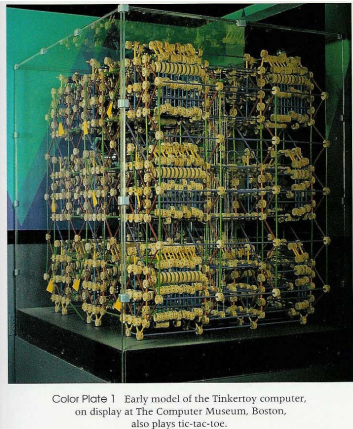


Glue, Photons, P=NP?



Color Plate 1 Early model of the Tinkertoy computer, on display at The Computer Museum, Boston, also plays tic-tac-toe.

DNA Helix Photomosaic from cover of *Nature*, 15 Feb 2001 (made by Eric Lander)

Summer CS Classes

- CS 4630 - Defense against the Dark Arts
 - Aaron Bloomfield, M-F 10:30am - 12:45
- CS 2110 - Software Development Methods
 - Mark Sherriff, M -W 10:30am - 12:45
 - You should probably take CS 2220 with Dave Evans in the fall instead, but taking CS 2110 instead will work for the CS major
- **CS 1010 - Introduction to Information Technology**
 - **Kinga Dobolyi, M-F 1pm - 3:15pm**
 - **Tell your friends!**
 - **Small class, easy way to satisfy some requirements**

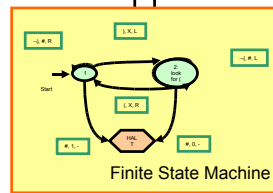
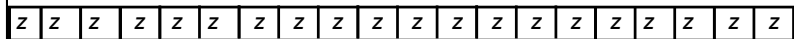
#2

One-Slide Summary

- The **lambda calculus** is a universal, fundamental model of computation. You can view it as “the essence of Scheme”. It contains terms and rules describing variables, function abstraction, and function application.
- It is possible to **encode** programming concepts, such as true, false, if, numbers, lists, etc., in the lambda calculus. Lambda calculus can simulate Turing machines.
- **Quantum computers** and **non-deterministic** Turing machines can try many options at once. They are **not more powerful** than normal Turing machines.
- The **Complexity Class P** contains tractable problems that can be solved in polynomial time. The **Complexity Class NP** contains problems for which solutions can be verified in polynomial time.
- **Does P = NP?** We don't know! \$1+ million if you know.

#3

Lambda Calculus is a Universal Computer?



- Read/Write Infinite Tape
- **Mutable Lists**
- Finite State Machine
- **Numbers**
- Processing
- **Way to make decisions (if)**
- **Way to keep going**

#4

What is 42?

42

forty-two

XLII

cuarenta y dos

Meaning of Numbers

- “42-ness” is something who’s successor is “43-ness”
- “42-ness” is something who’s predecessor is “41-ness”
- “Zero” is special. It has a successor “one-ness”, but no predecessor.

#5

#6

Meaning of Numbers

$\text{pred (succ } N) \rightarrow N$

$\text{succ (pred } N) \rightarrow N$

$\text{succ (pred (succ } N)) \rightarrow \text{succ } N$

$\text{zero? zero} \rightarrow \text{T}$

$\text{zero? (succ zero)} \rightarrow \text{F}$

#7

Is this enough?

Can we define **add** with **pred**, **succ**, **zero?** and **zero**?

$\text{add} \equiv \lambda xy. \text{if (zero? } x) y$
 $\text{(add (pred } x) (\text{succ } y))$

#8

Can we define lambda terms that behave like **zero**, **zero?**, **pred** and **succ**?

Hint: what if we had **cons**, **car** and **cdr**?

#9

Numbers are Lists...

$\text{zero?} \equiv \text{null?}$

$\text{pred} \equiv \text{cdr}$

$\text{succ} \equiv \lambda x. \text{cons F } x$

The *length* of the list corresponds to the number value.

#10

Liberal Arts Trivia: Religious Studies

- In Sunni Islam, the Five Pillars of Islam are five duties incumbent on Muslims. They include the Profession of Faith, Formal Prayers, and Giving Alms. Name the remaining two pillars.

