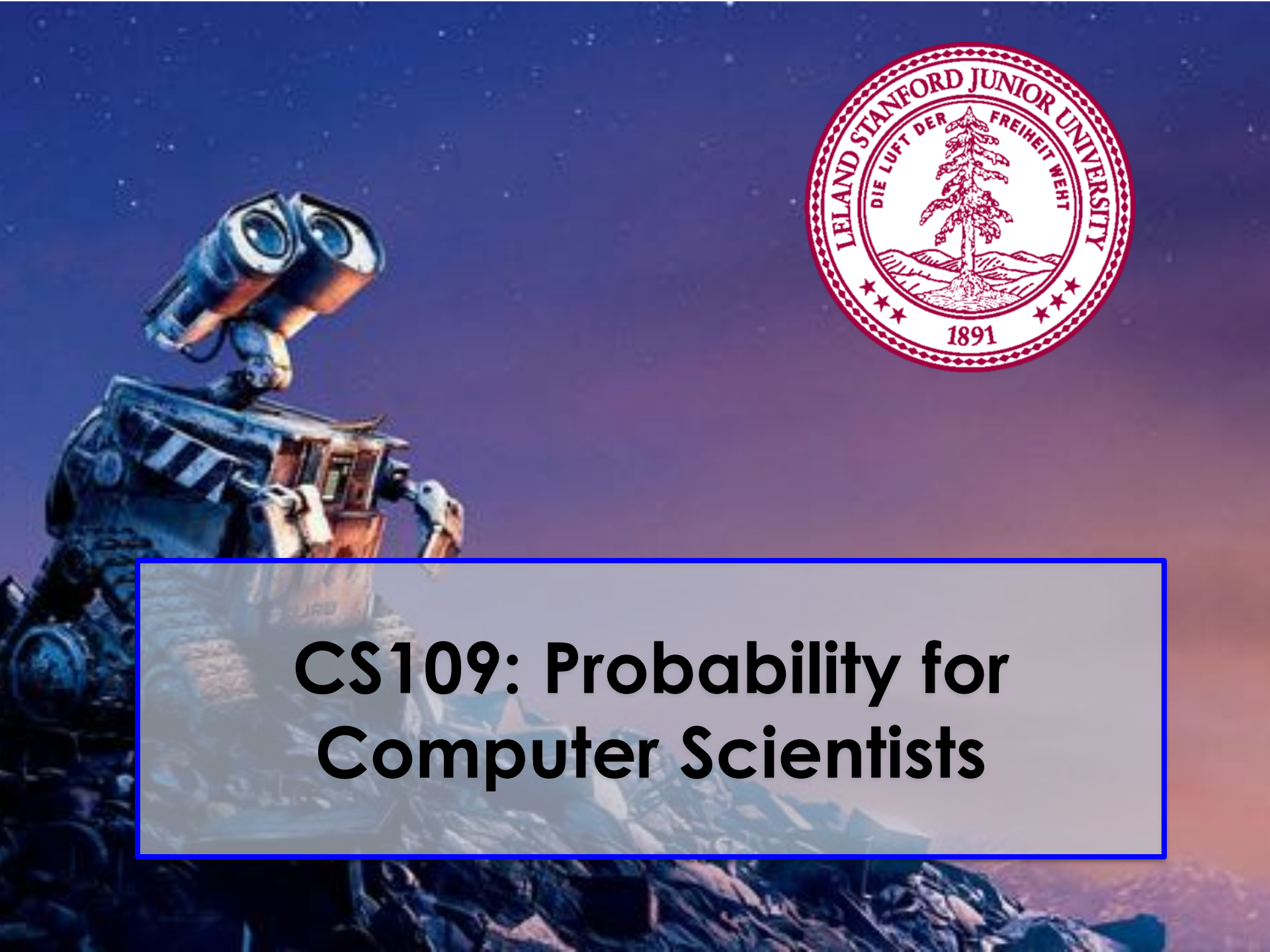


CS109: Probability for Computer Scientists

I am not Chris

Noah Arthurs

$\neq$

Chris Piech

I am not Chris

Noah Arthurs

≠

Chris Piech

April Fools?

