CS109: Probability for Computer Scientists
Instructors
Chris Piech

Created a research lab in: 
AI for Social Good (esp Education)

Grew up in Nairobi, Kuala Lumpur before Stanford!

Piech + Cain, CS109, Stanford University
Long History in CS109

I took the first CS109 back when I looked like this

Been teaching it since 2014
Probability in my Research: Better Eye Exam

Jan 2020, With a former CS109 student, Ali Malik

Math question: Estimate a continuous valued number. Get to run noisy experiments of your choosing.
Math question: Can you remove racism from a deep learning predictor?
So many things to love in this world
I went here from 1987 through 1991 and majored in chemistry.

Then I came here for a PhD in chem, switched to CS

Received MSCS 1998
Lecturer: not so new
Speculation: Taught more classes than anyone else here.
Why Jerry likes probability

- I majored in chemistry, and my undergraduate research was rooted in surface science and **statistical** mechanics.

- When I switched to CS as a grad student here, I focused on CS theory and all the beautiful mathematics that comes with it. Mathematics feeds the soul.

- Probability has revived parts of AI and information theory that were thought to be borderline dead when I was getting my MSCS degree here.

- \[ PV = \frac{1}{3} N m v_{\text{rms}}^2 \]
  \[ f(v) = 4\pi \left( \frac{m}{2\pi kT} \right)^{\frac{3}{2}} v^2 e^{-\frac{mv^2}{2kT}} \]
  \[ v_{\text{rms}}^2 = \int_0^\infty v^2 f(v) \, dv = 4\pi \left( \frac{m}{2\pi kT} \right)^{\frac{3}{2}} \int_0^\infty v^4 e^{-\frac{mv^2}{2kT}} \, dv \]
Amazing Teaching Team
Course mechanics

(this is a light version. Please read the handout for details).
What makes this quarter important

We are seeing a huge surge in **statistics, predictions, and probabilistic models** shared through global news, governing bodies, and social media.

Global cases of COVID-19 as of September 14th (JHU)
https://coronavirus.jhu.edu/map.html

National Weather Service Alerts
https://www.weather.gov/

US Presidential Election 2020 prediction forecasts
https://fivethirtyeight.com/
https://www.nytimes.com/
What makes this quarter important

We are seeing a huge surge in statistics, predictions, and probabilistic models shared through global news, governing bodies, and social media.

The challenge of delivering Stanford-class education online reflects our university’s commitment to fostering a diverse body of students.
Essential Information

cs109.stanford.edu
• **Camera on** during breakouts.

• **Camera off** during lecture is just fine!

• Come to lecture, it’s a great time.

• **Ask questions** in the chat.

• Use Ed for questions after lecture.
Are you in the right place?
Prerequisites

What you really need:

**CS106B/X (important):**
- Recursion
- Hash Tables
- Binary Trees
- Programming

**CS103 (ok as a corequisite):**
- Proof techniques (induction)
- Set theory
- Math maturity

**Math 51 or CME 100 (important):**
- Multivariate differentiation
- Multivariate integration
- Basic facility with linear algebra (vectors)
Coding in CS109

Review session on Friday

Piech + Cain, CS109, Stanford University
CS109 Units

Hours per week = Units $\times 3$

Average about 10 hours / week for assignments

Start Here

Are you an Undergrad?

No

Do you want to take CS109 for fewer units?

Yes

3 Units -or- 4 Units

Yes

No

5 Units
## Class Breakdown

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>42%</td>
<td>6 Assignments</td>
</tr>
<tr>
<td>42%</td>
<td>3 Quizzes</td>
</tr>
<tr>
<td></td>
<td>48 hour take home tests</td>
</tr>
<tr>
<td>10%</td>
<td>Concept Checks</td>
</tr>
<tr>
<td></td>
<td>Daily after lecture</td>
</tr>
<tr>
<td>6%</td>
<td>Section Participation</td>
</tr>
</tbody>
</table>
Ask questions