#11

Making Pairs

(define (**make-pair** x y)
 (**lambda** (selector) (**if** selector x y)))

(define (**car-of-pair** p) (p #t))
(define (**cdr-of-pair** p) (p #f))

A **pair** is just an **if** statement that chooses between the **car** (**then**) and the **cdr** (**else**).

#12

cons and car

cons $\equiv \lambda x. \lambda y. \lambda z. zxy$

Example: $\text{cons } M \ N = (\lambda x. \lambda y. \lambda z. zxy) \ M \ N$

$\rightarrow_{\beta} (\lambda y. \lambda z. zMy) \ N$

$\rightarrow_{\beta} \lambda z. zMN$

T $\equiv \lambda xy. x$

car $\equiv \lambda p. p \ T$

Example: $\text{car } (\text{cons } M \ N) \equiv \text{car } (\lambda z. zMN) \equiv (\lambda p. p \ T)$

$(\lambda z. zMN) \rightarrow_{\beta} (\lambda z. zMN) \ T \rightarrow_{\beta} \ TMN$

$\rightarrow_{\beta} (\lambda xy. x) \ MN$

$\rightarrow_{\beta} (\lambda y. M)N$

$\rightarrow_{\beta} M$

#13

cdr too!

cons $\equiv \lambda xyz. zxy$

car $\equiv \lambda p. p \ T$

cdr $\equiv \lambda p. p \ F$

Example: $\text{cdr } (\text{cons } M \ N)$

$\text{cdr } \lambda z. zMN = (\lambda p. p \ F) \ \lambda z. zMN$

$\rightarrow_{\beta} (\lambda z. zMN) \ F$

$\rightarrow_{\beta} \ FMN$

$\rightarrow_{\beta} \ N$

#14

Null and null?

null $\equiv \lambda x. T$

null? $\equiv \lambda x. (x \ \lambda y. \lambda z. F)$

Example:

$\text{null? } \text{null} \rightarrow \lambda x. (x \ \lambda y. \lambda z. F) \ (\lambda x. T)$

$\rightarrow_{\beta} (\lambda x. T)(\lambda y. \lambda z. F)$

$\rightarrow_{\beta} T$

#15

Null and null?

null $\equiv \lambda x. T$

null? $\equiv \lambda x. (x \ \lambda y. \lambda z. F)$

Example:

$\text{null? } (\text{cons } M \ N) \rightarrow \lambda x. (x \ \lambda y. \lambda z. F) \ \lambda z. zMN$

$\rightarrow_{\beta} (\lambda z. z \ MN)(\lambda y. \lambda z. F)$

$\rightarrow_{\beta} (\lambda y. \lambda z. F) \ MN$

$\rightarrow_{\beta} F$

#16

Counting

0 $\equiv \text{null}$

1 $\equiv \text{cons } F \ 0$

2 $\equiv \text{cons } F \ 1$

3 $\equiv \text{cons } F \ 2$

...

succ $\equiv \lambda x. \text{cons } F \ x$

pred $\equiv \lambda x. \text{cdr } x$

#17

42 $= \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y$

$\lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy.$

$(\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy)$

$\lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y$

$\lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy.$

$(\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy)$

$\lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y$

$\lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy.$

$(\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy)$

$\lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y$

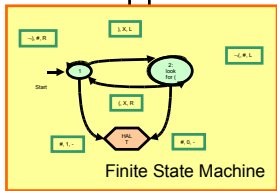
$(\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy)$

$\lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy) \ \lambda xy. y \ \lambda xy. (\lambda z. z \ xy)$

$\lambda xy. y \ \lambda x. T$

#18

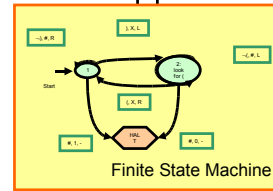
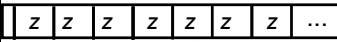
Lambda Calculus is a Universal Computer



- Read/Write Infinite Tape
 - ✓ Mutable Lists
- Finite State Machine
 - ✓ Numbers to keep track of state
- Processing
 - ✓ Way of making decisions (if)
 - ✓ Way to keep going

#19

Equivalent Computers!



