LECTURE 3: DIVIDE & CONQUER

ADVANCED ALGORITHMS
ANNOUNCEMENTS

▸ Homework 0 grading

▸ Homework 1 posted; due on Monday Sep 10, 11:59 PM

▸ Start early! (Post comments/questions on canvas…)
  ▸ Unless immediate, argue why algorithm is correct
  ▸ Three steps – algorithm description, proof of correctness, complexity analysis
LAST CLASS

▸ Basics of data structures
  ▶ What data? What operations supported? Time/space complexity
  ▶ Prefix trees (aka tries)

▸ Connectivity in graphs via depth-first-search (DFS)
  ▶ Proved using induction on the length of the path
TODAY’S CLASS

- Divide and conquer
- Analysis: recurrences and how to solve them
DIVIDE AND CONQUER

Divide and rule (or divide and conquer, from Latin divide et impera) in politics and sociology is gaining and maintaining power by breaking up larger concentrations of power into pieces that individually have less power than the one implementing the strategy. The concept refers to a strategy that breaks up existing power structures, and especially prevents smaller power groups from linking up, causing rivalries and fomenting discord among the people.

Traiano Boccalini cites "divide et impera" in La bilancia politica, 1,136 and 2,225 as a common principle in politics. The use of this technique is meant to empower the sovereign to control subjects, populations, or factions of different interests, who collectively might be able to oppose his rule. Machiavelli identifies a similar application to military strategy, advising in Book VI of The Art of War[^1] (Dell'arte della guerra),[^2] that a Captain should endeavor with every art to divide the forces of the enemy, either by making him suspicious of his men in whom he trusted, or by giving him cause that he has to separate his forces, and, because of this, become weaker.

The maxim divide et impera has been attributed to Philip II of Macedon, and together with the maxim divide ut regnes was utilised by the Roman ruler Caesar and the French emperor Napoleon.

The strategy, but not the phrase, applies in many ancient cases: the example of Gabinius exists, parting the Jewish nation into five conventions, reported by Flavius Josephus in Book I, 169-170 of The Wars of the Jews (De bello Judaico).[^3] Strabo also reports in Geography, 8.7.3[^4] that the Achaean League was gradually dissolved under the Roman possession of the whole of Macedonia, owing to them not dealing with the several states in the same way, but wishing to preserve some and to destroy others.
DIVIDE AND CONQUER

PROBLEM INSTANCE

- Divide step
- Solve sub-problems
- Combine/"conquer"

FULL SOLUTION
WHEN AND WHY?

- Problem “cleanly” divides into not-too-many sub-problems
- Solutions easily (efficiently) combined
- Leads to simple, efficient solutions
- Analysis easier – correctness by induction, complexity by recurrences

Sometimes difficult to see how D&C can be used...
**Basic question:** given unsorted array $A[0], \ldots, A[n-1]$, re-order elements in increasing order (produce new array)

- Divide step?
- Conquer step

| 3 | 5 | 4 | 2 | 9 | 8 | 5 | 6 | 10 |

- Find smallest element.

Sort recursively
MERGING SORTED LISTS

- Shift the array $B^{(j)}$ to the left & repeated

Time for merging:

\[ \frac{n}{k} \text{ steps, each taking time } \frac{1}{\log k}. \]

Observation: Smallest among $B^{(1)}, B^{(2)}, \ldots, B^{(k)}$ is the smallest elt in the entire array.
OVERALL ALGORITHM — MERGE SORT

$T(n)$:

- Divide $A[]$ into $k$ pieces each of size $[n/k]$
- Sort each one recursively
- Use merge procedure to create new sorted list

What is the overall time complexity?

Let $T(m)$ denote running time of procedure on an instance of size $m$. 
BASIC "RECURRENCE"

\[ T(n) = k \cdot T(n/k) + kn + n \]

Setting \( k = 2 \)

\[ T(n) = 2T(n/2) + 3n \]

**Goal:** come up with a closed form solution

Minor: is \( n \) divisible by \( k \)?

\[ T(n) \leq n^{\log_2 2} = n \]
ASIDE: HOW TO SOLVE RECURRENCES?