I am not Chris

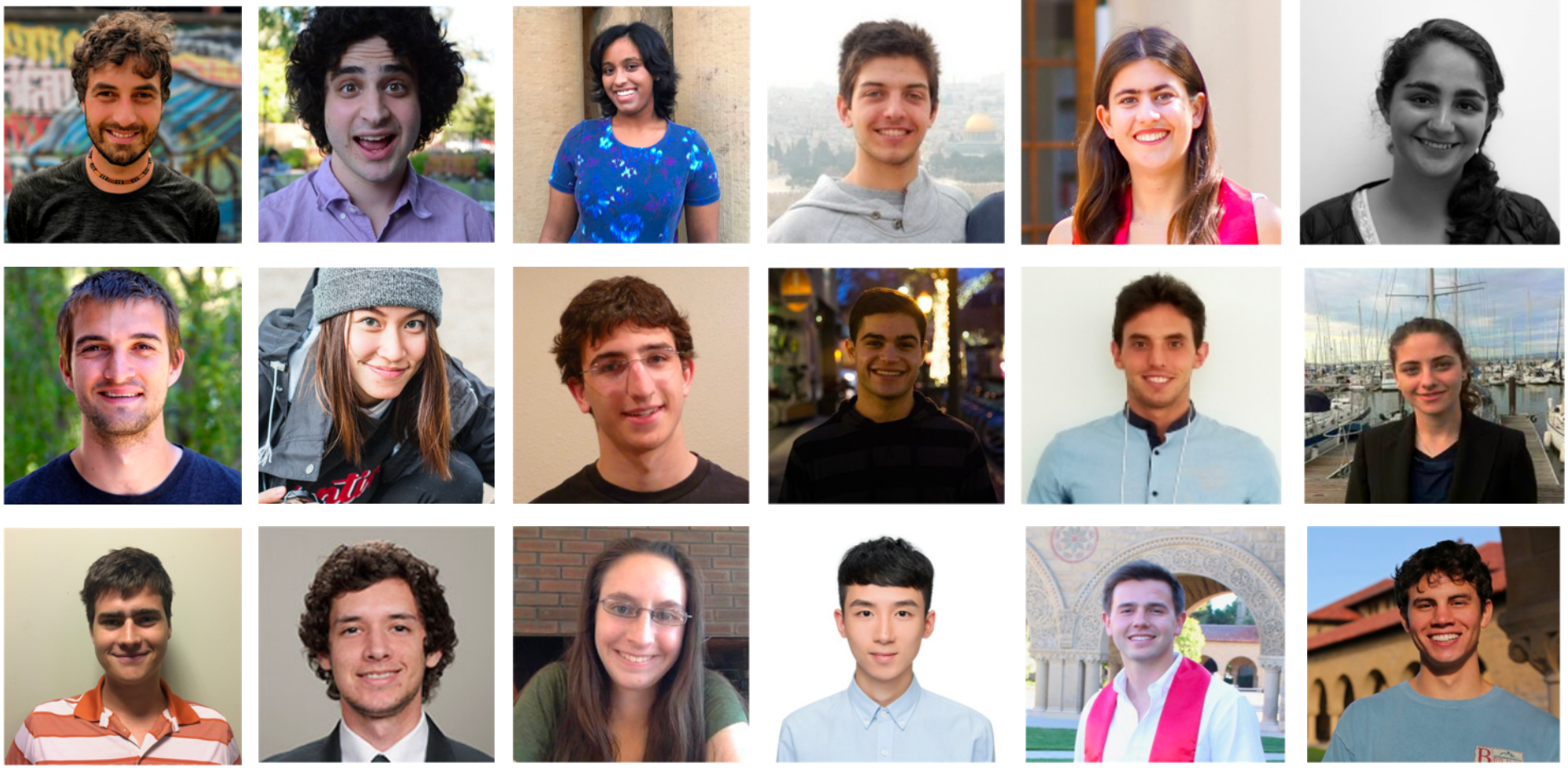
Noah Arthurs

≠

Chris Piech

~~April Fools?~~

Teaching Team



We love questions!

=

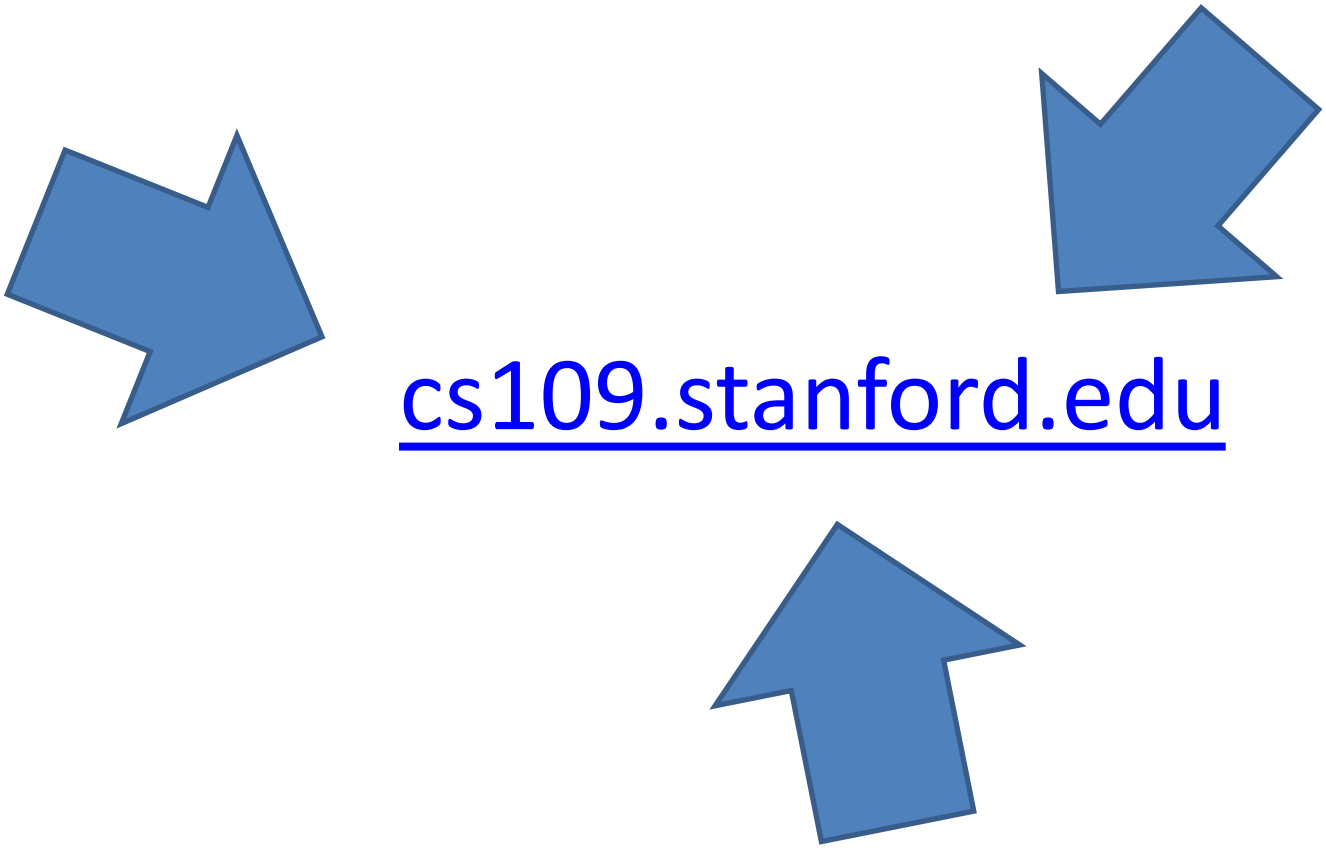
This Lecture

1. Course Mechanics
2. Why you should take CS109
3. Counting!

Course Mechanics

(This is a light version. Please read the handout for details.)

