CS398: Computational Education
[warmup activity]

What was your favorite learning experience. What was your least favorite? Why?
Chris Piech

Created a research lab in:
**Computational Education**

Grew up in Nairobi, Kuala Lumpur before Stanford!

Piech, CS109, Stanford University
What Sort of Thing Does Chris Research?
Deployed to grade 16,000 submissions in Code.org

Algorithm uses attention to highlight where in the code the error comes from

AI generated feedback

Students evaluate the feedback

Syntax error (missing ") here would prevent auto graders from being useful.
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.
So many things to love in this world
Who are you!?
Course mechanics

(this is a light version. Please read the handout for details).
Essential Information

cs398.stanford.edu
What is CS398?
Graduate Level Class

**CS198**: Teach in small groups

**CS298**: Teach in big groups

**CS398**: Teaching ideas that scale

Can be code based, but doesn’t have to be
Special offering
Special Theme this Quarter
Prereqs

What you really need:

Coding:
• Very helpful

No coding?
• We can design an experience. You will need to work.
Learning Goal

Students should be ready to produce a cutting edge project in computational education. Either:

(1) A novel research paper or
(2) A novel policy proposal or
(3) A novel learning tool
Path

Exploratory Learning

Common experience with great ideas

Final Project

We have done a good job if you are taking steps towards your potential
Assignments in class
Staff Contact

• Meet me in class! Ask me questions. Get to know me.
• Go to Office Hours
• Post a question to the slack channel
• Email piech@cs.stanford.edu
The Grand Challenge
People have valued and thought deeply about education for a long time

John Dewey, 1923
90% of children enroll in primary education [1]

40% in secondary education [1]

20% in tertiary education [1]

Dramatic quality differences

375 million workers need to be retrained by 2030 [3]

Half a million unfilled computer science jobs (60% of STEM jobs) [2]

For all learners we want quality.

Cost Disease of Education

Direct Costs Per Student, Compared with an Economy-Wide Cost Index

Index, 1904-05 = 100

Fiscal Year, beginning 1904-05

In 2022, there is a unique opportunity to make progress.
Told Speech was 30 years out

Almost perfect...
The Last Remaining Boardgame
Self Driving Cars
Computers Making Art
Smart Phone Access

Advanced Economies

Emerging Economies

- Smartphone
- Mobile
- No phone
Code.org

= 500,000 learners
50,000,000 K12 students in the US

1,234,127 teachers
41,909,399 unique enrolled students
Used in 180+ countries
832M hour of code sessions
Clear Societal Need

Increase education productivity

Grand Challenge in Education

Use the power of computers and computation to support learning

=> Open problems in education
=> Open problems in AI

MOOC assignments

Deep Learning

Digital Learning

AI Renaissance
In 2022, there is a unique opportunity to make progress.
Pandemic showed downsides and upsides!
But what should educations **digital future** look like?
The New York Times

The Year of the MOOC

By LAURA PAPPANO  NOV. 2, 2012
Two stories of technology that can help
Story #1: Technology that can help teachers understand their students
Feedback is Labor Intensive

Online classes have not solved the feedback problem [1].

Computational Education Paradigm

A student’s past → Understanding of the student → Predictions of their future

→ Give feedback (hints + summative)
→ Visualize knowledge to aid teaching
→ Make optimal learning policies
Many domains of student work

Why did the original colonists come to America?
Chapter 1: Deep Knowledge Tracing

*Feedback in a simple context*
Story of Riley

Exercise Type:
- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
- Slope of a line

Answer:
- Correct
- Incorrect
Story of Riley

Exercise Type:  
- Solving for x-intercept  
- Solving for y-intercept  
- Graphing linear equations  
- Square roots  
- Slope of a line

Answer:  
- Correct  
- Incorrect

Exercise index:  
1 10 20 30 40 50
What should Riley do next?

