Reading Assignment

Read Chapter 1

Mirror…mirror on the wall, who's the biggest trader of'em all?

Suppose you are hired by Warren Buffet to a well-paid job at Berkshire Hathaway.

Suppose on your first day Mr. Buffet shows up at your cubicle and tells you: "write a program to find out the fifth largest stock trade that occurred in the NY Stock Exchange today. If you get me the info tomorrow morning, I will give you one stock in my company as a bonus!"

Warren is a tough guy…If you don't deliver, you can kiss your job goodbye!!

The Problem

How many trades occur in NYSE each day?

    several millions

Suppose 10 million trades occurred on the day Buffett came to you.

All you have to do is to find the name of the person who did the fifth largest trade in terms of $ amount.

The Problem…contd

Suppose you can download the trading data to your computer, put it in a two-dimensional (names and trade values) array of 10 million cells.

How will you solve the problem?

Will you be done in 24 hours?

The Problem…contd

Suppose you decide to sort the array using Bubble sort.

How many steps will your algorithm have to execute?

10^7 array writes

Bubblesort is a quadratic ($O(n^2)$) algorithm. What does this mean?

It will carry out approximately $n^2 - 10^{14}$ – operations to sort the array

Then one print operation to print out the name in the fifth cell of the array

The Problem…contd

Say Buffet gave you a really fast machine running at 4 GHz, i.e., $4 * 10^9$ clock cycles per second.

A machine typically requires about 200 clock cycles to execute one computation step.

So your computer can execute $2 * 10^7$ steps in each second.

How long will it take to do just the sorting (execute 10^{14} steps)?

Approximately $5 * 10^6$ seconds.

The Problem…contd

How long is $5 * 10^6$ seconds?

There are $8.6 * 10^4$ seconds in a day

58 days...so you'll have to tell Buffet the next morning: "come back after two months!"

You better head over to the unemployment benefits office!

The Problem…contd

What if you used a different sorting algorithm – Merge Sort which has O(nlogn) complexity?

How many steps are needed to sort 10 million items using Merge Sort?

$24 * 10^7$ steps

How long will it take to execute $24 * 10^7$ steps?

About 12 seconds!

Selection Problem

This example is actually an instance of a general problem called the Selection Problem:

Given n items and a way to compare any two items (to decide if one is >, = or < the other), find the k-th largest or smallest item, $1 \leq k \leq n$

How fast do you think you can solve this problem?

Selection Problem

Fastest known algorithm (median-of-median algorithm) is O(n).

10^7 steps

How long will it take to execute 10^7 steps?

Recall that your computer can execute $2 * 10^7$ steps in each second.

half-a-second!

So, your choices are…

Sort data using Bubble Sort and ask Mr. Buffett to come back after two months…lose your job

Sort data using Merge Sort and find the answer in 12 seconds

Use the fastest known algorithm to solve the problem in half of a second

This is why efficiency is so important!