Essential Information



cs109.stanford.edu

Are you in the right place?

Prereqs

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

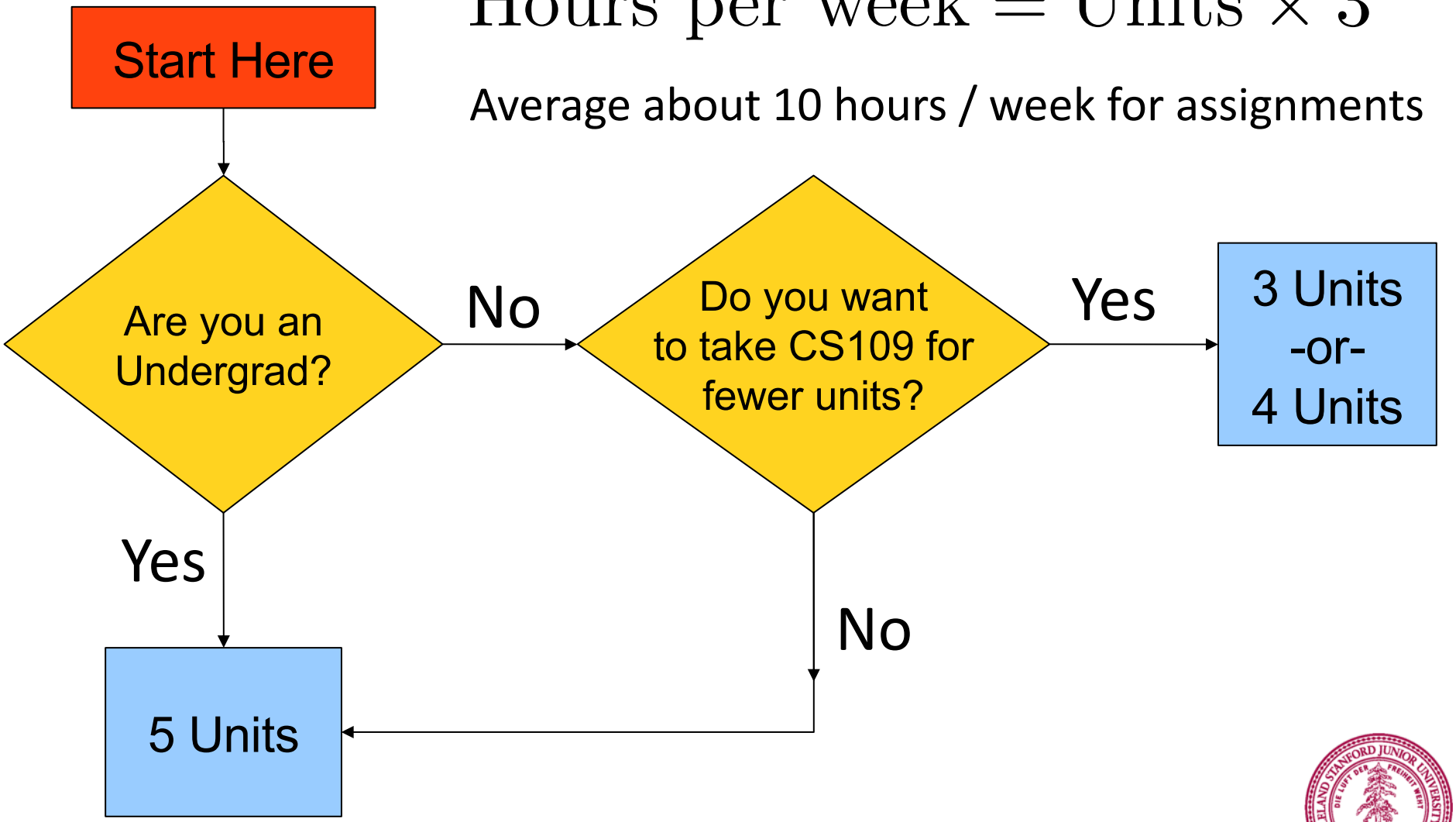
Staff Contact

- Post to Piazza for clarification
- Go to Working Office Hours
- Email cs109@cs.stanford.edu
- Email Chris or go to his office for course level issues.

CS109 Units

$$\text{Hours per week} = \text{Units} \times 3$$

Average about 10 hours / week for assignments



Videotaped*

*but there are many reasons you should come to lecture

Class Breakdown

45% **6 Assignments**

20% **Midterm**
Tuesday May 7th, 7-9pm

30% **Final**
Tuesday June 11th, 8:30-11:30m

5% **Section Participation**

Late Days

2


LELAND STANFORD JUNIOR UNIVERSITY
DIE LUFT DER FREIHEIT WEHT
1891

Textbook

Chris Piech
CS 109

Lecture Notes #1
Sept 24, 2018

Counting

Based on a handout by Mehran Sahami with examples by Peter Norvig

Although you may have thought you had a pretty good grasp on the notion of counting at the age of three, it turns out that you had to wait until now to learn how to really count. Aren't you glad you took this class now?! But seriously, below we present some properties related to counting which you may find helpful in the future.

Counting is important in the world of computer science for a few reasons. In order to understand probability on a fundamental level, it is useful to first understand counting. Moreover, while computers are fast, some problems require so much work that they would take an unreasonable amount of time to complete. We can use counting theory to reason about complexity.