Exercise Type:  
- Purple: Solving for x-intercept  
- Blue: Solving for y-intercept  
- Green: Graphing linear equations  
- Yellow: Square roots  
- Red: Slope of a line

Answer:  
- Correct  
- Incorrect
Given $n$ historical answers:

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

Predict the next one

$x_{n+1}$
Build the **first** deep learning algorithm for human learning
Understanding Students

**Benchmark AUC**

<table>
<thead>
<tr>
<th></th>
<th>Marginal</th>
<th>BKT</th>
<th>BKT*</th>
<th>DKT</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
</tr>
</tbody>
</table>

**Khan AUC**

<table>
<thead>
<tr>
<th></th>
<th>Marginal</th>
<th>BKT</th>
<th>DKT</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
</tr>
</tbody>
</table>
Learns Concept Relationships
Finding the x-intercept
Finding the y-intercept
Learns Concept Relationships

Slope of a line
Graphing Linear Equations
Learns Concept Relationships

Graphing Systems of Equations
Can it help make decisions?
Predicting Next Item

Exercise attempted: • correct, ○ incorrect

Exercise index

Predicted Probability

- Line graph intuition
- Slope of a line
- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
Predicting Next Item

Exercise attempted: • correct • incorrect

- Line graph intuition
- Slope of a line
- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
Predicting Next Item

Exercise attempted: • correct, ○ incorrect

Exercise index

Predicted Probability

0.0

0.5

1.0

Line graph intuition
Solving for x-intercept
Solving for y-intercept
Graphing linear equations
Square roots

The deal maker
Optimal Teaching

Exercise Index

Average Predicted Probability

Maximize knowledge after 30 questions

Deep Knowledge Tracing

Previous Algorithms

Blocking

Mixing
Chapter 2: Zero Shot Feedback
Online Classes Haven’t Solved the Problem

Feedback is hard
Online Classes Haven’t Solved the Problem

This is hard.

Introduction to Python

```
1  my_name = "Codecademy"
2  print("Hello and welcome " + my_name + "!")
3  SyntaxError: unexpected EOF while parsing
```

```
Hi! I'm Elsa from Arendelle. Help me create a single line.

Not quite. You have to use a block you aren't using yet.

Block: 1 / 2 blocks
```

```
move forward 100 pixels
when run
  turn left by 90 degrees
  turn right by 90 degrees
  turn left by 90 degrees
```

The Code.org Dataset

- Students learning nested loops
- 50k students with 1.5 million submissions to a curriculum of 8 exercises.
- 800 human labels across 2 of the exercises.
Let's put it all together!

Using your knowledge of for loops and the counter variable, create this drawing where each shape has two more sides than the last. Make sure that each side is 10 times as long as the number of sides in the polygon.

Very little of the code has been provided for you.
Let's put it all together!

Using your knowledge of `for` loops and the `counter` variable, create this drawing where each shape has two more sides than the last. Make sure that each side is 10 times as long as the number of sides in the polygon.

Very little of the code has been provided for you.

```
when run
  for counter from 3 to 9 count by 2
```

---

**P8** 59,615 / 36,727
Let's put it all together!

Using your knowledge of `for` loops and the `counter` variable, create this drawing where each shape has two more sides than the last. Make sure that each side is 10 times as long as the number of sides in the polygon.

Very little of the code has been provided for you.
Let's put it all together!

Using your knowledge of for loops and the counter variable, create this drawing where each shape has two more sides than the last. Make sure that each side is 10 times as long as the number of sides in the polygon.

Very little of the code has been provided for you.
**Instructions**

- If there are many moves, focus on the first one.
- Random code strategy is for when the student seems to be trying things randomly.
- Lookout for students who don't get nesting or pre/post conditions.
- Often extra blocks in a body is an indication that they don't get that the post of the loop has to match the precondition.

