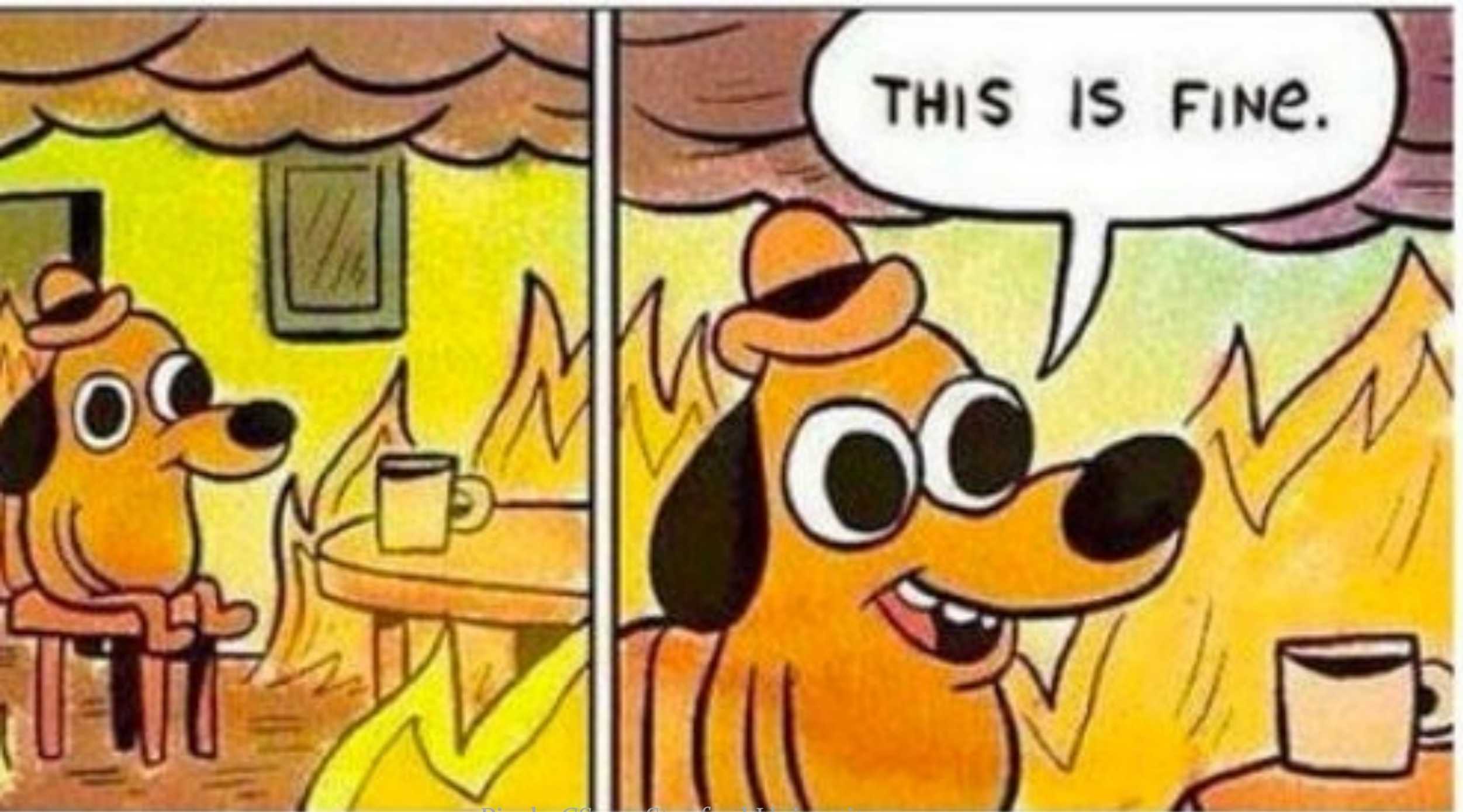
Looking for a challenge? 😊

The fun never stops with hash tables

- m strings are hashed (unequally) 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ at least 1 of buckets 1 to k has ≥ 1 string hashed into it?
3. $E =$ **each** of buckets 1 to k has ≥ 1 string hashed into it?

