# Computer Science I 

5 December 2022

# How Technology Changes Our Perceptions and Decision-Making Capabilities 

## Prairie Goodwin, Vassar College

4:30 pm, 5 December 2022
SP 105


What have we been doing this semester?



We're not especially interested in Pyret - or Python!
If you're programming 20 years from now, it'll be in a different language, using different tools.

What have we been doing in these languages?


We've been practicing computational thinking.
"Modern computer science is the last 1 percent of the historical timeline of computational thinking. Computer scientists inherited and then perfected computational thinking from a long line of mathematicians, natural philosophers, scientists, and engineers all interested in performing large calculations and complex inferences without error."

Peter J. Denning \& Matti Tedre, Computational Thinking

## Origins of computational thinking

Before the modern computer age, there was a profession of mathematically trained experts who performed complex calculations as teams.

They were called "computers".


Bonus Bureau, Computing Division, 1924, loc.gov/pictures/item/2016838906

Teams of human computers engaged in computational thinking long before the invention of electronic computers.

Early computational thinking can be seen going back to the records of the Babylonians, who wrote down general procedures for solving mathematical problems around starting around 1800 BCE.

Long before this class, you learned these kind of computational methods.

## 27182818284590 +31415926535897

## 27182818284590 +31415926535897 <br> 7

$$
\begin{array}{r}
271828182845^{1} 90 \\
+31415926535897 \\
\hline 87
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 0487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 20487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 820487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 4820487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 44820487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 744820487
\end{array}
$$

$$
\begin{array}{r}
27182^{1} 818284590 \\
+31415926535897 \\
\hline 8744820487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 98744820487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 598744820487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 8598744820487
\end{array}
$$

$$
\begin{array}{r}
27182818284590 \\
+31415926535897 \\
\hline 58598744820487
\end{array}
$$

$$
\begin{array}{r}
2718281828411119 \\
+31415926535897 \\
\hline 58598744820487
\end{array}
$$

## Euclid's algorithm

Around 300 BCE, the Greek mathematician Euclid gave a method to find the greatest common divisor (GCD) of two numbers, which is the largest integer that divides both numbers.

## Euclid's algorithm

Euclid noticed that the GCD of two numbers divides their difference.

So, he repeatedly replaced the larger number with their difference until both were the same.

## Euclid's algorithm

Euclid noticed that the GCD of two numbers divides their difference.

So, he repeatedly replaced the larger number with their difference until both were the same.

$$
\begin{aligned}
& \operatorname{gcd}(48,18) \\
\rightarrow & \operatorname{gcd}(30,18) \\
\rightarrow & \operatorname{gcd}(12,18) \\
\rightarrow & \operatorname{gcd}(12,6) \\
\rightarrow & \operatorname{gcd}(6,6) \\
\rightarrow & 6
\end{aligned}
$$

## Sieve of Eratosthenes

This is another famous method dating back to the ancient Greeks, used to find all the prime numbers up to some limit.

## Sieve of Eratosthenes

23456789101112131415161718

We begin with a list of all the integers, from 2 to the limit.

## Sieve of Eratosthenes

23456789101112131415161718

We cross out all the multiples of 2 .

## Sieve of Eratosthenes

```
2 3 4 5 6 7 8 9 10 11 12 1314 15 16 17 18
```

Then all the multiples of 3 .

## Sieve of Eratosthenes

```
2 3 4 5 6 7 8 9 10 11 12 1314 15 16 17 18
```

Then all the multiples of 5 .

## Sieve of Eratosthenes

23456789101112131415161718

And so on, leaving you with only the primes between 2 and the limit you chose.

## Sieve of Eratosthenes

After each round of elimination, a new prime will be revealed, and the next round crosses out all its multiples.

These are computational procedures, carried out by hand!

## Programmable computers

No matter how simple and unambiguous the steps are made, human computers make mistakes - and lots of them!

So, inventors through the ages have sought to make computing machines to allow people to perform longer computations with fewer errors.

This was a slow process, taking us from...


Slide rule
c. 1620


Blaise Pascal's mechanical calculator

Precursors to the idea of a programmable computer originated well before the electronic computing age.

In the early 1700 s, French textile weavers experimented with machines that could weave complex patterns using an automatic loom.


One of the more well known is the Jacquard loom, which was controlled by long chains of punched cards.


Plan for Babbage's Analytical Engine

Babbage collaborated with a gifted mathematician, Ada Lovelace, who designed algorithms for the Analytical Engine, even though there was no machine to run them on.


Lovelace saw the Analytical Engine not as a mere calculator but as a processor of any information that could be encoded in symbols.

This insight, that computing programs can calculate not only over numbers but over symbols that can stand for anything in the world, anticipated by a hundred years a key tenet of the modern computer age.

Lovelace saw the computer as an information machine.