**Question**

```
when run
move forward ▼ by 100 pixels
repeat 5 times
do
  turn right ▼ by 90 degrees
  move forward ▼ by 10 pixels
```

**Label Console**

- **Strategy**
  - Beeper Boundary (most people do this)
  - Triangle Strategy
  - Recursive Strategy

- **Looping**
  - Correct use of looping
  - Doesn’t use a while
  - Doesn’t have correct stop condition
  - Body is missing statements
  - Body has extra statements
  - Body order is incorrect
  - Sets up initial precondition
  - Doesn’t get nesting
  - Loop post condition doesn’t match preconditions

- **Cleanup**
  - Repetition of bodies
It is a very hard problem
Terribly Clever Static Analysis

A student’s program

A bug

Nguyen et Al, WWW 2014
// User defined method
private void run() {
    while(isClear()){
        putBeeper();
        move();
    }
    putBeeper();
}
method step() {
    putBeeper();
    move();
}

Autoencoding Loss

Prediction Loss

Piech et al., ICML 2014
Can we provide feedback by dynamic analysis?
Each edge is what a teacher suggested.

Pink dots are students.

Each node is a unique partial solution.

Solution
http://stanford.edu/~cpiech/demos/research/blossoms.html
The Crowd is Un-wise

Temporal methods tried:
- Shortest path
- Min Time
- Expected Success
- Reinforcement learning
- Most Common Next
- Most Popular Path

18%
45%
12%
SOTA uses RNNs to vectorize programs and classify among K feedback classes.

These models are ...

- *Far from human accuracy.*
- *Uninterpretable... why pick the feedback it did?*
- *Require lots of labeled examples.*

Wang et al, EDM 2017
What Matters for Students?

1. Two compound errors
2. Solves first error
3. Starts reasonable attempt
4. Completes attempt
5. Backtracks
6. Finds solution

Lisa Wang, Chris Piech. Deep Knowledge Tracing on Programming Exercises. L@S 2017
Highly Rates Grit

1. Two compound errors

2. Solves first error

3. Starts reasonable attempt

4. Completes attempt

5. Backtracks

6. Finds solution

Lisa Wang, Chris Piech. Deep Knowledge Tracing on Programming Exercises. L@S 2017
Highly Rates Grit

1. Two compound errors

2. Solves first error

3. Starts reasonable attempt

4. Completes attempt

5. Backtracks

6. Finds solution

Lisa Wang, Chris Piech. Deep Knowledge Tracing on Programming Exercises. L@S 2017
Highly Rates Grit

1. Two compound errors

2. Solves first error

3. Starts reasonable attempt

4. Completes attempt

5. Backtracks

6. Finds solution

Lisa Wang, Chris Piech. Deep Knowledge Tracing on Programming Exercises. L@S 2017
Highly Rates Grit

1. Two compound errors
2. Solves first error
3. Starts reasonable attempt

4. Completes attempt
5. Backtracks

6. Finds solution

Lisa Wang, Chris Piech. Deep Knowledge Tracing on Programming Exercises. L@S 2017
Traditional Deep Learning Doesn’t Work

Label student code

Feedback F1 Score

Last Problem (P8)

Old Gaurd  Humans

0.0  0.2  0.4  0.6  0.8  1.0

Old Gaurd

Humans
Inaccurate, Uninterpretable, and Data Hungry

Label student code

Last Problem (P8)

Feedback F1 Score

Old Gaurd | Deep Learning | Humans

Inaccurate, Uninterpretable, and Data Hungry

Piech et al., ICML 2014
Label student code

Inaccurate, Uninterpretable, and Data Hungry

Last Problem (P8)

Feedback F1 Score

Old Gaurd
Deep Learning
Humans
Label student code

Data Hungry

Last Problem (P8)

Feedback F1 Score

- Old Gaurd
- Deep Learning
- Humans

We need one shot learning
We need verifiability
[suspense]
Taste of the future of AI
Machine Learning Uses a Lot of Data
Humans Don’t Need Much Data

Single training example:

Test set:
Bayesian Program Learning

A. i) primitives

B. procedure GENERATETYPE

   \[ \kappa \leftarrow P(\kappa) \]  
   > Sample number of parts

   for \( i = 1 \ldots \kappa \) do
     \[ n_i \leftarrow P(n_i|\kappa) \]  
     > Sample number of sub-parts

     for \( j = 1 \ldots n_i \) do
       \[ s_{ij} \leftarrow P(s_{ij}|n_i) \]  
       > Sample sub-part sequence

     end for

     \[ R_\psi \leftarrow P(R_\psi|S_1, \ldots, S_{\kappa-1}) \]  
     > Sample relation

   end for

   \[ \psi \leftarrow \{\kappa, R, \psi\} \]  
   return \@GENERATETOKEN(\psi)  
   > Return program

procedure GENERATETOKEN(\psi)

   for \( i = 1 \ldots \kappa \) do
     \[ S_i^{(m)} \leftarrow P(S_i^{(m)}|S_i) \]  
     > Add motor variance

     \[ L_i^{(m)} \leftarrow P(L_i^{(m)}|R_\psi, T_1^{(m)}, \ldots, T_{\kappa-1}^{(m)}) \]  
     > Sample part’s start location

     \[ T_i^{(m)} \leftarrow f(L_i^{(m)}, S_i^{(m)}) \]  
     > Compose a part’s trajectory

   end for

   \[ A^{(m)} \leftarrow P(A^{(m)}) \]  
   > Sample affine transform

   \[ f^{(m)} \leftarrow P(f^{(m)}|T^{(m)}, A^{(m)}) \]  
   > Sample image

return \( f^{(m)} \)
1. Enabled teachers to write a student Bayesian Program.

2. We learned to infer into their Bayesian Program (auditable, novel).

- Struggle with double for loops
- Confuses logic for deleting bricks

```plaintext
when run
move forward by 100 pixels
repeat 5 times
do
  turn right by 90 degrees
  move forward by 10 pixels
```

*  

*
Generative Understanding

Label student code

Feedback F1 Score

Last Problem (P8)

- Old Guard
- Deep Learning
- Humans

Struggle with double for loops
- Confuses logic for deleting bricks
Generative Understanding

Label student code

Last Problem (P8)

Feedback F1 Score

0.0

0.2

0.4

0.6

0.8

1.0

Old Gaurd

Deep Learning

Zero Shot Learning

Humans

Outstanding student paper award, AAAI 2019

Struggle with double for loops
Confuses logic for deleting bricks

Outstanding student code

• stubborn

Students

• stubborn

Gaurd

Humans
Not just for code
Not just for code

Results from early 2019
Many domains of student work

Why did the original colonists come to America?
Can we *grade* my CS1 * midterm more accurately than Stanford TAs?
Deployed to grade 16,000 submissions in Code.org

AI generated feedback

Students evaluate the feedback

Algorithm uses attention to highlight where in the code the error comes from

Syntax error (missing ") here would prevent auto graders from being useful.
Impact and Fairness Analysis (work in progress)

Rubric
- RGM (ours)
- Supervised
- * Autograde
- Code in Place TAs

Humans
- Stanford TAs
- Code in Place TAs

The New York Times

Stanford News

Among others
Story #1: Technology that can help teachers understand their students
Chapter 3: Helping Student Teachers

Presented at SIGCSE 2019

Lisa Yan

Annie Hu
Students should learn the *process* of how to solve programming problems.
But we aren’t providing feedback on process.

(We are providing feedback on final solutions)
Pensieve Tool
Students scoring in 99th percentile on midterm exam

Students scoring in ≤3rd percentile on midterm exam

Error
Students work faster and learn more

- **Baseline**: 7.0 hours
- **Exp Term**: 6.3 hours

**Pensieve Intervention**

- Hours Spent Coding (minus baseline)
- Probability Mass

**Homework**

**Midterm Ability**
Very Intelligent Tutor in Our Pockets
This will change who can be an amazing teaching.
Story #2: The near-peer teacher idea
China Reports First Death From New Virus

The coronavirus, which surfaced in the city of Wuhan, has put the region on alert, but there is no evidence among humans.

W.H.O. Declares Global Emergency as Wuhan Coronavirus Spreads

The announcement came as nearly 10,000 cases have been reported worldwide.
Multiple pandemics

Pandemic of *loneliness*

Pandemic of *inequality*
What would it take to scale up high-quality, joyful education?
What do these two people have in common?

Marissa Meyers

Mike Schroepfer
What is a Section Leader?