Turing Machine

term = variable
 | term term
 | (term)
 | λ variable . term

← can simulate
 → can simulate

$$\lambda y. M \Rightarrow_{\alpha} \lambda v. (M [y \rightarrow v])$$

where v does not occur in M .

$$(\lambda x. M)N \Rightarrow_{\beta} M [x \rightarrow N]$$

Lambda Calculus

#20

Liberal Arts Trivia: Biology, Chemistry

- While rendered fat obtained from pigs is known as lard, rendered beef or mutton fat is known as *this*. It is used to make animal feed and soap. Historically, it was used to make candles: it provided a cheaper alternative to wax. Before switching to vegetable oil in 1990, McDonald's cooked fries in 93% *this* and 7% cottonseed oil.

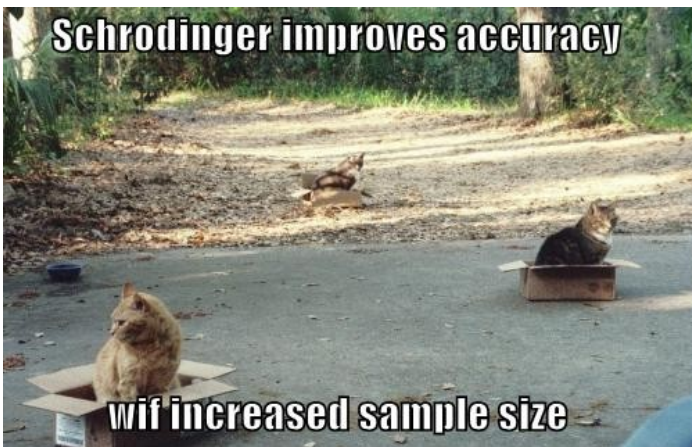
#21

Universal Computer

- Lambda Calculus can simulate a Turing Machine
 - Everything a Turing Machine can compute, Lambda Calculus can compute also
- Turing Machine can simulate Lambda Calculus (we didn't prove this)
 - Everything Lambda Calculus can compute, a Turing Machine can compute also
- **Church-Turing Thesis**: this is true for **any** other mechanical computer also

#22

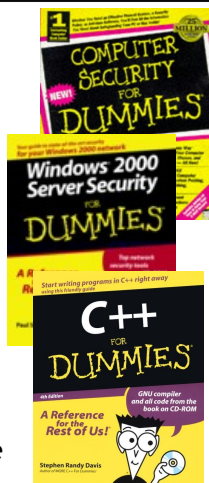
What about "non-mechanical" computers?



#23

Quantum Physics for Dummies

- Light behaves like both a **wave** and a **particle** at the same time
- A single photon is in **many states** at once
- Can't observe its state without forcing it into one state
- Schrödinger's Cat
 - Put a live cat in a box with cyanide vial that opens depending on quantum state
 - Cat is both dead and alive at the same time until you open the box



#24

Quantum Computing

- Feynman, 1982
- Quantum particles are in all possible states
- Can try lots of possible computations at once with the same particles
- In theory, can test all possible factorizations/keys/paths/etc. and get the right one!
- In practice, very hard to keep states entangled: once disturbed, must be in just one possible state

#25

Qubit

- Regular bit: either a 0 or a 1
- **Quantum bit:** 0, 1 or **in between**
 - p% probability it is a 1
- A single qubit is in 2 possible states at once
- If you have 7 bits, you can represent any one of 2^7 different states
- If you have 7 qubits, you have 2^7 different states (at once!)



