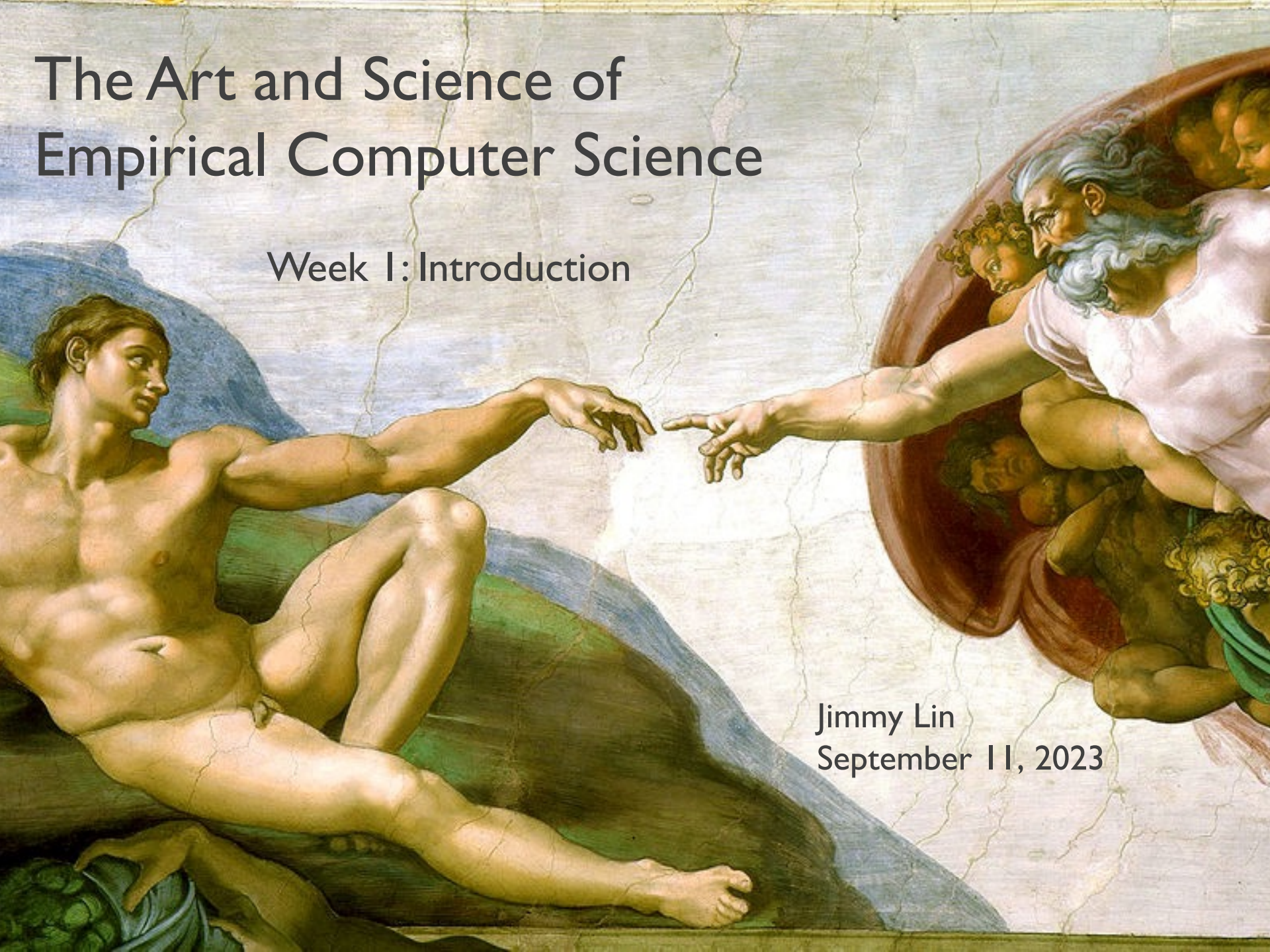
The Art and Science of Empirical Computer Science



Week 1: Introduction

Jimmy Lin
September 11, 2023

What does a grad student do?

How do you learn
to do these things?