1 Sum Rule

Sum Rule of Counting
If the outcome of an experiment can either be one of m outcomes or one of n outcomes, where none of the outcomes in the set of m outcomes is the same as the any of the outcomes in the set of n outcomes, then there are $m + n$ possible outcomes of the experiment.

Modern AI
or, How we learned to combine probability and programming

Something is going on in the world of AI

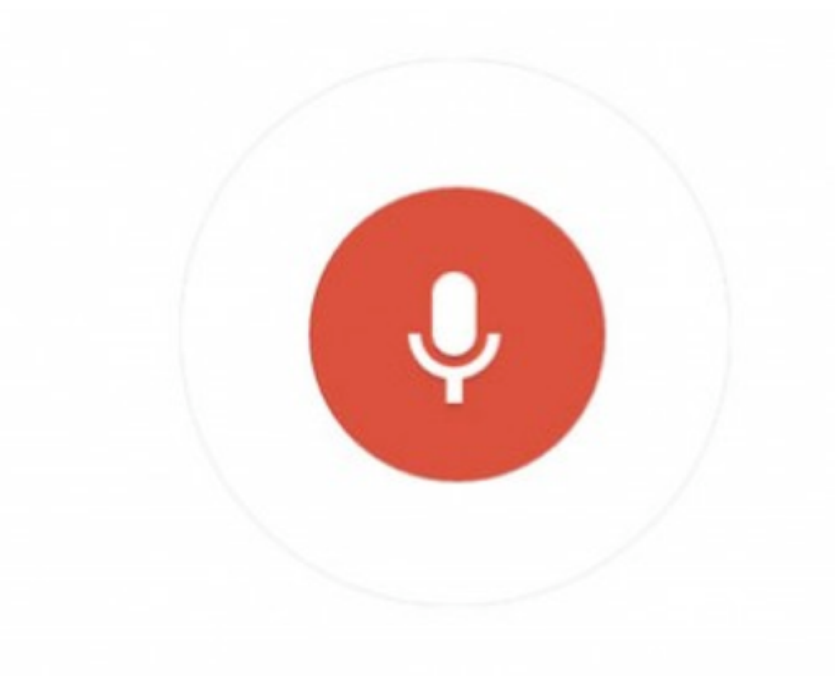
Big Milestones Pt 1

1997 Deep Blue

2005 Stanley

2011 Watson

Told Speech Was 30 Years Out



Almost perfect…

The Last Remaining Board Game

Computers Making Art

Self Driving Cars

What is going on?