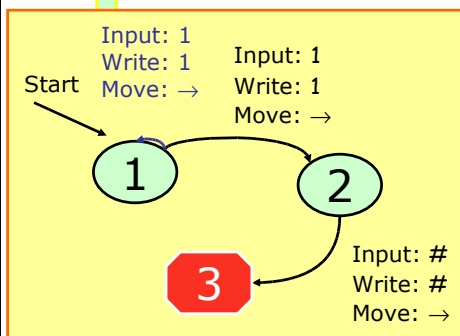
#26

Quantum Computers Today

- Several quantum algorithms
 - Shor's algorithm: factoring using a quantum computer
- Actual quantum computers
 - 5-qubit computer built by IBM (2001)
 - Implemented Shor's algorithm to factor:
 - "World's most complex quantum computation" $15 (= 5 * 3)$
 - D-Wave 16-qubit quantum computer (2007)
 - Solves Sudoku puzzles
- To exceed practical normal computing need > 50 qubits
 - Adding another qubit is more than twice as hard

#27

Nondeterministic Computing



There can be multiple transitions on the same input. Nondeterministic TM takes *all* of them at once. Each gets its own independent copy of the tape. If *any* path finds halting state, that is the result.

#28

Two Ways of Thinking about Nondeterministic Computing

- **Omniscient** (all-knowing): machine always guesses right (the right guess is the one that eventually leads to a halting state)
- **Omnipotent** (all-powerful): machine can split in two every step, all resulting machines execute on each step, if one of the machines halts its tape is the output

#29

Computability

Is a nondeterministic TM more powerful than a deterministic TM?

#30

Liberal Arts Trivia: Geography

- This second-longest river in the United States flows from Lake Itasca in Minnesota to the Gulf of Mexico. Forty percent of North America's ducks, geese, swan and wading bird species use it as a migration corridor. It serves as the shared border for ten states, contains over 29 locks and dams, and generates more than a billion dollars a year in revenue from recreational uses, including over 600 water-oriented sites.

#31

Computability

Is a nondeterministic TM more powerful than a deterministic TM?

No! We can simulate a nondeterministic TM with a regular TM.

#32

Speed

Is a nondeterministic TM **faster** than a deterministic TM?

#33

Speed

Is a nondeterministic TM **faster** than a deterministic TM?

Unknown! This is the most famous open problem in CS.

#34

Pegboard Problem



#35

Pegboard Problem

- Input: a configuration of n pegs on a cracker barrel style pegboard
- Output: if there is a sequence of jumps that leaves a single peg, output that sequence of jumps. Otherwise, output false.

How hard is the Pegboard Problem?

#36

Problems and Procedures

- To know a $O(f)$ **bound for a problem**, we need to find a $\Theta(f)$ *procedure* that solves it
 - The sorting **problem** is $O(n \log n)$ since we know a **procedure** that solves it in $\Theta(n \log n)$
- To know a $\Omega(f)$ **bound for a problem**, we need to prove that there is no procedure faster than $\Theta(f)$ that solves it
 - We proved sorting is $\Omega(n \log n)$ by reasoning about the number of decisions needed

#37

How much work is the Pegboard Problem?

- Upper bound: (O)
 $O(n!)$
Try all possible permutations
- Lower bound: (Ω)
 $\Omega(n)$
Must at least look at every peg
- Tight bound: (Θ)
 - What do you think?

#38

How much work is the Pegboard Problem?

- Upper bound: (O)
 $O(n!)$
Try all possible permutations
- Lower bound: (Ω)
 $\Omega(n)$
Must at least look at every peg
- Tight bound: (Θ)

No one knows!

#39

Complexity Class P “Tractable”

Class P: problems that can be solved in a **polynomial** ($O(n^k)$ for some constant k) number of steps by a **deterministic** TM.

Easy problems like sorting, making a photomosaic using duplicate tiles, and simulating the universe are all in P.

#40

Liberal Arts Trivia: Astronomy

- This space telescope was launched into orbit by the Space Shuttle Discovery in 1990. After having its incorrectly-ground main mirror replaced in 1993, it has helped to make breakthroughs in astrophysics. Notable are its Ultra Deep Field image, which details the universe's most distant objects, and its help in determining the ultimate expansion of the universe (the eponymous constant).

