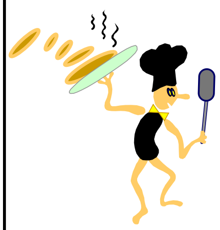
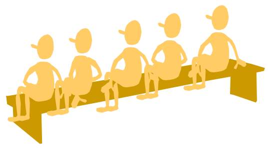
Great Theoretical Ideas In Computer Science

Victor Adamchik CS 15-251 Fall 2013

Lecture 1 Aug 27, 2013 Carnegie Mellon University

Cooking for Computer Scientists





Course Staff

Victor Adamchik

Ariel Procaccia

TAs

A: Nathan Dobson
B: Jason Harding
C: John Retterer-Moore
D: Tim Broman

Office hours start Wednesday
Timings are posted.

Web Sites

http://www.andrew.cmu.edu/course/15-251

Calendar, Slides, Notes, Homeworks, Course Policy, Grades, …

https://piazza.com/cmu/fall2013/15251

Questions, Comments, Announcements, …

Textbook

There is no textbook.

Slides will be posted on the website.

Some supplementary notes will also be posted.

Grading

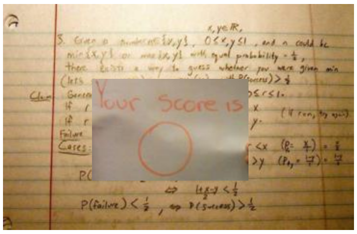
35%	Homework	(11, lowest one dropped)
10%	Quizzes	(12, lowest *two* dropped; no make-ups)
30%	Tests	(2 midterms)
25%	Final	

Homework

Homeworks roughly every week
(see currently planned schedule on calendar).
Out/due at 11:59pm on respective date.

Must be typeset
submit pdf via "handin",
returned via "handback"
(read FAQ on website).

Homework



Homework Late Policy

You have 8 late days (total),
but you cannot use more
than 2 late days
per homework.

Collaboration

You may work in a group of ≤ 4 people.

You *must* report who you worked with.

You must think about *each of the problems* by yourself for ≥ 30 minutes before discussing them with others.

You must write up *all* solutions by yourself.

Cheating

You MAY NOT

Share written work.

Get help from anyone besides your collaborators, staff.

Refer to solutions/materials from earlier versions of 251 or the web

Quizzes

Every Tuesday, beginning of class

The quiz will be DUE AT 3:10 pm.
Therefore, do NOT be late to class.

Tested on material from the previous 2-3 lectures.

These are designed to be easy, assuming you are keeping up with the lectures.

Midterm tests

Designed to be doable in 1 hour.

You will have 1.5 hours.

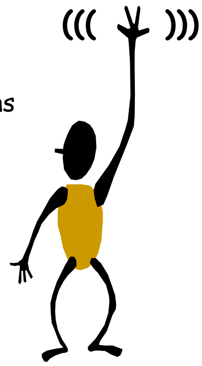
"Semi-cumulative."

Given in lectures.

Oct 1, Nov. 5

Mark these dates on your calendar *now*!

Feel free to ask questions

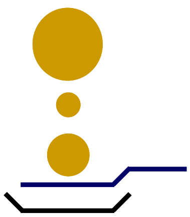


The chef at our place is sloppy: when he prepares pancakes, they come out all different sizes

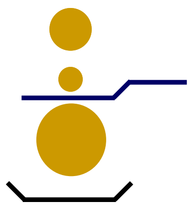
When the waiter delivers them to a customer, he rearranges them (so that smallest is on top, and so on, down to the largest at the bottom)

He does this by grabbing several from the top and flipping them over, to perform a prefix reversal, repeating this as many times as necessary

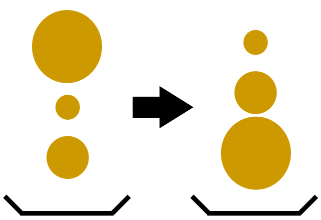
How do we sort this stack?
How many flips do we need?



How do we sort this stack?
How many flips do we need?

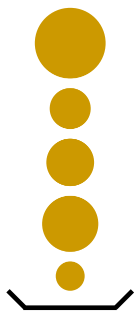


How do we sort this stack?
How many flips do we need?