Q&A forum
All announcements

“Working” office hours
start after Wednesday class

Email cs109@cs.stanford.edu
Companion class: CS109A

- CS109A is an extra 1-unit “Pathfinders” or “ACE” section with additional support, practice, and instruction
- Meets for an additional weekly section and has additional review session
- Entry by application — see course website for details
Brand new this quarter: Course Reader!
Story of Modern AI
Modern AI
or, How we learned to combine probability and programming
Brief History
Early Optimism 1950s

1952

1955

Axioms \( \models C \)

ATP System (theorem prover)

Yes (proof/answer)

No

Timeout
Early Optimism 1950s

“Machines will be capable, within twenty years, of doing any work a man can do.”
–Herbert Simon, 1952
Underwhelming Results 1950s to 1980s

The spirit is willing but the flesh is weak.

(Russian)

The vodka is good but the meat is rotten.

The world is too complex
BRACE YOURSELVES

WINTER IS COMING
Something is going on in the world of AI
Big Milestones Part 1

1997 Deep Blue

2005 Stanley

2011 Watson
I was told speech was 30 years out

Almost perfect...
The last remaining board game
Computers Making Art
Self Driving Cars
What is going on?
Focus on one problem
Computer Vision

Chihuahua or muffin?

Piech + Cain, CS109, Stanford University
Can you do it?
Chihuahua or Muffin?
Chihuahua or Muffin?
How about now?

<p>| | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What a computer sees

What a human sees

Piech + Cain, CS109, Stanford University
public class Chihuahua extends ConsoleProgram {

    public void run() {
        println("Todo: Write program");
    }

}
Two Great Ideas

1. Probability from Examples

2. Artificial Neurons
Two Great Ideas

1. Probability from Examples

2. Artificial Neurons
Probability from Examples
When does the magic happen?

Lots of Data + Sound Probability
Basically just a rebranding of statistics and probability.
Computer Vision is Still Hard

You see this:

But the camera sees this:

<table>
<thead>
<tr>
<th>194</th>
<th>210</th>
<th>201</th>
<th>212</th>
<th>199</th>
<th>213</th>
<th>215</th>
<th>195</th>
<th>178</th>
<th>158</th>
<th>182</th>
<th>209</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>189</td>
<td>190</td>
<td>221</td>
<td>209</td>
<td>205</td>
<td>191</td>
<td>167</td>
<td>147</td>
<td>115</td>
<td>129</td>
<td>163</td>
</tr>
<tr>
<td>114</td>
<td>126</td>
<td>140</td>
<td>188</td>
<td>176</td>
<td>165</td>
<td>152</td>
<td>140</td>
<td>170</td>
<td>106</td>
<td>78</td>
<td>88</td>
</tr>
<tr>
<td>87</td>
<td>103</td>
<td>115</td>
<td>154</td>
<td>143</td>
<td>142</td>
<td>149</td>
<td>153</td>
<td>173</td>
<td>101</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>102</td>
<td>112</td>
<td>106</td>
<td>131</td>
<td>122</td>
<td>130</td>
<td>152</td>
<td>147</td>
<td>128</td>
<td>84</td>
<td>58</td>
<td>66</td>
</tr>
<tr>
<td>94</td>
<td>95</td>
<td>79</td>
<td>104</td>
<td>105</td>
<td>124</td>
<td>129</td>
<td>113</td>
<td>107</td>
<td>87</td>
<td>69</td>
<td>67</td>
</tr>
<tr>
<td>60</td>
<td>71</td>
<td>69</td>
<td>98</td>
<td>89</td>
<td>92</td>
<td>98</td>
<td>95</td>
<td>89</td>
<td>80</td>
<td>76</td>
<td>67</td>
</tr>
<tr>
<td>41</td>
<td>56</td>
<td>68</td>
<td>99</td>
<td>63</td>
<td>45</td>
<td>60</td>
<td>82</td>
<td>58</td>
<td>76</td>
<td>75</td>
<td>65</td>
</tr>
<tr>
<td>20</td>
<td>43</td>
<td>69</td>
<td>75</td>
<td>56</td>
<td>41</td>
<td>51</td>
<td>73</td>
<td>55</td>
<td>70</td>
<td>63</td>
<td>44</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>57</td>
<td>69</td>
<td>75</td>
<td>75</td>
<td>74</td>
<td>53</td>
<td>53</td>
<td>60</td>
<td>59</td>
<td>37</td>
</tr>
<tr>
<td>72</td>
<td>59</td>
<td>53</td>
<td>66</td>
<td>84</td>
<td>92</td>
<td>84</td>
<td>74</td>
<td>57</td>
<td>72</td>
<td>63</td>
<td>42</td>
</tr>
<tr>
<td>67</td>
<td>61</td>
<td>58</td>
<td>65</td>
<td>75</td>
<td>75</td>
<td>76</td>
<td>73</td>
<td>59</td>
<td>75</td>
<td>69</td>
<td>50</td>
</tr>
</tbody>
</table>