Course Description

Graduate students in computer science aspire to "do computer science" (research), but what exactly does that mean? It involves, among a multitude of activities, reading papers, learning the "state of the art", advancing knowledge, writing papers, and (hopefully) getting them published. Graduate students learn how to do these things under tutelage of professors, but rarely is there explicit or deliberate instruction on these myriad activities. With a focus on empirical computer science, this course covers elements that comprise the research enterprise, synthesizing both "art" — personal experiences I have accumulated over the years — as well as "science" — insights derived from quantitative analyses. The hope is that knowledge and actionable advice from this course will help graduate students better understand research, hopefully leading to more productive and fulfilling careers.

Course Description

Graduate students in computer science aspire to "do computer science" (research), but what exactly does that mean? It involves, among a multitude of activities, reading papers, learning the "state of the art", advancing knowledge, writing papers, and (hopefully) getting them published. Graduate students learn how to do these things under tutelage of professors, but rarely is there explicit or deliberate instruction on these myriad activities. With a focus on empirical computer science, this course covers elements that comprise the research enterprise, synthesizing both "art" — personal experiences I have accumulated over the years — as well as "science" — insights derived from quantitative analyses. The hope is that knowledge and actionable advice from this course will help graduate students better understand research, hopefully leading to more productive and fulfilling careers.

Art / Science

What's the point of data science?

(and big data)

Data-driven decision making

(vs. HiPPO)

Evidence-based medicine

Evidence-based policymaking

Meta-research

Meta-research is the study of research itself: its methods, reporting, reproducibility, evaluation, and incentives. Given that science is the key driver of human progress, improving the efficiency of scientific investigation and yielding more credible and more useful research results can translate to major benefits.

Ioannidis. Meta-research: Why research on research matters. PLoS Biol 16(3):e2005468, 2018
https://doi.org/10.1371/journal.pbio.2005468

Scoping

Course Description

Graduate students in computer science aspire to "do computer science" (research), but what exactly does that mean? It involves, among a multitude of activities, reading papers, learning the "state of the art", advancing knowledge, writing papers, and (hopefully) getting them published. Graduate students learn how to do these things under tutelage of professors, but rarely is there explicit or deliberate instruction on these myriad activities. With a focus on empirical computer science, this course covers elements that comprise the research enterprise, synthesizing both "art" — personal experiences I have accumulated over the years — as well as "science" — insights derived from quantitative analyses. The hope is that knowledge and actionable advice from this course will help graduate students better understand research, hopefully leading to more productive and fulfilling careers.

Computer Science silos

Source: Wikipedia

silo noun

si·lo | \ ˈsī-(ˌ)lō \

plural **silos**

Definition of *silo* (Entry 1 of 2)

1 : a trench, pit, or especially a tall cylinder (as of wood or concrete) usually sealed to exclude air and used for making and storing silage

2 **a** : a deep bin for storing material (such as coal)

b : an underground structure for housing a guided missile

// These are weapons that can hit military targets, such as missile *silos* and headquarters, swiftly and accurately ...

— Leslie H. Gelb

3 : an isolated grouping, department, etc., that functions apart from others especially in a way seen as hindering communication and cooperation

// Big, complex companies are typically structured so that decision making is separated according to function, geography and product. That naturally creates *silos*.

— John H. Howard

// Examples of *silos* in blue-chip firms abound: Sony once had two separate divisions working on creating the same electrical plug without anyone realizing it.

— Rana Foroohar

// To break down *silos* individual officials need a stronger sense of belonging to a bigger whole, with norms and expectations spanning all departments and agencies.

— David Walker

// Moorhouse came to the paper determined to break down its "*silos*"—MBA-speak for self-contained departments that don't have anything to do with other departments in a business.

— Mark Fitzgerald

Source: Merriam-Webster Dictionary

Empirical Computer Science

Course Description

Graduate students in computer science aspire to "do computer science" (research), but what exactly does that mean? It involves, among a multitude of activities, reading papers, learning the "state of the art", advancing knowledge, writing papers, and (hopefully) getting them published. Graduate students learn how to do these things under tutelage of professors, but rarely is there explicit or deliberate instruction on these myriad activities. With a focus on empirical computer science, this course covers elements that comprise the research enterprise, synthesizing both "art" — personal experiences I have accumulated over the years — as well as "science" — insights derived from quantitative analyses. The hope is that knowledge and actionable advice from this course will help graduate students better understand research, hopefully leading to more productive and fulfilling careers.