Two Great Ideas

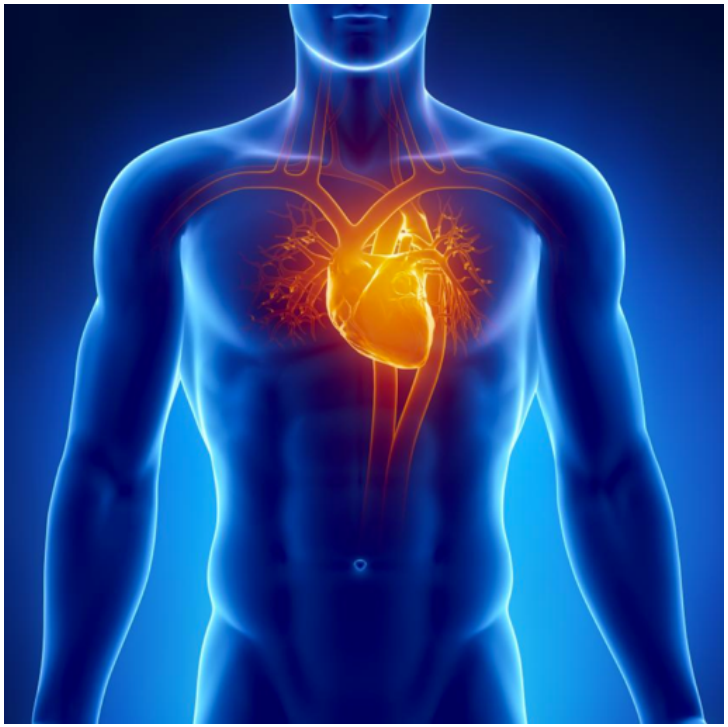
1. Probability from Examples

2. Artificial Neurons

Where you will be by the end of class

CS View of Probability

Heart



Ancestry

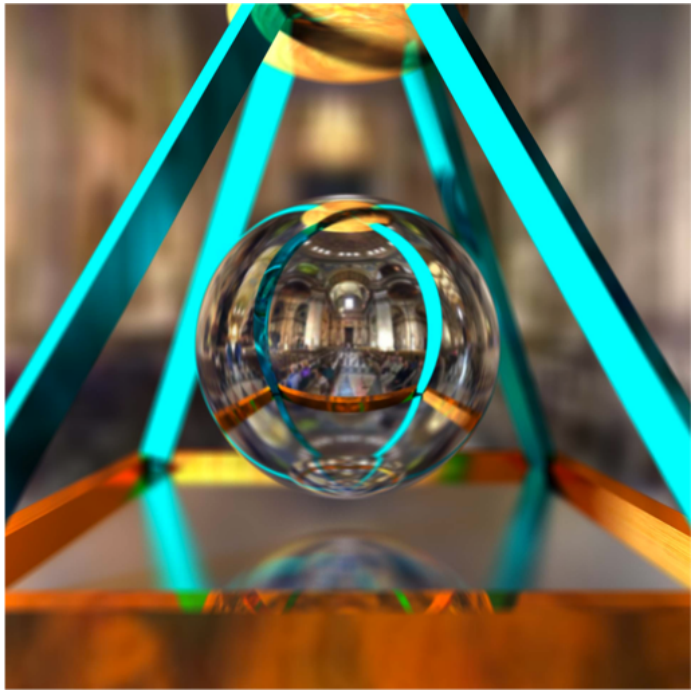
Netflix

Probability is more than just machine learning

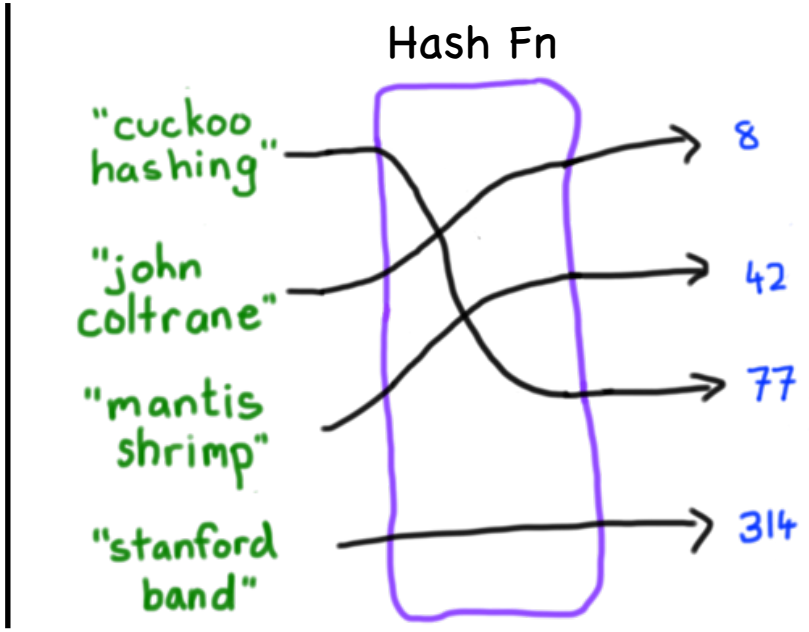
Abundance of Important Problems

Algorithms and Probability

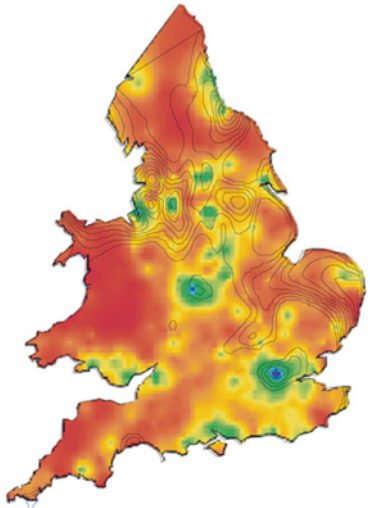
Eg Raytracing



Eg HashMaps



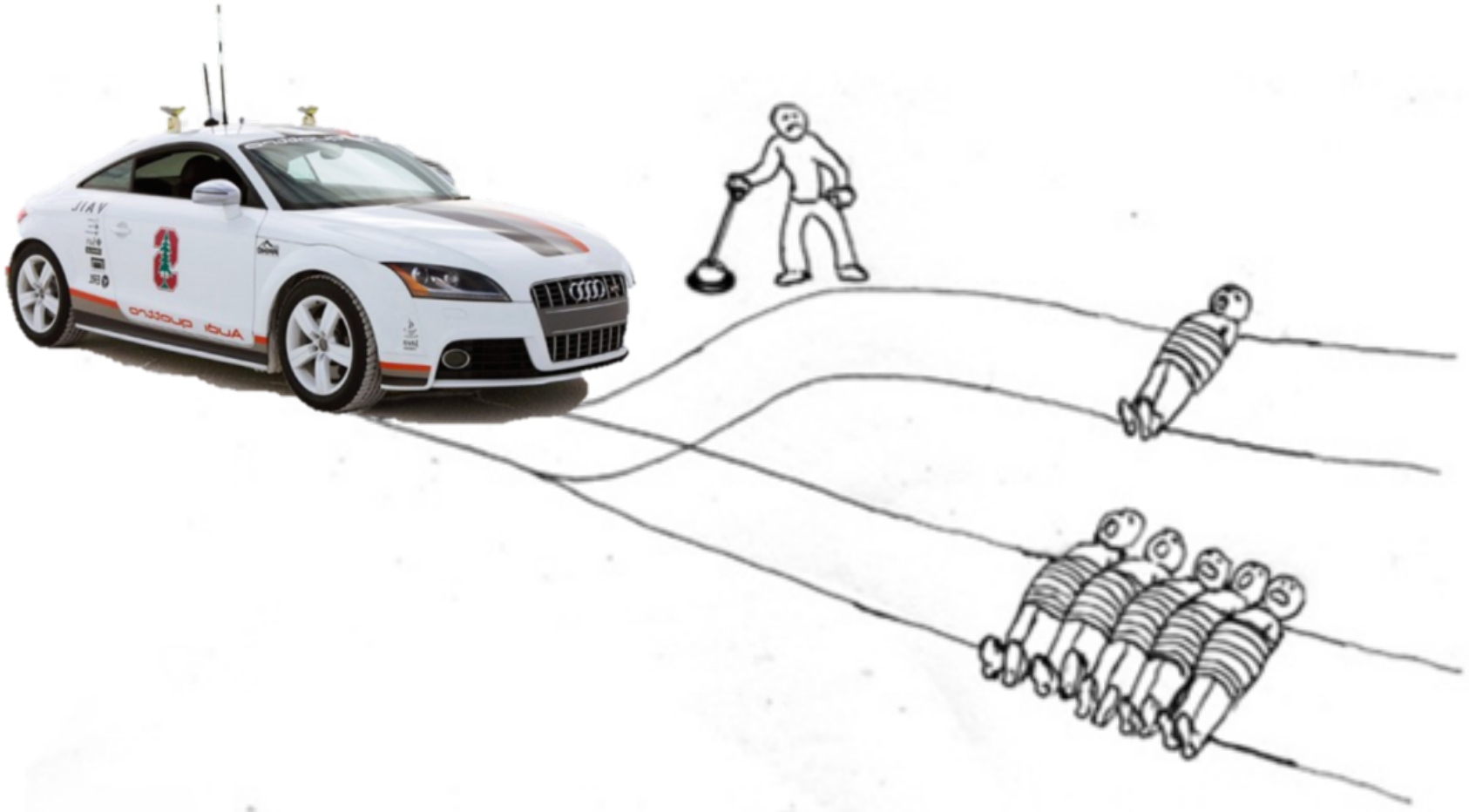
