CSE 521
Algorithms:
Divide and Conquer

Larry Ruzzo

Thanks to Paul Beame, Kevin Wayne for some slides
algorithm design paradigms: divide and conquer

Outline:

General Idea

Review of Merge Sort

Why does it work?
  Importance of balance
  Importance of super-linear growth

Some interesting applications
  Closest points
  Integer Multiplication

Finding & Solving Recurrences
Divide & Conquer

Reduce problem to one or more sub-problems of the same type

Typically, each sub-problem is at most a constant fraction of the size of the original problem

Subproblems typically disjoint

Often gives significant, usually polynomial, speedup

Examples:

  - Binary Search, Mergesort, Quicksort (roughly),
  - Strassen’s Algorithm, …
MS(A: array[1..n]) returns array[1..n] {
    If(n=1) return A;
    New U:array[1..n/2] = MS(A[1..n/2]);
    New L:array[1..n/2] = MS(A[n/2+1..n]);
    Return(Merge(U,L));
}

Merge(U,L: array[1..n]) {
    New C: array[1..2n];
    a=1; b=1;
    For i = 1 to 2n
        C[i] = “smaller of U[a], L[b] and correspondingly a++ or b++”;
    Return C;
}
why balanced subdivision?

Alternative "divide & conquer" algorithm:
Sort n-1
Sort last 1
Merge them

\[ T(n) = T(n-1) + T(1) + 3n \quad \text{for } n \geq 2 \]
\[ T(1) = 0 \]

Solution: \[ 3n + 3(n-1) + 3(n-2) \ldots = \Theta(n^2) \]
Suppose we've already invented DumbSort, taking time $n^2$

Try *Just One Level* of divide & conquer:

- DumbSort(first $n/2$ elements)
- DumbSort(last $n/2$ elements)

Merge results

Time: $2 \left(\frac{n}{2}\right)^2 + n = \frac{n^2}{2} + n \ll n^2$

*Almost twice as fast!*
Moral 1: “two halves are better than a whole”
Two problems of half size are better than one full-size problem, even given $O(n)$ overhead of recombining, since the base algorithm has super-linear complexity.

Moral 2: “If a little's good, then more's better”
Two levels of D&C would be almost 4 times faster, 3 levels almost 8, etc., even though overhead is growing. Best is usually full recursion down to some small constant size (balancing "work" vs "overhead"). In the limit: you’ve just rediscovered mergesort!
Moral 3: unbalanced division less good:

\[(.1n)^2 + (.9n)^2 + n = .82n^2 + n\]

The 18% savings compounds significantly if you carry recursion to more levels, actually giving $O(n \log n)$, but with a bigger constant. So worth doing if you can’t get 50-50 split, but balanced is better if you can.

This is intuitively why Quicksort with random splitter is good – badly unbalanced splits are rare, and not instantly fatal.

\[(1)^2 + (n-1)^2 + n = n^2 - 2n + 2 + n\]

Little improvement here.
Mergesort: (recursively) sort 2 half-lists, then merge results.

\[ T(n) = 2T(n/2) + cn, \quad n \geq 2 \]
\[ T(1) = 0 \]

Solution: \( \Theta(n \log n) \)
(details later)
A Divide & Conquer Example: Closest Pair of Points
closest pair of points: non-geometric version

Given \( n \) points and \textit{arbitrary} distances between them, find the closest pair. (E.g., think of distance as airfare – definitely \textit{not} Euclidean distance!)

\[ \binom{n}{2} \]

\( \cdots \) and all the rest of the \( \binom{n}{2} \) edges…

\textit{Must} look at all \( n \) choose 2 pairwise distances, else any one you didn’t check might be the shortest.

Also true for Euclidean distance in 1-2 dimensions?
closest pair of points: 1 dimensional version

Given n points on the real line, find the closest pair

Closest pair is *adjacent* in ordered list
Time $O(n \log n)$ to sort, if needed
Plus $O(n)$ to scan adjacent pairs
Key point: do *not* need to calc distances between all pairs: exploit geometry + ordering
Closest pair of points: 2 dimensional version

Closest pair. Given \( n \) points in the plane, find a pair with smallest Euclidean distance between them.

Fundamental geometric primitive.