#41

Complexity Class NP

Class NP: Problems that can be solved in a **polynomial** number of steps by a **nondeterministic** TM.

Omnipotent: If we could try all possible solutions at once, we could identify the solution in polynomial time.

Omniscient: If we had a magic guess-correctly procedure that makes every decision correctly, we could devise a procedure that solves the problem in polynomial time.

#42

NP Problems

- Can be solved by just **trying all possible answers** until we find one that is right
- Easy to quickly check if an answer is right
 - Checking an answer is in P
- The pegboard problem is in NP
 - We can easily try $\sim n!$ different answers
 - We can check if a guess is correct in $O(n)$ (check all n jumps are legal)

#43

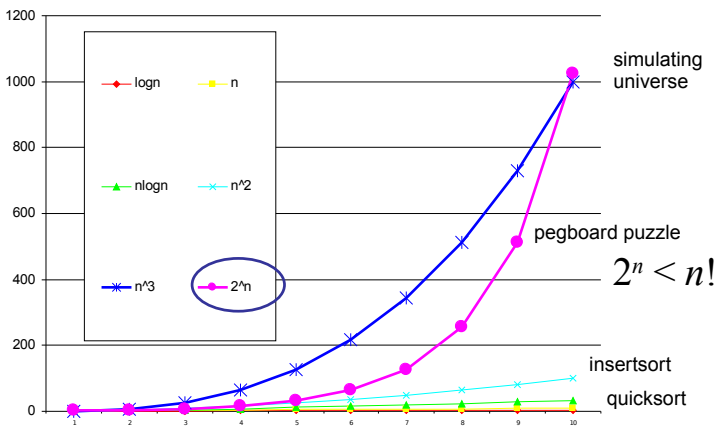
Is the Pegboard Problem in P?

No one knows!

We can't find a $O(n^k)$ solution.
We also can't prove one doesn't exist.