Medicine and Probability



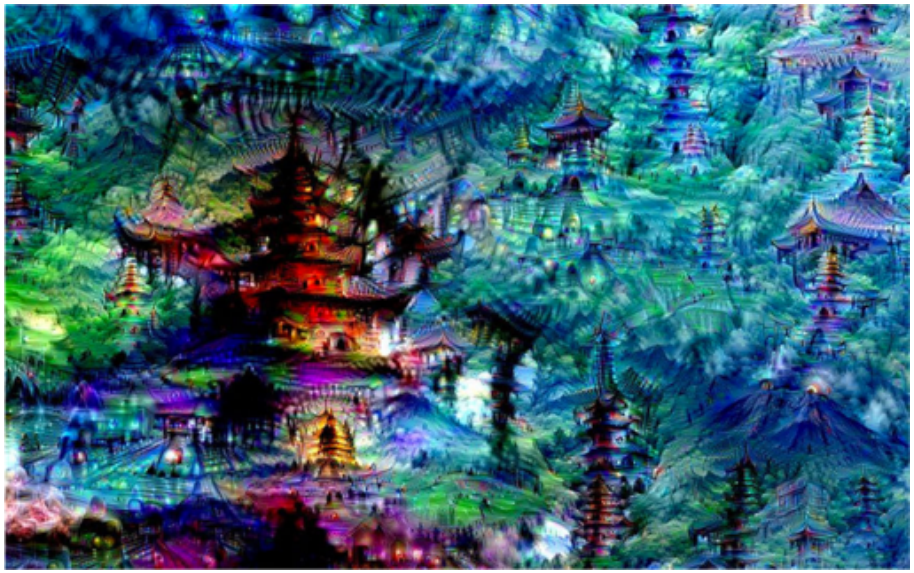
Autocomplete

Probability in Practice

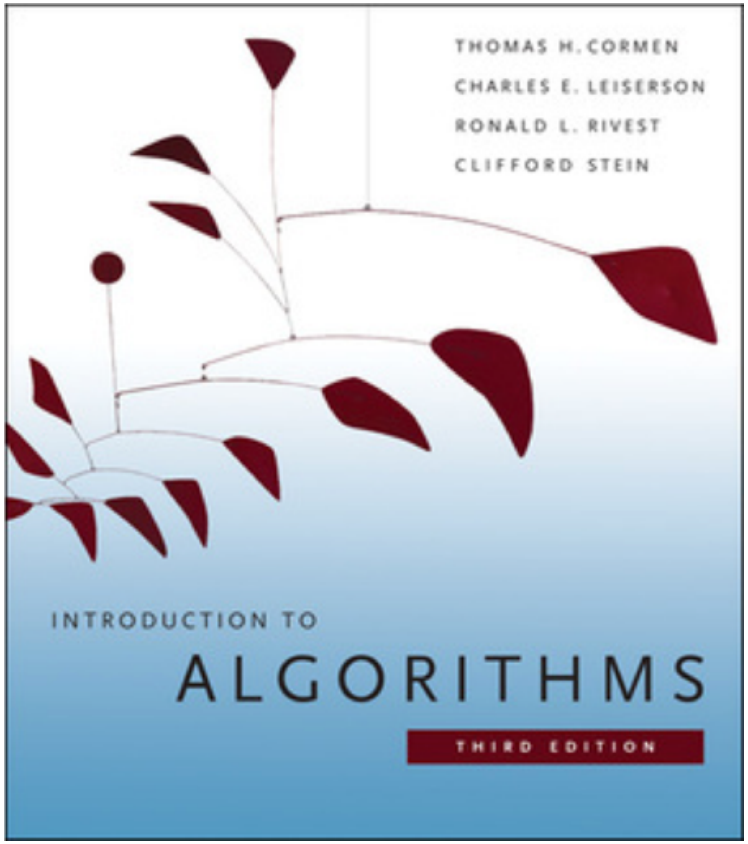
Philosophy and Probability

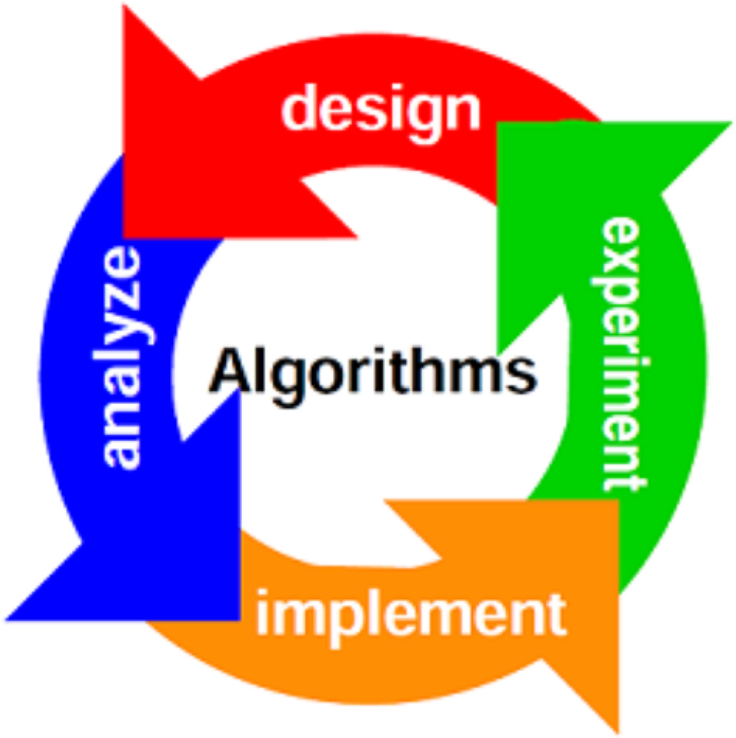


Art and Probability

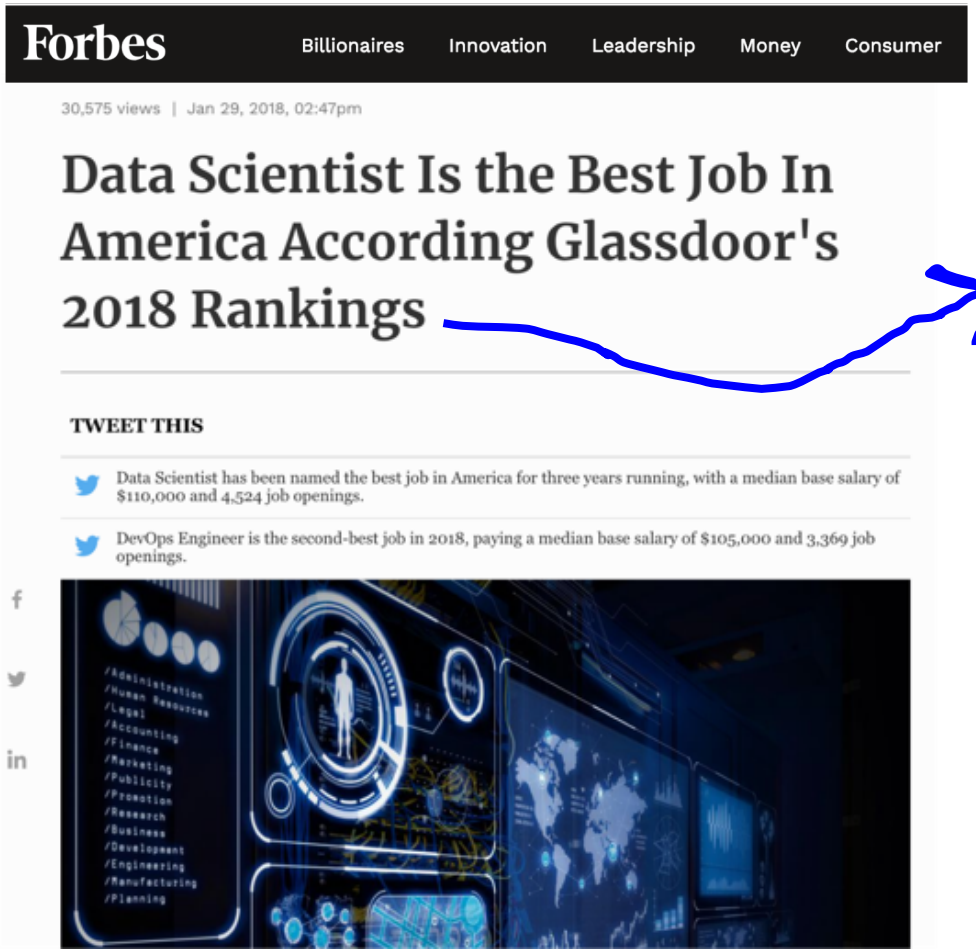


Probabilistic Analysis of Algorithms





#1 Most Desired Skill in Industry



Forbes Billionaires Innovation Leadership Money Consumer

30,575 views | Jan 29, 2018, 02:47pm

Data Scientist Is the Best Job In America According Glassdoor's 2018 Rankings