[Andrew Ng]
Computer Vision is Still Hard

Find edges at four orientations

Sum up edge strength in each quadrant

Final feature vector

[Andrew Ng]
Straight ML not perfect
Some Great Thinkers

Daphne Koller
Two Great Ideas

1. Probability from Examples

2. Artificial Neurons
Artificial Neurons
Human Neuron
Human Neuron
Human Neuron
Human Neuron
Human Neuron
Artificial Neuron
Artificial Neural Network

Each node represents a neuron

Each edge represents the weight of the interaction

Pixel
Artificial Neural Network

Each node represents a neuron

Each edge represents the weight of the interaction

Piech + Cain, CS109, Stanford University
Artificial Neural Network

Each node represents a neuron

Each edge represents the weight of the interaction
Artificial Neural Network

Each node represents a neuron

Each edge represents the weight of the interaction
Artificial Neural Network
Interpret the last neuron as the “probability” that the image is of a chihuahua.
Artificial Neural Network

The image had a 0 but we predicted a high probability that it was a 1
Let’s update our weights to make our probabilities better match reality.

The image had a muffin but we predicted a high probability that it was a chihuahua.

Artificial Neural Network

Piech + Cain, CS109, Stanford University
Artificial Neural Network

Let’s update our weights to make our probabilities better match reality.

The image had a muffin but we predicted a high probability that it was a chihuahua.

Update the parameter at each connection.
Update Neural Network

\[ P(Y = 1|X = x) = \hat{y} \]

\[ \hat{y} = \sigma \left( \sum_{j=0}^{m_h} h_j \theta_j^{(y)} \right) \]

For one datum

\[ P(Y = y|X = X) = (\hat{y})^y (1 - \hat{y})^{1-y} \]

For IID data

\[ L(\theta) = \prod_{i=1}^{n} P(Y = y^{(i)}|X = x^{(i)}) \]

\[ = \prod_{i=1}^{n} (\hat{y}^{(i)})^{y^{(i)}} \cdot \left[ 1 - (\hat{y}^{(i)}) \right]^{(1-y^{(i)})} \]
Gradient Descent

Walk uphill and you will find a local maxima
(if your step size is small enough)

Piech + Cain, CS109, Stanford University
Gradient of Probability

\[
\frac{\partial L}{\partial \theta_i^{(\hat{y})}} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial \theta_i^{(\hat{y})}}
\]

\[
\hat{y} = \sigma \left( \sum_{j=0}^{m_h} h_j \theta_j^{(\hat{y})} \right)
\]

\[
\frac{\partial \hat{y}}{\partial \theta_i^{(\hat{y})}} = \sigma \left( \sum_{j=0}^{m_h} h_j \theta_j^{(\hat{y})} \right) \left[ 1 - \sigma \left( \sum_{j=0}^{m_h} h_j \theta_j^{(\hat{y})} \right) \right] \cdot \frac{\partial}{\partial \theta_i^{(\hat{y})}} \sum_{j=0}^{m_h} h_j \theta_j^{(\hat{y})}
\]

\[
= \hat{y}[1-\hat{y}] \cdot \frac{\partial}{\partial \theta_i^{(\hat{y})}} \sum_{j=0}^{m_h} h_j \theta_j^{(\hat{y})}
\]

\[
= \hat{y}[1-\hat{y}] \cdot h_i
\]

You will be able to do this.
Where you will be by the end of class
CS109: Theory Class focused on Applications

Heart

Ancestry

Netflix
A machine learning algorithm performs better than the best dermatologists.