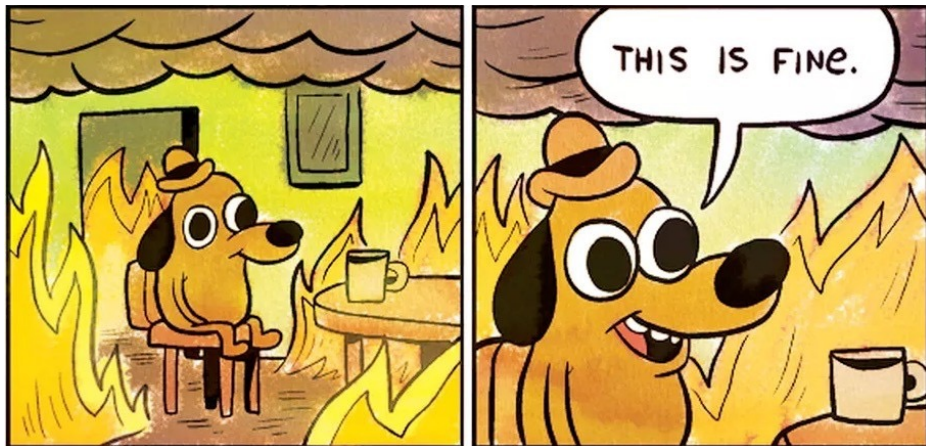


The fun never stops with hash tables

- m strings are hashed (unequally) 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

1. $E =$ bucket 1 has ≥ 1 string hashed into it?
2. $E =$ at least 1 of buckets 1 to k has ≥ 1 string hashed into it?
3. $E =$ **each** of buckets 1 to k has ≥ 1 string hashed into it?



Hint: Use Part 2's event definition:

Define $F_i =$ bucket i has at least one string in it

Hint: Try $k = 2$, then $k = 3$, then generalize.

The fun never stops with hash tables

Solution

- F_i = at least one string hashed into i -th bucket
- $P(E) = P(F_1 F_2 \dots F_k) = 1 - P((F_1 F_2 \dots F_k)^c)$
 $= 1 - P(F_1^c \cup F_2^c \cup \dots \cup F_k^c)$ (DeMorgan's Law)
 $= 1 -$

where
$$P\left(\bigcup_{i=1}^k F_i^c\right) = 1 - \sum_{r=1}^k (-1)^{(r+1)} \sum_{i_1 < \dots < i_r} P(F_{i_1}^c F_{i_2}^c \dots F_{i_r}^c)$$

$$P(F_{i_1}^c F_{i_2}^c \dots F_{i_r}^c) = (1 - p_{i_1} - p_{i_2} - \dots - p_{i_r})^m$$

Here we are



Source: The Hobbit

G_1

G_2

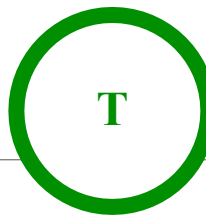
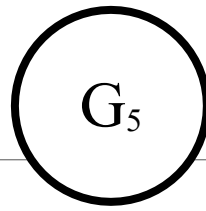
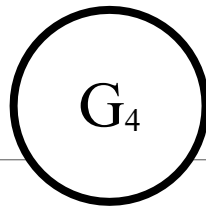
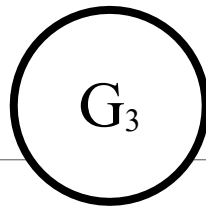
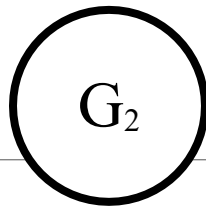
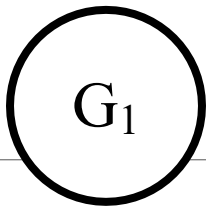
G_3

G_4

G_5

T





```

dna.txt — dna
dna.txt
1 False, True, False, False, True, False
2 True, True, False, True, True, False
3 True, True, False, True, True, True
4 False, True, False, True, True, False
5 False, True, False, False, True, False
6 True, True, False, True, True, True
7 False, False, True, False, False, False
8 False, False, True, False, True, False
9 True, False, False, True, False, False
10 False, True, False, True, True, False
11 True, False, False, True, False, False
12 True, False, True, True, False, False
13 False, True, False, False, True, False
14 False, False, True, True, False, False
15 True, True, False, False, True, True
16 True, False, True, True, False, False
17 True, True, True, True, True, True |
18 True, False, True, False, False, True
19 False, True, False, True, True, True
20 False, False, True, False, False, False
21 False, False, False, True, True, False
22 False, True, False, False, True, False
23 True, True, False, True, True, True
24 False, True, False, True, True, False
25 True, False, False, False, False, True
26 False, False, True, True, False, True
27 False, False, False, True, False, False
28 False, True, True, False, False, True
29 False, True, False, False, True, True
30 False, False, False, False, False, True
31 False, True, False, True, True, False
32 True, False, False, True, False, False
33 True, True, False, True, True, True
34 True, True, False, False, True, True
35 True, True, False, True, True, True
36 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) = 0.200 , P(T)p(G2) = 0.212
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) = 0.200 , P(T)p(G2) = 0.212
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
```

Only Causal Structure That Fits