1 Section Leader for every 10 students

Why do our intro classes feel so personal?

What undergrads are getting the best education?
We drastically underutilize human potential.

The magnitude of people who want to teach is roughly proportional to the magnitude of people who want to learn.

Teaching is learning.

Teaching is joyful.
We are confronted with the fierce urgency of now. In this unfolding conundrum of life and history, there “is” such a thing as being too late – Martin Luther King Jr
Stanford Code in Place

2,400+ novice teachers teach 24,000+ students 1/2 of CS106A As Community Service
10x the retention of the equivalent online course

99.7% teacher completion rate

9.7 average recommender strength

Hundreds of jobs found

6 experiments in Human-augmented AI
Do you think that education is solved?
If not why?
My job this quarter

Design

Coding Place

2023

(CP3)
Data Science in Place: Welcome Barbara
Spring 2022. April 19th to May 28th

Learn Collaboratively, now

523 other learners and teachers online

Data Science Mastery

Teaching Mastery
Overview of the Process

1. Learner who wants to improve
2. Find a good match
3. Co Learning
4. Learning experience
5. Generate insights
6. Repeat
7. Augment co-learning
Open Research Problems

Feedback to teachers?
Open Research Problems

Optimal learning ecosystem for refugees
Open Research Problems

Autonomous Tutor Gym
Open Research Problems

Feedback on the process of how someone solves a problem
Open Research Problems

Ad technology for more engaging handouts?
Open Research Problems

Can you translate educational material into different languages?

Open Research Problems

Can you give feedback to someone engaging in open-ended, unstructured learning?
Can your AI with **domain expertise** to give feedback?

A Breakthrough for A.I. Technology: Passing an 8th-Grade Science Test

By Cade Metz

Sept. 4, 2019

SAN FRANCISCO — Four years ago, more than 700 computer scientists competed in a contest to build artificial intelligence that
Hi Ayesha, you like skateboarding. If you jump off a half pipe at 2 km/h how high will you go?

What is the relationship between dinosaurs and birds?

Dinosaurs are a diverse group of reptiles of the clade Dinosauria. They first appeared during the Triassic period, between 243 and 233.23 million years ago, although the exact origin and timing of the evolution of dinosaurs is the subject of active research. They became the dominant terrestrial vertebrates after the Triassic–Jurassic extinction event 201 million years ago; their dominance continued through the Jurassic and Cretaceous periods. The fossil record demonstrates that birds are modern feathered dinosaurs, having evolved...
Automatic Generation of Contrasting Cases

Like tasting wines side-by-side, contrasts can improve discernment
Can you measure informal learning?

Greenhouse Gas - USA
Can you measure informal learning?

Greenhouse Gas - USA
Is it learned in school? When?

<table>
<thead>
<tr>
<th>Concept</th>
<th>When its taught (Spring, days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global warming</td>
<td>105</td>
</tr>
<tr>
<td>Fossil fuel</td>
<td>111</td>
</tr>
<tr>
<td>Sea level rise</td>
<td>111</td>
</tr>
<tr>
<td>Greenhouse Effect</td>
<td>114</td>
</tr>
<tr>
<td>Ocean Acidification</td>
<td>118</td>
</tr>
</tbody>
</table>

**Graph:**
- **General Topics:**
  - Global warming
  - Fossil fuel
  - Sea level rise
  - Greenhouse Effect
  - Ocean Acidification
- **School Topics:**
  - Global warming
  - Fossil fuel
  - Sea level rise
  - Greenhouse Effect
  - Ocean Acidification
Where are people learning about Greenhouse Effect
Human Education with AI

UNESCO
United Nations Educational, Scientific and Cultural Organization

Global Education Monitoring Report

[Image of a man]
A Vision of AI for Joyful Education

Here’s how we can avert the dangers and maximize the benefits of this powerful but still emerging technology

By Chris Piech, Lisa Einstein on February 26, 2020
What problems do you find interesting?

What is your vision for educations digital future?
Joy of Building
Now is the Time
Oh and Its Useful
Everyone is Welcome
Learn By Doing
Let's dive in
Brainstorm open problems