Developed recently, at Stanford.

AI Augmented Education?

50 thousand students
Given $n$ historical answers:

$$x_1 \quad x_2 \quad \ldots \quad x_n$$

Answer is a tuple:

$$x_i = \{q_i, a_i\}$$

Predict the next one
Prediction Results

Huge improvement in ability to predict for real students

Piech + Cain, CS109, Stanford University

Piech et al., 2015
Learns Concept Relations

Scatter plots

Pythagorean theorem

Line graphs

Systems of Equations

Lines

Angles

Functions

Exponents

Fractions

Piech + Cain, CS109, Stanford University
Not once, but twice, AI was revolutionized by people who understood probability theory.
End of Story
Except it isn’t the end of the story...
Probability is more than just machine learning
Abundance of important problems
Algorithms and Probability

Eg Raytracing

Eg HashMaps

Hash Fn

- "cuckoo hashing"
- "john coltrane"
- "mantis shrimp"
- "stanford band"

8
42
77
314
The next medical revolution?
Autocomplete
Recommender Systems

Piech + Cain, CS109, Stanford University
Philosophy and Ethics
Art and Probability
Most Desired Skill in Industry

Job Score is based on:

• Earning potential
• Number of jobs
• Job satisfaction rating

“Data science and machine learning are generating more jobs than candidates right now, making these two areas the fastest growing employment areas.”

9.8 times more jobs than five years ago.

LinkedIn’s 2017 U.S. Emerging Jobs Report
Most Desired Skill in Academia

Most CS PhD students list their highest desiderata upon graduation as:

“Better understanding of probability”
Open Problem: One Shot Learning

Human-level concept learning through probabilistic program induction.

Current deep learning methods are not enough to move the needle as far as we want, especially on socially relevant problems that often do not have the benefit of massive public datasets. The best new ideas are coming from probability theory.
Open Problem: One Shot Learning

Human-level concept learning through probabilistic program induction.

Current deep learning methods are not enough to move the needle as far as we want, especially on socially relevant problems that often do not have the benefit of massive public datasets. The best new ideas are coming from probability theory.
Open Problem: One Shot Learning

Human-level concept learning through probabilistic program induction.

Current deep learning methods are not enough to move the needle as far as we want, especially on socially relevant problems that often do not have the benefit of massive public datasets. The best new ideas are coming from probability theory.
Open Problem: One Shot Learning

Human-level concept learning through probabilistic program induction.

Current deep learning methods are not enough to move the needle as far as we want, especially on socially relevant problems that often do not have the benefit of massive public datasets. The best new ideas are coming from probability theory.
Open Problem: One Shot Learning


Human-level concept learning through probabilistic program induction.

Current deep learning methods are not enough to move the needle as far as we want, especially on socially relevant problems that often do not have the benefit of massive public datasets. The best new ideas are coming from probability theory.
Foundation for your future
But its not always intuitive
A patient has a positive Zika test. What is the probability they have zika?

- 0.8% of people have zika
- Test has 90% positive rate for people with zika
- Test has 7% positive rate for people without zika

The right answer is 9%
Probability = Important + Needs Study

Delayed gratification
CS109 View of Probability

Teach you how to write programs that most people are not able to write.
CS109 View of Probability

Teach you the theory you need to do the math that most people are not able to do.
CS109

AI

Uncertainty Theory

Single Random Variables

Probabilistic Models

Counting

Probability Fundamentals
Let's dive in...
2 min pedagogic pause.
Counting I
What is Counting?

An experiment in probability:

Counting: How many possible outcomes can occur from performing this experiment?
What is Counting?

Roll

\[ \{1, 2, 3, 4, 5, 6\} \]

Roll even only

\[ \{2, 4, 6\} \]

\[
\{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), \\
(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), \\
(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), \\
(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\
(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), \\
(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}
\]
Step Rule of Counting (aka Product Rule of Counting)

If an experiment has two steps, where

The first step’s outcomes are from Set $A$, where $|A| = m$,  
and the second step's outcomes are from Set $B$, where $|B| = n$,  
and $|B|$ is unaffected by outcome of first step.