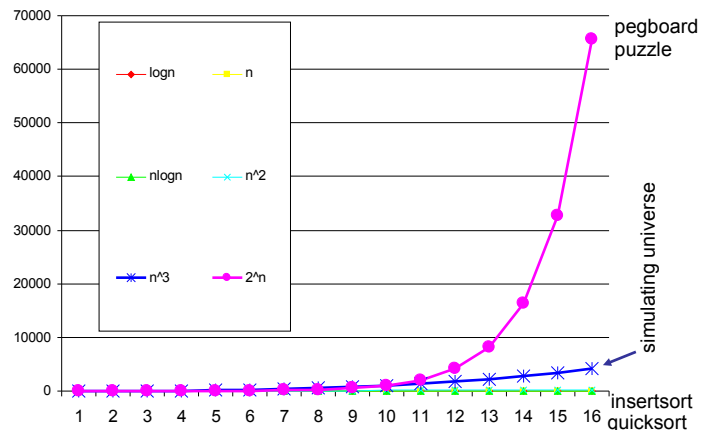
#44

Orders of Growth



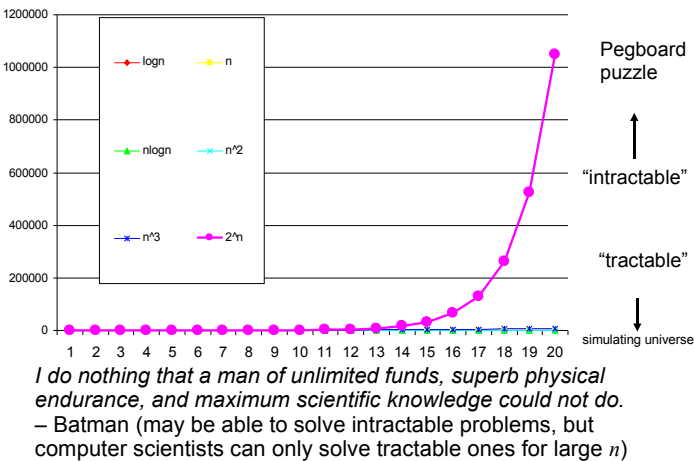
#45

Orders of Growth



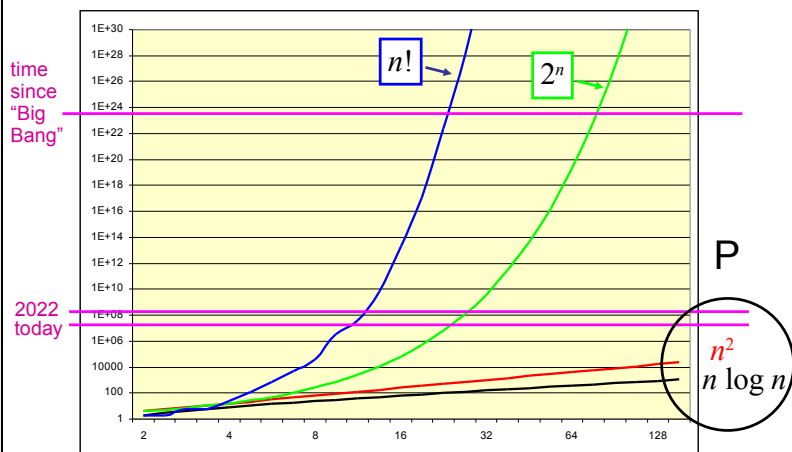
#46

Orders of Growth



#47

Intractable Problems log-log scale



#48

Moore's Law Doesn't Help

- If the fastest procedure to solve a problem is $\Theta(2^n)$ or worse, Moore's Law doesn't help much.
- Every **doubling** in computing power only increases the solvable problem size by 1.

#49

Complexity Classes

Class P: problems that can be solved in polynomial time by deterministic TM

Easy problems like simulating the universe are all in P.

Class NP: problems that can be solved in polynomial time by a nondeterministic TM.

Includes all problems in P and some problems possibly outside P like the Pegboard puzzle.

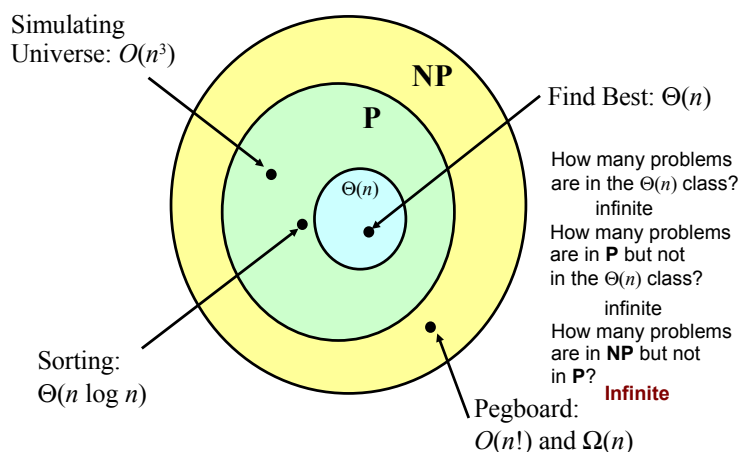
#50

Liberal Arts Trivia: Psychology

- This Swiss philosopher and natural scientist is known as "the great pioneer of the constructivist theory of knowing." His four stages of development include infancy, preschool, childhood and adolescence. Each stage corresponds to the child's understanding of reality (e.g., conservation, abstract reasoning) during that period.

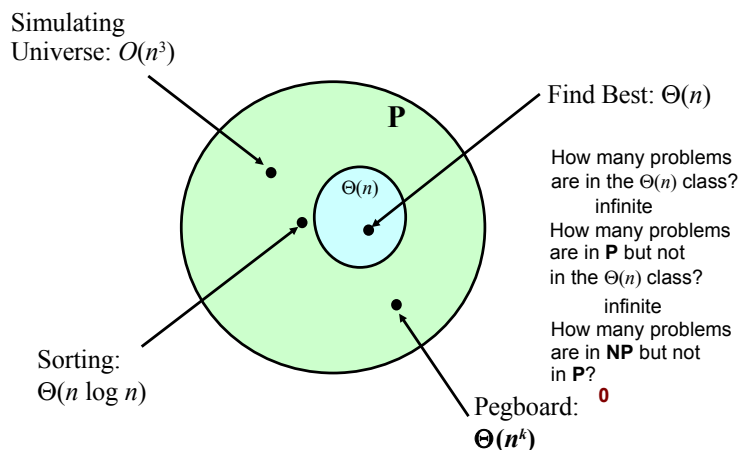
#51

Problem Classes if $P \neq NP$:



#52

Problem Classes if $P = NP$:



#53

P = NP?