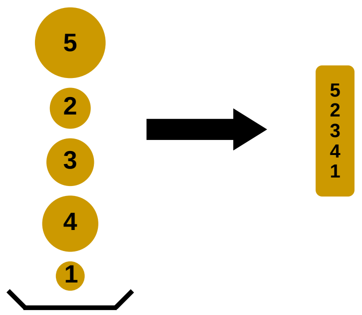


2 flips sufficient
2 flips necessary

How do we sort this stack?
How many flips do we need?

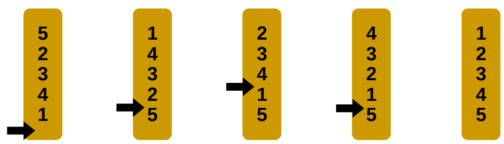


Developing Notation: Turning pancakes into numbers



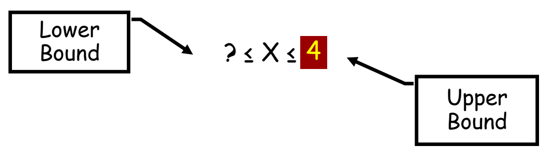
How do we sort this stack? How many flips do we need?

Sorting by prefix reversal: 4 flips are sufficient



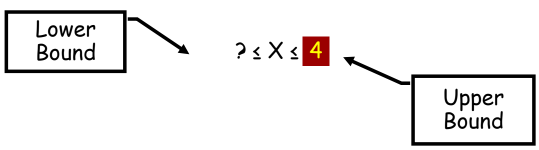
Best way to sort this stack?

Let X be the smallest number of flips that can sort this specific stack.

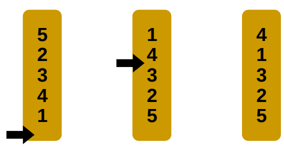


Is 4 a lower bound? What would it take to show that?

A convincing argument that every way of sorting the stack uses at least 4 flips.



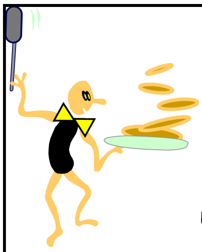
Four Flips Are Necessary



If we could do it in three flips:

First flip has to put 5 on bottom, because…

Second flip has to bring 4 to the top, because…



Best way to sort?

5
2
3
4
1

Let X be the smallest number of flips that can sort this specific stack.

Lower Bound → $4 \le X \le 4$ ← Upper Bound

$X = 4$

Pancake Number

5 2 3 4 1	5 4 3 2 1	1 2 3 4 5	5 4 1 2 3	? ? ? ? ?
4	1	0	2	P_5

5th Pancake Number

The cook chooses the "worst possible" stack of five pancakes (from 5! = 120), and the waiter sorts the stack using the "fewest possible" flips.

P_5 = MAX over all 5-stacks S of MIN # of flips to sort S

5th Pancake Number

Lower Bound → $4 \le P_5 \le ?$ ← Upper Bound

Fact: $P_5 = 5$

To show $P_5 \ge 5$?

1. Show a specific 5-stack.
2. Argue that every way of sorting this stack uses at least 5 flips.

To show $P_5 \le 5$?

Give a way of sorting every 5-stack using at most 5 flips.

$P_3 = 3$

To show $P_3 \ge 3$?

1. Show a specific 3-stack.
2. Argue that every way of sorting this stack uses at least 3 flips.

1
3
2

To show $P_3 \le 3$?

Give a way of sorting every 3-stack using at most 3 flips.

- Biggest one to bottom using ≤ 2 flips.
- Smallest one to top using ≤ 1 flip.

nth Pancake Number

Lower Bound → $? \le P_n \le ?$ ← Upper Bound

"What is the best upper bound and lower bound I can prove?"

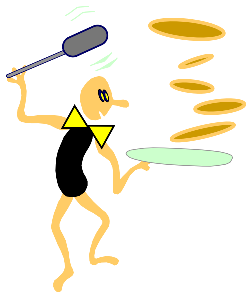
n^{th} Pancake Number

$\leq P_n \leq ?$

Upper Bound

Let's start by thinking about an <u>upper</u> bound.