Then the number of outcomes of the experiment is

$$|A||B| = mn.$$
How Many Unique Images?

Each pixel can be one of 17 million distinct colors

(a) 12 million pixels  
(b) 300 pixels  
(c) 12 pixels

$(17 \text{ million})^n$
How Many Unique Images?

Each pixel can be one of 17 million distinct colors

(a) 12 million pixels
\[ \approx 10^{86696638} \]

(b) 300 pixels
\[ \approx 10^{2167} \]

(c) 12 pixels
\[ \approx 10^{86} \]

(17 million)^

Piech + Cain, CS109, Stanford University
Sum Rule of Counting

If the outcome of an experiment can be either from
Set $A$, where $|A| = m$,  

or Set $B$, where $|B| = n$,  

where $A \cap B = \emptyset$,  

Then the number of outcomes of the experiment is  
$|A| + |B| = m + n$.  

One experiment

A

B
How many routes

**Question:** All routes considered by google maps from Nairobi to Dar es Salaam go through either Mt Kilimanjaro or Mombasa. How many total routes are considered?

- 20 routes go through Mt Kili
- 10 routes go through Mombasa
- 0 go through both

**Answer:** 20 + 10
## How Many Bit Strings?

**Problem:** A 6-bit string is sent over a network. The valid set of strings recognized by the receiver must either start with "01" or end with "10". How many such strings are there?

### Answer

<table>
<thead>
<tr>
<th>2⁴ start with 01</th>
<th>2⁴ end with 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>010000</td>
<td>000010</td>
</tr>
<tr>
<td>010001</td>
<td>000110</td>
</tr>
<tr>
<td>010010</td>
<td>001010</td>
</tr>
<tr>
<td>010011</td>
<td>001110</td>
</tr>
<tr>
<td>010100</td>
<td>010010</td>
</tr>
<tr>
<td>010101</td>
<td>010110</td>
</tr>
<tr>
<td>010110</td>
<td>011010</td>
</tr>
<tr>
<td>010111</td>
<td>011110</td>
</tr>
<tr>
<td>011000</td>
<td>100010</td>
</tr>
<tr>
<td>011001</td>
<td>100110</td>
</tr>
<tr>
<td>011010</td>
<td>101010</td>
</tr>
<tr>
<td>011011</td>
<td>101110</td>
</tr>
<tr>
<td>011100</td>
<td>110010</td>
</tr>
<tr>
<td>011101</td>
<td>110110</td>
</tr>
<tr>
<td>011110</td>
<td>111010</td>
</tr>
<tr>
<td>011111</td>
<td>111110</td>
</tr>
</tbody>
</table>

Set $A$  
Set $B$
# How Many Bit Stings?

**Problem:** A 6-bit string is sent over a network. The valid set of strings recognized by the receiver must either start with "01" or end with "10". How many such strings are there?

<table>
<thead>
<tr>
<th>2⁴ start with 01</th>
<th>2⁴ end with 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>010000</td>
<td>000010</td>
</tr>
<tr>
<td>010001</td>
<td>000110</td>
</tr>
<tr>
<td>010010</td>
<td>001010</td>
</tr>
<tr>
<td>010011</td>
<td>001110</td>
</tr>
<tr>
<td>010100</td>
<td>010010</td>
</tr>
<tr>
<td>010101</td>
<td>010110</td>
</tr>
<tr>
<td>010110</td>
<td>011010</td>
</tr>
<tr>
<td>010111</td>
<td>011110</td>
</tr>
<tr>
<td>011000</td>
<td>100010</td>
</tr>
<tr>
<td>011001</td>
<td>100110</td>
</tr>
<tr>
<td>011010</td>
<td>101010</td>
</tr>
<tr>
<td>011011</td>
<td>101110</td>
</tr>
<tr>
<td>011100</td>
<td>110010</td>
</tr>
<tr>
<td>011101</td>
<td>110110</td>
</tr>
<tr>
<td>011110</td>
<td>111010</td>
</tr>
<tr>
<td>011111</td>
<td>111110</td>
</tr>
</tbody>
</table>

Set A

Set B

**Answer**
### How Many Bit Strings?

**Problem:** A 6-bit string is sent over a network. The valid set of strings recognized by the receiver must either start with "01" or end with "10". How many such strings are there?