While Babbage's designs for a programmable computer weren't realized at the time, the age of electronics opened new possibilities.


Harvard Mark I
1944


ENIAC
c. 1945

Early computers were very difficult to program, working in languages that were closely tied to the hardware.

Grace Hopper '28 popularized the idea of a compiler for machine-independent programming languages and defined FLOW-MATIC, the first English-like data processing language in the early 1950s.

Those ideas were later folded into the popular COBOL language (1959).

Since the 1950s, many programming languages have been defined, experienced popularity, and then been supplanted by new designs.

Today, Python is the programming language most often used for work in data science and artificial intelligence.

Programming no longer involves plugging in wires or punching cards, but it's still hard!
"The programmer, like the poet, works only slightly removed from pure thought-stuff. He builds castles in the air, from air, creating by exertion of the imagination...

Fred Brooks, The Mythical Man Month, 1975
"Few media of creation are so flexible, so easy to polish and rework, so readily capable of realizing grand conceptual structures. Yet the program construct, unlike the poet's words, is real in the sense that it moves and works, producing visible outputs

Fred Brooks, The Mythical Man

Month, 1975 separate from the construct itself...
"One types the correct incantation on a keyboard, and a display screen comes to life, showing things that never were nor could be... It prints results, draws pictures, produces sounds, moves arms. The magic of myth and legend has come true in our time...
"The computer resembles the magic of legend in this respect, too. If one character, one pause, of the incantation is not strictly in proper form, the magic doesn't work. Human beings are not accustomed to being perfect, and few areas of human activity demand it. Adjusting to the requirement for perfection is, I think, the most difficult part of learning to program."

Fred Brooks, The Mythical Man<br>Month, 1975

## Computing with data

We've seen some cool data sets during this semester, but there are many, many more you can explore.

## Data Is Plural

... is a weekly newsletter of useful/curious datasets, published by Jeremy SingerVine. There have been 311 editions, dating from October 21, 2015 to November 30 2022. To receive future editions, sign up here:


## Some Nice Things People Have Said

- "a treasure trove of interesting datasets" - Julia Silge
- "delivers exactly what it promises, it's delightful" - Simon Willison
- "required reading for anyone interested in data journalism" - Julia Angwin
- "consistently fascinating" - Robin Sloan
- "the only newsletter I open immediately" - Paul Ford
- "the newsletter I forward most frequently to colleagues who like datasets" - Melody Joy Kramer
- "definitely my favorite newsletter about data and possibly my favorite newsletter period" - Melanie Walsh
- "the best data newsletter ever" - François Briatte


## Recent Editions

2022.11.30 • Pills, per-pupil spending, travelers' coronavirus variants, Indonesia earthquake intensities, and more roadkill.
2022.11.23 - Presidential pardons, radioactive waste, songs of the world, semiconductor logistics, and "every Star Trek ever." craters, and tinned fish.

Building data models

Data is situated in the environment where it was gathered.

Consider Galton's child-height data.
He gathered the data in England c. 1886.
What would happen if you tried to use it to predict heights in Poughkeepsie today? In Guatemala? In China?


When we collect data, it's like making a map:
We're constructing a model, where we choose what to represent, and how to represent it.
"...most of the data and data models we have inherited deal with structures of power, like gender and race, with a crudeness that would never pass muster in a peer-reviewed humanities publication."

Miriam Posner, "What's Next: The Radical, Unrealized Potential of
Digital Humanities", 2016


"I want us to be more ambitious, to hold ourselves to much higher standards when we are claiming to develop data-based work that depicts people's lives."

Miriam Posner, "What's Next: The Radical, Unrealized Potential of Digital Humanities", 2016

Data and privacy

help. sentiment140.com

SUM 122 AVERAGE 61 MIN 0 MAX 122
"The words and phrases we search for on Google, the times of day we are most active on Facebook, and the number of items we add to our Amazon carts are all tracked and stored as data - data that are then converted into corporate financial gain."

D'Ignazio \& Klein, Data Feminism, 2020

"As Pole's computers crawled through the data, he was able to identify about 25 products that, when analyzed together, allowed him to assign each shopper a 'pregnancy prediction' score. More important, he could also estimate her due date to within a small window, so Target could send coupons timed to very specific stages of her pregnancy.
"One Target employee I spoke to provided a hypothetical example. Take a fictional Target shopper named Jenny Ward, who is 23, lives in Atlanta and in March bought cocoa-butter lotion, a purse large enough to double as a diaper bag, zinc and magnesium supplements and a bright blue rug. There's, say, an 87 percent chance that she's pregnant and that her delivery date is sometime in late August."

Kashmir Hill, "How Target Figured Out a Teen Girl was Pregnant Before Her Father Did", Forbes, 2012

Computing with data is complex, and it's not just "technical" issues we need to concern ourselves with!

Computing with data gives us a lot of power!
We can do a lot of harm, but we can also do a lot of good!