TWEET THIS

Data Scientist has been named the best job in America for three years running, with a median base salary of $110,000 and 4,524 job openings.

DevOps Engineer is the second-best job in 2018, paying a median base salary of $105,000 and 3,369 job openings.

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

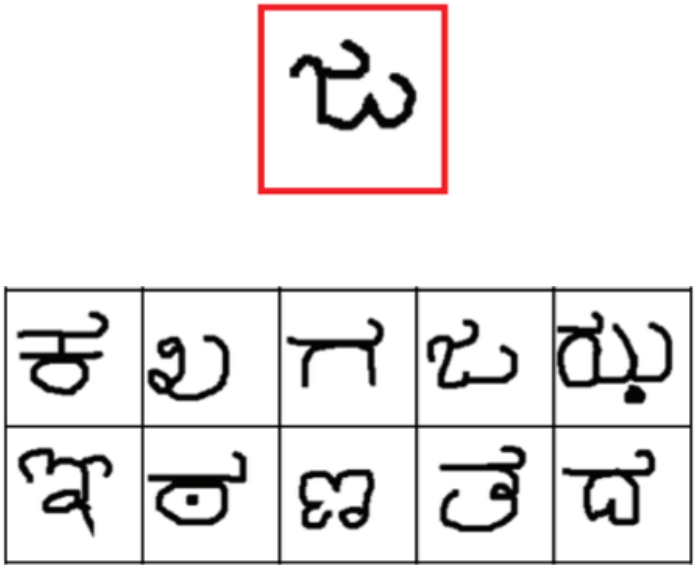
#1 Most Desired Skill in Academia

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

"Better understanding of probability"