<table>
<thead>
<tr>
<th>2^4 start with 01</th>
<th>2^4 end with 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>010000</td>
<td>000010</td>
</tr>
<tr>
<td>010001</td>
<td>000110</td>
</tr>
<tr>
<td><strong>010010</strong></td>
<td><strong>001010</strong></td>
</tr>
<tr>
<td>010011</td>
<td>001110</td>
</tr>
<tr>
<td>010100</td>
<td>010010</td>
</tr>
<tr>
<td>010101</td>
<td>010110</td>
</tr>
<tr>
<td><strong>010110</strong></td>
<td><strong>011010</strong></td>
</tr>
<tr>
<td>010111</td>
<td>011110</td>
</tr>
<tr>
<td>011000</td>
<td>100010</td>
</tr>
<tr>
<td>011001</td>
<td>100110</td>
</tr>
<tr>
<td><strong>011010</strong></td>
<td><strong>101010</strong></td>
</tr>
<tr>
<td>011011</td>
<td>101110</td>
</tr>
<tr>
<td>011100</td>
<td>110010</td>
</tr>
<tr>
<td>011101</td>
<td>110110</td>
</tr>
<tr>
<td><strong>011110</strong></td>
<td><strong>111010</strong></td>
</tr>
<tr>
<td>011111</td>
<td>111110</td>
</tr>
</tbody>
</table>

**Answer**

Set A

Set B
## How Many Bit Stings?

**Problem:** A 6-bit string is sent over a network. The valid set of strings recognized by the receiver must either start with "01" or end with "10". How many such strings are there?

<table>
<thead>
<tr>
<th>2⁴ start with 01</th>
<th>2⁴ end with 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>010000</td>
<td>000010</td>
</tr>
<tr>
<td>010001</td>
<td>000110</td>
</tr>
<tr>
<td><strong>010010</strong></td>
<td><strong>001010</strong></td>
</tr>
<tr>
<td>010011</td>
<td>001110</td>
</tr>
<tr>
<td>010100</td>
<td><strong>010010</strong></td>
</tr>
<tr>
<td>010101</td>
<td><strong>010110</strong></td>
</tr>
<tr>
<td><strong>010110</strong></td>
<td><strong>011010</strong></td>
</tr>
<tr>
<td>010111</td>
<td><strong>011110</strong></td>
</tr>
<tr>
<td>011000</td>
<td>100010</td>
</tr>
<tr>
<td>011001</td>
<td>100110</td>
</tr>
<tr>
<td><strong>011010</strong></td>
<td><strong>101010</strong></td>
</tr>
<tr>
<td>011011</td>
<td>101110</td>
</tr>
<tr>
<td>011100</td>
<td>110010</td>
</tr>
<tr>
<td>011101</td>
<td>110110</td>
</tr>
<tr>
<td><strong>011110</strong></td>
<td><strong>111010</strong></td>
</tr>
<tr>
<td>011111</td>
<td>111110</td>
</tr>
</tbody>
</table>

Set $A$

Set $B$

### Answer

$$N = |A| + |B| - |A \text{ and } B|$$

$$= 16 + 16 - 4$$

$$= 28$$
Or Rule of Counting (aka Inclusion/Exclusion)

If the outcome of an experiment can be either from

Set $A$, where $|A| = m$,

or Set $B$, where $|B| = n$,

where $A \cap B$ may not be empty,

Then the number of outcomes of the experiment is

$N = |A| + |B| - |A \cap B|$. 
Challenge Problem

1. Strings
   • How many *different* orderings of letters are possible for the string BAYES?
   • How about BOBA?

BOBA, ABOB, OBBA...
Concept Check!
Incredible time and school at which to study probability! Exciting.