Bring-to-top Method



Bring the biggest to the top
Place it on the bottom

Bring the next largest to the top. Place it on the bottom

And so on...

Bring-to-top Method For n Pancakes

If n=2, at most one flip and we are done!

Otherwise, flip pancake n to the top and then flip it to the bottom

Now use:

Bring-To-Top Method for n-1 Pancakes

If T(n) is the number of flips then

$T(n) = 2 + T(n-1)$

$T(2) = 1$

Recurrence Equation
$T(n) = 2 + T(n-1)$

Solving by iteration

$T(n) = 2 + T(n-1)$

$= 2 + 2 + T(n-2)$

$= 2 + 2 + 2 + T(n-3)$

… … … … after k steps

$= 2 * k + T(n-k)$

For $k = n-2$

$T(n) = 2(n-2) + T(2) = 2n - 4 + 1 = 2n-3$

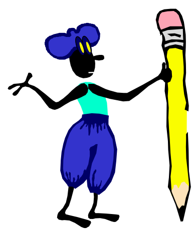
Bring-to-top Method For n Pancakes

$? \leq P_n \leq T(n) = 2n-3$

Observe,

$P_5 \leq T(5) = 7$

$? \leq P_n \leq 2n-3$



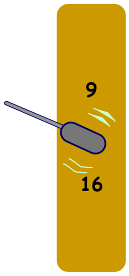
Let's think about a lower bound for P_n

Breaking Apart Argument

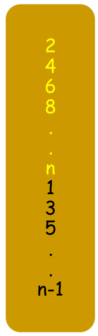
Suppose a stack S has a pair of adjacent pancakes that will not be adjacent in the sorted stack

Any sequence of flips that sorts stack S must have one flip that inserts the spatula between that pair and breaks them apart

Each flip can achieve at most 1 "break-apart".



$n \le P_n$



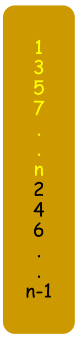
Suppose n is even

S contains n adjacent pairs that will need to be broken apart during any sequence that sorts it

Detail: This construction works when n>=2.

"Adjacent pair" includes bottom pancake and the plate.

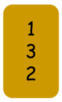
$n \le P_n$



Suppose n is odd

S contains n pairs that will need to be broken apart during any sequence that sorts it

Detail: This construction works when n>=3.



$n \le P_n \le 2n - 3$ for n >= 3

Upper and lower bounds are within a factor of 2.

The Known Pancake Numbers

n	P_n
5	5
6	7
7	8
8	9
9	10
10	11
11	13
12	14
13	15
14	16
15	17
16	18
17	19
18	20
19	22

P_{20} is unknown

It is either 23 or 24, we don't know which.

$20 \cdot 19 \cdot 18 \cdots 2 \cdot 1 = 20!$ possible 20-stacks

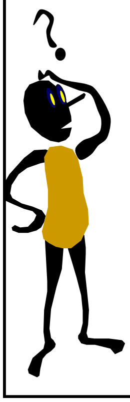
$20! = 2.43 \times 10^{18}$

(2.43 exa-pancakes)

Brute-force analysis would take forever!

Sorting by prefix reversal (with a min number of flips) is NP-hard

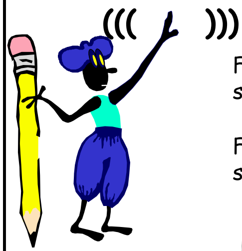
Any Stack to Any Stack



Suppose we didn't want to sort the stack of pancakes:
instead we wanted to go from a "source" stack (4 3 5 1 2)
to some "target" stack (5 1 4 3 2).

How should we do this?

Any Stack to Any Stack



From any stack to a sorted stack in ≤ P_n

From a sorted stack to any stack in ≤ P_n (by reversing)

Hence, from ANY stack to ANY stack in ≤ 2 P_n

Can you find a faster way than $2P_n$ flips to go from ANY to ANY?

Any Stack S to Any Stack T in ≤ P_n

S: 4,3,5,1,2 T: 5,1,4,3,2

3,4,1,2,5 ← "new S" 1,2,3,4,5

Rename the pancakes in T to be 1,2,3,…,n

Rewrite S using the new naming scheme

P_n flips can sort "new S".