Open Problem: One Shot Learning

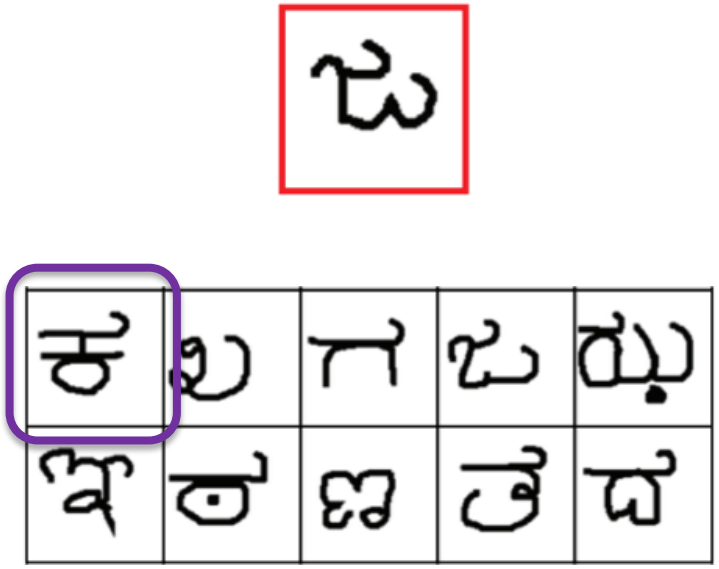
B Lake, R Salakhutdinov, J Tenenbaum. Science 2015.
Human-level concept learning through probabilistic program induction.



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

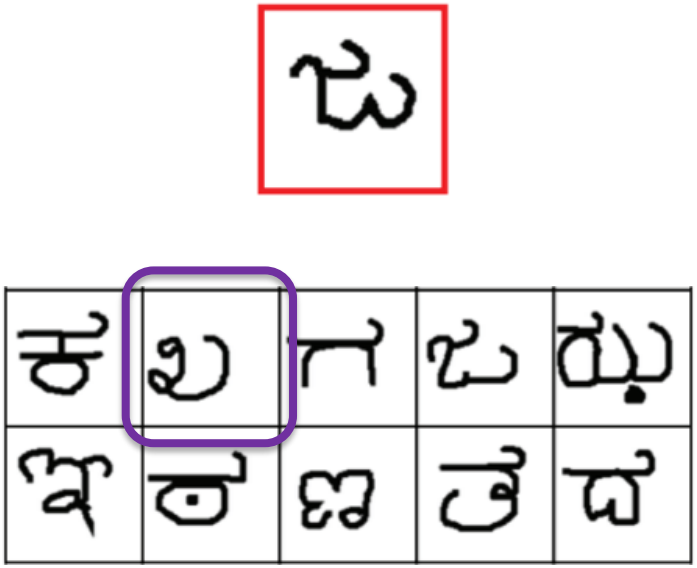
B Lake, R Salakhutdinov, J Tenenbaum. Science 2015.
Human-level concept learning through probabilistic program induction.



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

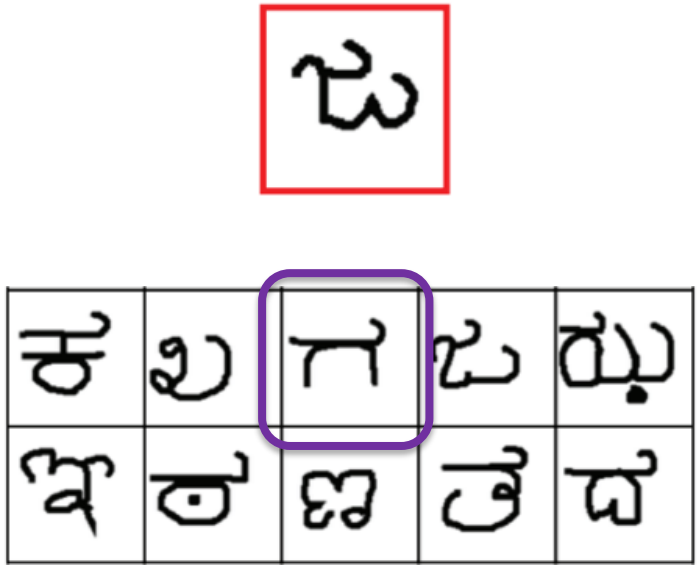
B Lake, R Salakhutdinov, J Tenenbaum. Science 2015.
Human-level concept learning through probabilistic program induction.



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

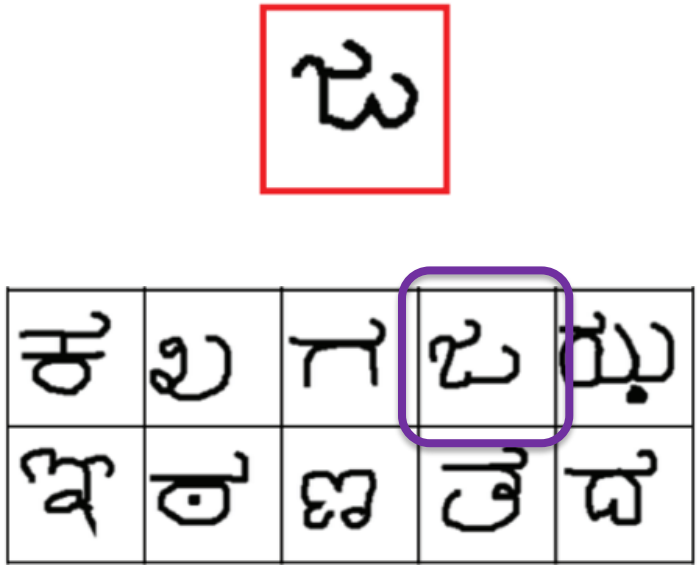
B Lake, R Salakhutdinov, J Tenenbaum. Science 2015.
Human-level concept learning through probabilistic program induction.



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

B Lake, R Salakhutdinov, J Tenenbaum. Science 2015.
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

Zika Test

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.

Lets dive in…

Counting