- Is P different from NP: is there a problem in NP that is not also in P
 - If there is one, there are infinitely many
- Is the "hardest" problem in NP also in P
 - If it is, then every problem in NP is also in P
- **The most famous unsolved problem in computer science and math**
 - Listed first on Millennium Prize Problems
 - \$1M + an automatic A+ in this course
 - (and probably all CS courses at UVA ...)

#54

NP-Complete

- **NP-Complete:** is the class of problems that are the *hardest* problems in NP
- Cook and Levin proved that 3SAT was NP-Complete (1971)
 - If 3SAT can be transformed into a different problem in polynomial time, than that problem must also be NP-complete.
 - Pegboard \Leftrightarrow 3SAT
- **Either all NP-complete problems are tractable (in P) or none of them are!**

#55

NP-Complete Problems

- Easy way to solve by trying all possible guesses
- If given the “yes” answer, quick (in P) way to check if it is right
- If given the “no” answer, no quick way to check if it is right
 - No solution (can't tell there isn't one)
 - No way (can't tell there isn't one)

This part is hard to prove: requires showing you could use a solution to the problem to solve a known NP-Complete problem.

#56

Most Important Science/Technology Races

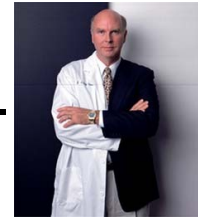
- 1930-40s: **Decryption** Nazis vs. British
 Winner: British
 Reason: Bletchley Park had computers (and Alan Turing), Nazi's didn't
- 1940s: **Atomic Bomb** Nazis vs. US
 Winner: US
 Reason: Heisenberg miscalculated, US had better physicists, computers, resources
- 1960s: **Moon Landing** Soviet Union vs. US
 Winner: US
 Reason: Many, better computing was a big one
- 1990s-2001: **Sequencing Human Genome**

#57

Human Genome Race



Francis Collins
 (Director of public National Center for Human Genome Research)
 (Picture from UVa Graduation 2001)



Craig Venter
 (President of Celera Genomics)

VS.

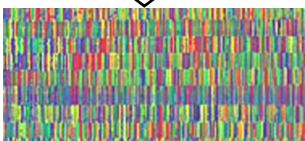
- UVa CLAS 1970
- Yale PhD
- Tenured Professor at U. Michigan
- San Mateo College
- Almost court-martialed
- Denied tenure at SUNY Buffalo

#58

Reading the Genome



Whitehead Institute, MIT



Gene Reading Machines

- One read: about 700 base pairs
- But...don't know where they are on the chromosome

Read 3 **TACCCGTGATCCA**

Read 2 **TCCAGAATAA**

Read 1 **ACCAGAATACC**

AGGCATACCAGAATACCCGTGATCCAGAATAAGC

Actual Genome

#59

#60

Liberal Arts Trivia: Greek Myths

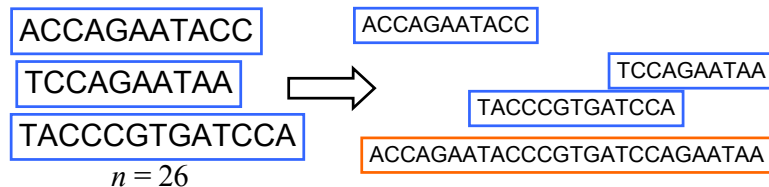
- The Sphinx is said to have guarded the entrance to the city of Thebes and to have asked a riddle of would-be travelers. The Sphinx, originally from Ethiopia, is said to have been sent by Hera or Ares. Oedipus solved her riddle, and is thus seen as a threshold figure, helping to transition between the old religious practices and the new Olympian gods.
- State the Riddle of the Sphinx and its answer.