Computer Science I
-or, where do you go from here?
Data typesNaming values
Evaluation
Conditionals (if andcases)
Function signaturesTesting functionsTabular data
Sanitizing real-worlddatadata
Visualization

Higher-order functions

Lambda expressions
Linked lists
Defining structured data

Structurally recursive data and functions

Trees (binary, n-ary)
Python lists and arrays

Iteration (for)

Side effects and mutation

Debugging using print statements

Functional vs imperative languages

Accumulators
Memory and aliasing
Hash tables (dictionaries)

Web APIs and JSON
Numpy arrays

Congratulations on making it this far!

CS courses at Vassar

Major-required courses
CMPU 101 - Computer Science I: Problem-Solving and Abstraction
CMPU 102 - Computer Science II: Data Structures and Algorithms
CMPU 145 - Foundations of Computer Science
CMPU 203 - Computer Science III: Software Design and Implementation
CMPU 224 - Computer Organization
CMPU 240 - Theory of Computation
CMPU 241 - Analysis of Algorithms
CMPU 331 or CMPU 334 - Compilers or Operating Systems
CMPU-3XX - Three 300-level units

## Correlate-required courses

CMPU 101 - Computer Science I: Problem-Solving and Abstraction CMPU 102 - Computer Science II: Data Structures and Algorithms CMPU 145 - Foundations of Computer Science
CMPU 240 or 241 - Theory of Computation or Analysis of Algorithms CMPU 2xx - Any other 200-level course
CMPU 3xx - Any 300-level course

| 300-level electives |
| :--- |
| CMPU 324 - Computer Architecture |
| CMPU 353 - Bioinformatics |
| CMPU 365 - Artificial Intelligence |
| CMPU 366 - Computational Linguistics |
| CMPU 375 - Computer Networks |
| CMPU 377 - Parallel Programming |
| CMPU 378 - Graphics |
| CMPU 379 - Computer Animation: Art, Science and Criticism |
| CMPU 395 - Advanced Special Topics |

## 200-level electives

CMPU 245 - Declarative Programming Models

## Intensives

CMPU 310 - Topics in Virtualization
CMPU 311 - Database Systems
CMPU 312 - Applications of Artificial Intelligence CMPU 314 - Projects in Digital Media Production CMPU 315 - Computer Security

## Extra-departmental

MATH 221 - Linear Algebra
*At least two CMPU-200 level courses required for every CMPU-300 level course. †Prerequisites vary depending on topic.


| CMPU 102 |
| :---: |
| Data Structures and Algorithms |

Any
CMPU
200


Four 300-level units required for major

Major-required courses
CMPU 101 - Computer Science I: Problem-Solving and Abstraction CMPU 102 - Computer Science I: Problem-Solving and Abstraction CMPU 145 - Foundations of Computer Science
CMPU 203 - Computer Science III: Software Design and Implementation CMPU 224 - Computer Organization
CMPU 240 - Theory of Computation
CMPU 241 - Analysis of Algorithms
CMADU 331 or CMMPU 331 - Compilere or Operating Systems CMPU-3XX - Three 300-level units

CMPU 101 - Computer Science I: Problem-Solving and Abstraction CMPU 102 - Computer Science II: Data Structures and Algorithms CMPU 145 - Foundations of Computer Science
CMPU 240 or 241 - Theory of Computation or Analysis of Algorithms CMPU 2xx - Any other 200-level course
CMPU 3xx - Any 300-level course

CMPU 324 - Computer Architecture
CMPU 353 - Bioinformatics
CMPU 365 - Artificial Intelligence
CMPU 366 - Computational Linguistics
CMPU 375 - Computer Networks
CMPU 377 - Parallel Programming CMPU 378 - Graphics
CMPU 379 - Computer Animation: Art, Science and Criticism CMPU 395 - Advanced Special Topics

200-level electives
CMPU 245 - Declarative Programming Models

Intensives

CMPU 310 - Topics in Virtualization CMPU 311 - Database Systems CMPU 312 - Applications of Artificial Intelligence CMPU 314 - Projects in Digital Media Production CMPU 315 - Computer Security
Extra-departmental

MATH 221 - Linear Algebra


## Try them out!

If you keep going with the CS major sequence, you work your way up to some really exciting courses, including...

CMPU 353 Bioinformatics
CMPU 377 Parallel Programming
And, you know,
probably some cool
courses I don't teach
as well!

Further reading


data-feminism.mitpress.mit.edu

That's it!

On Wednesday, we'll have a review for Exam 3 where we'll work through practice problems and answer your questions.

## go.vassar.edu/course/evals

## Acknowledgments

This lecture incorporates material from:
Peter J. Denning \& Matti Tedre, Computational Thinking
Miram Posner, UCLA
Edward Tufte, Visual Explanations
Jonathan Gordon, Vassar College