- No silver bullet!
- Guess, prove by induction – sometimes proof gives “feedback”
- “Plug and chug” [Lehman, Leighton]
- Recurrence tree

$\Rightarrow$ keep plugging back into equation, see a pattern, prove by induction
PLUG AND CHUG

\[ T(1) = 1 \]

\[ T(n) = 2T(n/2) + 3n = \frac{3n}{2} + 2T\left(\frac{n}{4}\right) \]

\[ T(n) = 3n + \frac{n}{2} \left[ 3 \cdot \frac{n}{2} + 2T\left(\frac{n}{8}\right) \right] = 3n + 3n + 4T\left(\frac{n}{4}\right) \]

\[ = 3n + 3n + 4 \left[ 3 \cdot \frac{n}{4} + 2T\left(\frac{n}{8}\right) \right] = 3n + 3n + 3n + 8T\left(\frac{n}{8}\right) \]

\[ \vdots \]

\[ r \text{ steps} \]

\[ = r \cdot 3n + 2^r T\left(\frac{n}{2^r}\right) \rightarrow 3n \log_2 n + n - 1 \]

Base case: \( \frac{n}{2^r} = 1 \) \( \Rightarrow \) \( r = \log_2 n \)
Recurrence Tree

\[ T(n) = 2T\left(\frac{n}{2}\right) + 3n \]

2 terms

\[ T\left(\frac{n}{2}\right) \]

4 terms

\[ T\left(\frac{n}{4}\right) \]

3 terms

\[ T(1) \]

Total extra work = \( 3n \)

\[ 3n \cdot \log_2 n + n \cdot T(1) \]
EXAMPLE 2: MULTIPLYING INTEGERS

MULTIPLICATION
Each intermediate sum is shifted left!

\[
\begin{array}{c}
1932 \\
\times 2142 \\
\hline
3864 \\
7728 \\
1932 \\
3864 \\
\hline
4138344
\end{array}
\]

\[
\begin{array}{c}
a_1 a_2 \ldots a_n \\
\times b_1 b_2 \ldots b_n \\
\hline
\end{array}
\]

Run time of elementary school algo: time to create “table” + time to add

Question: given two \( n \) digit numbers, how efficiently can we multiply them?
DIVIDE AND CONQUER

\[ 500 = 5 \times 10^2 \]

\[
A = \underbrace{X000000 + Y}^{n/2} = (X \cdot 10^{n/2} + Y)
\]

\[
B = (Z \cdot 10^{n/2} + W)
\]

\[
A \cdot B = XZ \cdot 10^n + (XW +YZ) \cdot 10^{n/2} + YW \underbrace{0(n)}_{O(n)} + \underbrace{0(n)}_{O(n)}
\]

"conquer" step is \( O(n) \); \( T(n) = 4T\left(\frac{n}{2}\right) + O(n) \)
RUNNING TIME?

\[ T(n) = 4 \cdot T\left(\frac{n}{2}\right) + O(n) = C \cdot n . \]

\[
T(n) = 4 \cdot Cn + 4T\left(\frac{n}{2}\right)
\]

\[
= Cn + 4 \left[ C \cdot \frac{n}{2} + 4T\left(\frac{n}{4}\right) \right] = Cn + 2Cn + 4^2 \cdot T\left(\frac{n}{4}\right)
\]

\[
= Cn + 2Cn + 4^2 \left[ C \cdot \frac{n}{4} + 4^2 \cdot T\left(\frac{n}{8}\right) \right] = \ldots
\]

\[
= Cn + 2Cn + 2^2 Cn + \ldots + 2^{r-1} Cn + 4^{r-1} \cdot T\left(\frac{n}{2^r}\right)
\]

\[
\gamma = \log_2 n
\]

\[
4^{-\gamma} \cdot T(1) = \left(2^\gamma \right)^2 = n^{\frac{2}{\gamma}}
\]
CAN WE DO BETTER?

Reason to hope: we only need to find three quantities:

\[ XZ, (XW + YZ), YW \]
CAN WE DO BETTER?

**Reason to hope:** we only need to find three quantities:

$$XZ, (XW + YZ), YW$$

**Kartsuba’s observation:**

$$XW + YZ = (X + Y)(W + Z) − XZ − YW$$
OVERALL RUNNING TIME
EVEN BETTER?

► Can we do linear time?
► HW1 – will see something approaching linear
COMPUTING THE MEDIAN

**Question:** given an unsorted list A[0], A[1], ..., A[n-1], find the (n/2)th smallest element

*Divide and conquer?*
KEY INSIGHT: “APPROXIMATE MEDIAN” HELPS!
MEDIAN OF MEDIANS IS APPROXIMATE MEDIAN!