#61

Common Superstring

Input: A set of n substrings and a maximum length k .

Output: A string that contains all the substrings with total length $\leq k$, or no if no such string exists.

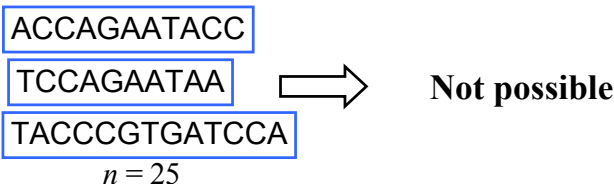


#62

Common Superstring

Input: A set of n substrings and a maximum length k .

Output: A string that contains all the substrings with total length $\leq k$, or no if no such string exists.



#63

Common Superstring

- In NP:
 - Easy to verify a “yes” solution: just check the letters match up, and count the superstring length
- In NP-Complete:
 - Similar to Pegboard Puzzle!
 - Could transform Common Superstring problem instance into Pegboard Puzzle instance!

#64

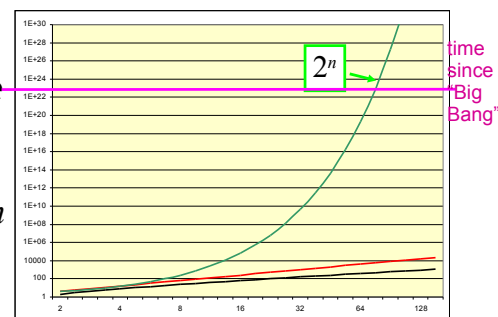
Human Genome

- 3 Billion base pairs
- 600-700 bases per read
- ~8X coverage required
 $> (1 / (* 8 3000000000)) 650$
 $36923076 \text{ } 12/13$
- So, $n \approx 37$ Million sequence fragments
- Celera used 27.2 Million reads (but could get more than 700 bases per read)

#65

Give up?

No way to solve an NP-Complete problem (best known solutions being $O(2^n)$ for $n \approx 20$ Million)



#66

Approaches

- Human Genome Project (Collins)
 - **Change problem:** Start by producing a genome map (using biology, chemistry, etc) to have a framework for knowing where the fragments should go
- Celera Solution (Venter)
 - **Approximate:** we can't guarantee finding the shortest possible, but we can develop clever algorithms that get close most of the time in $O(n \log n)$

#67

Result: Draw?



President Clinton announces Human Genome Sequence essentially complete, June 26, 2000

#68

CS 1120

- Language: Formal Systems, Rules of Eval
- Recursive Definitions
- Programming with Lists
- Programming with Mutation and Objects
- Interpreters, Lazy Eval, Type Checking
- Programming for the Internet
- Measuring Complexity
- Computability
- Models of Computation

#69

Homework

- PS 9 Presentation Requests due Today
- This Wednesday, Here, 5pm-...
 - Presentations
 - Extra Credit on PS9 for attending
- Wednesday May 12
 - Final Project Reports Due
- No final exam, as per class vote!

#70

Liberal Arts Trivia: Bias

- Weimer recommends that you take classes on **philosophy** until you've covered epistemology, free will, logic, the philosophy of science, and "what it is like to be a bat". Take **cognitive psychology** classes until you've covered perception and the Flynn effect. Take **speech** or **rhetoric** classes until you've covered persuasion. Take **anthropology** as well as **gender studies** classes until you've covered Mead and Freeman and you have a better feel for which behaviors are socially constructed and which may be essential. Take classes in **statistics** until you can avoid being fooled. Take classes in **religion** or **ethics** until you've covered the relationship between unhappiness and unrealized desires. Take classes in **physics** until you can explain how a microphone, radio and speaker all work. Take classes on **government** until you have an opinion about the feasibility of legislating morality. Take classes on **history** until you are not condemned to repeat the mistakes of the past.

#71