Course Description

Graduate students in computer science aspire to "do computer science" (research), but what exactly does that mean? It involves, among a multitude of activities, reading papers, learning the "state of the art", advancing knowledge, writing papers, and (hopefully) getting them published. Graduate students learn how to do these things under tutelage of professors, but rarely is there explicit or deliberate instruction on these myriad activities. With a focus on empirical computer science, this course covers elements that comprise the research enterprise, synthesizing both "art" — personal experiences I have accumulated over the years — as well as "science" — insights derived from quantitative analyses. The hope is that knowledge and actionable advice from this course will help graduate students better understand research, hopefully leading to more productive and fulfilling careers.

Evidence-based decision making

… for your own career

Be deliberate about your career choices!

Be deliberate about your life choices!

known vs. unknown unknowns

Logistics

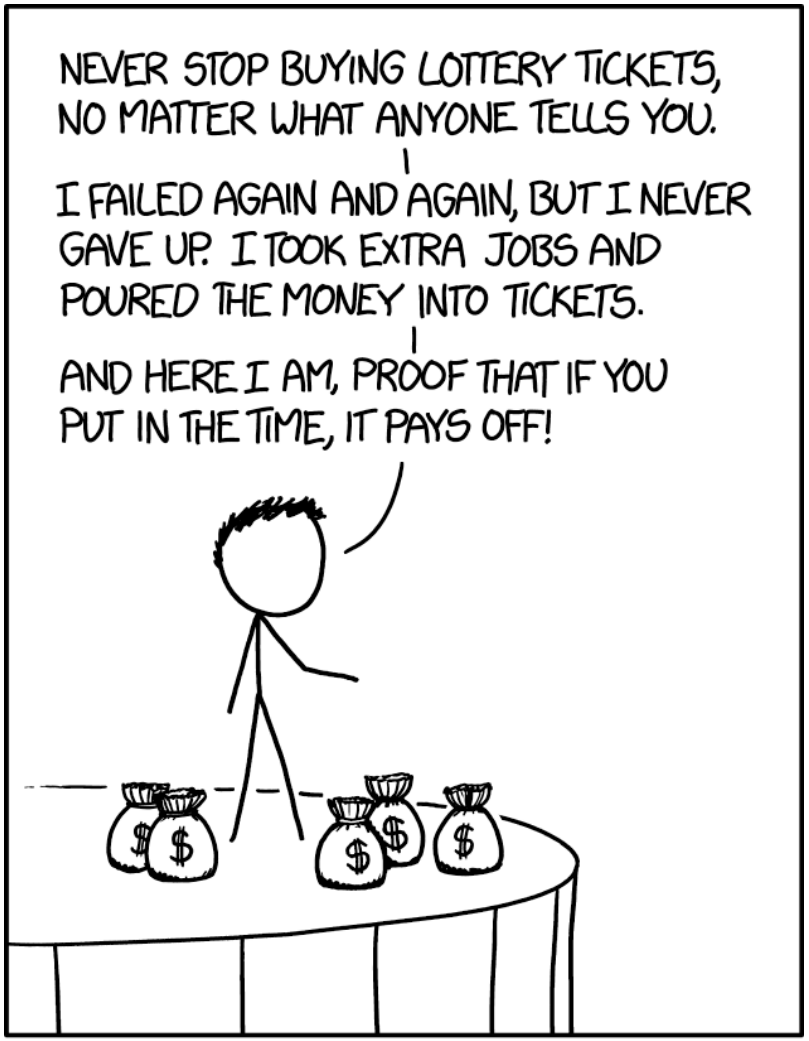
Course Description

Graduate students in computer science aspire to "do computer science" (research), but what exactly does that mean? It involves, among a multitude of activities, reading papers, learning the "state of the art", advancing knowledge, writing papers, and (hopefully) getting them published. Graduate students learn how to do these things under tutelage of professors, but rarely is there explicit or deliberate instruction on these myriad activities. With a focus on empirical computer science, this course covers elements that comprise the research enterprise, synthesizing both "art" — personal experiences I have accumulated over the years — as well as "science" — insights derived from quantitative analyses. The hope is that knowledge and actionable advice from this course will help graduate students better understand research, hopefully leading to more productive and fulfilling careers.

Be deliberate about your career choices!

Things to consider…

1. Recognize that there are multiple definitions of success and multiple ways to succeed.
2. Understand that choices come with opportunity costs.
3. Accept the role of skill as well as luck.
4. Think about distributions vs. samples.
5. Reflect on the differences between normative vs. positive approaches.



https://xkcd.com/1827/

That’s all for this week!