The same sequence of flips also brings S to T.

Is This Really Computer Science?

Is This Really Computer Science?

Posed in *Amer. Math. Monthly* 82(1), 1975, by "Harry Dweighter" (haha).

aka Jacob Goodman, a computational geometer.

Is This Really Computer Science?

Discrete Mathematics 27(1), 1979

$(17/16)\, n \leq P_n \leq (5/3)\, n + 5/3$

William H. Gates (Microsoft)
Christos Papadimitriou (Berkeley)

"On the Diameter of the Pancake Network"
Journal of Algorithms 25(1), 1997

$(15/14)\, n \leq P_n \leq (5/3)\, n + 5/3$

by Hossain Heydari and Hal Sudborough

"An (18/11)n Upper Bound For Sorting By Prefix Reversals"
Theoretical Computer Science 410(36), 2009

$(15/14)\, n \leq P_n \leq (18/11)\, n$

Upper and lower bounds are within a factor of 1.5

by B. Chitturi, W. Fahle, Z. Meng, L. Morales, C.O. Shields, I.H. Sudborough, W. Voit
@ UT Dallas

Burnt Pancakes

There are other variants of the problem: where the pancakes are burnt on one side, and the goal is not only to sort them but to also place them with the burnt side down.

The problem was introduced in the Gates & Papadimitriou paper.

$(3/2)\, n-1 \leq BP_n \leq 2\, n + 3$

Burnt Pancakes

$(3/2)\, n \leq BP_n \leq 2\, n - 2$

"On The Problem Of Sorting Burnt Pancakes"
Discrete Applied Math. 61(2), 1995

by David S. Cohen and Manuel Blum (cmu)

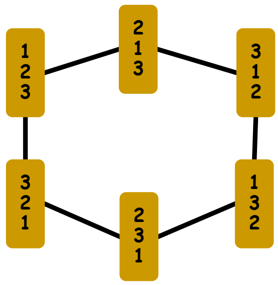
Application:

The Pancake Network

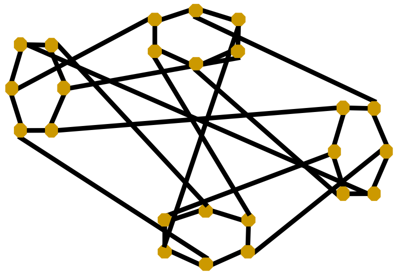
The Pancake Network

Nodes are named after the n! different stacks of n pancakes

Put a link between two nodes if you can go between them with one flip

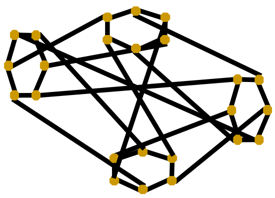


Network for n = 4



Pancake Network: Message Routing Delay

What is the maximum distance (a diameter) between two nodes in the pancake network?



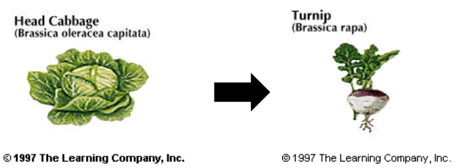
P_n

Pancake Network: Reliability

If up to n–2 nodes get hit by lightning, the network remains connected, even though each node is connected to only n–1 others

The Pancake Network is optimally reliable for its number of nodes and links

Computational Biology



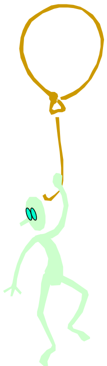
Studies convincingly proved that genome rearrangements is a common mode of molecular evolution. Genome rearrangements involves finding a shortest series of prefix reversals (and other transpositions) to transform one genome into another.

One "Simple" Problem

A host of problems and applications at the frontiers of science

High Level Point

Computer science is no more about computers than astronomy is about telescopes – E. Dijkstra



Computer Science is not merely about computers and programming, it is about mathematically modeling our world, and about finding better and better ways to solve problems

Study Bee

Definitions of:
- Pancake number
- Lower bound
- Upper bound

Proof of:
- Bring-To-Top
- Breaking-Apart
- ANY to ANY in $\leq P_n$