Graphics, computer vision, geographic information systems, molecular modeling, air traffic control.

Special case of nearest neighbor, Euclidean MST, Voronoi.

Brute force. Check all pairs of points \( p \) and \( q \) with \( \Theta(n^2) \) comparisons.

1-D version. \( O(n \log n) \) easy if points are on a line.

Assumption. No two points have same \( x \) coordinate.

Just to simplify presentation
closest pair of points. 2d, Euclidean distance: 1st try

Divide. Sub-divide region into 4 quadrants.
Divide. Sub-divide region into 4 quadrants.
Obstacle. Impossible to ensure n/4 points in each piece.
Algorithm.

Divide: draw vertical line $L$ with $\approx n/2$ points on each side.
closest pair of points

Algorithm.

Divide: draw vertical line $L$ with $\approx n/2$ points on each side.
Conquer: find closest pair on each side, recursively.
Algorithm.

Divide: draw vertical line \( L \) with \( \approx n/2 \) points on each side.
Conquer: find closest pair on each side, recursively.
Combine: find closest pair with one point in each side.
Return best of 3 solutions.
Find closest pair with one point in each side, *assuming* distance $< \delta$.

$\delta = \min(12, 21)$
Find closest pair with one point in each side, *assuming* distance $< \delta$.

Observation: suffices to consider points within $\delta$ of line $L$. 

$\delta = \min(12, 21)$
Find closest pair with one point in each side, *assuming* distance $< \delta$.

Observation: suffices to consider points within $\delta$ of line $L$.

Almost the one-D problem again: Sort points in $2\delta$-strip by their $y$ coordinate.

\[ \delta = \min(12, 21) \]
Find closest pair with one point in each side, *assuming* distance < \(\delta\).

Observation: suffices to consider points within \(\delta\) of line \(L\).

Almost the one-D problem again: Sort points in \(2\delta\)-strip by their \(y\) coordinate. Only check pts within 8 in sorted list!
Def. Let $s_i$ have the $i^{th}$ smallest $y$-coordinate among points in the $2\delta$-width-strip.

Claim. If $|i - j| > 8$, then the distance between $s_i$ and $s_j$ is $> \delta$.

Pf: No two points lie in the same $\frac{1}{2}\delta$-by-$\frac{1}{2}\delta$ box:

$$\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = \frac{\sqrt{2}}{2} \approx 0.7 < 1$$

so $\leq 8$ boxes within $+\delta$ of $y(s_i)$. 

---

**Closest pair of points**
closest pair algorithm

Closest-Pair(p₁, …, pₙ) {
    if(n <= ??) return ??

    Compute separation line L such that half the points
    are on one side and half on the other side.

    δ₁ = Closest-Pair(left half)
    δ₂ = Closest-Pair(right half)
    δ = min(δ₁, δ₂)

    Delete all points further than δ from separation line L

    Sort remaining points p[1]…p[m] by y-coordinate.

    for i = 1..m
        k = 1
        while i+k <= m && p[i+k].y < p[i].y + δ
            δ = min(δ, distance between p[i] and p[i+k]);
            k++;

    return δ.
}
Analysis, I: Let $D(n)$ be the number of pairwise distance calculations in the Closest-Pair Algorithm when run on $n \geq 1$ points.

$$D(n) \leq \begin{cases} 0 & n = 1 \\ 2D(n/2) + 7n & n > 1 \end{cases} \Rightarrow D(n) = O(n \log n)$$

BUT – that’s only the number of distance calculations.

What if we counted comparisons?
Analysis, II: Let $C(n)$ be the number of comparisons between coordinates/distances in the Closest-Pair Algorithm when run on $n \geq 1$ points

$$C(n) \leq \begin{cases} 
0 & n = 1 \\
2C(n/2) + kn \log n & n > 1 
\end{cases} \quad \Rightarrow \quad C(n) = O(n \log^2 n)$$

for some constant $k$

Q. Can we achieve $O(n \log n)$?

A. Yes. Don't sort points from scratch each time.
Sort by $x$ at top level only.
Each recursive call returns $\delta$ and list of all points sorted by $y$
Sort by merging two pre-sorted lists.

$$T(n) \leq 2T(n/2) + O(n) \quad \Rightarrow \quad T(n) = O(n \log n)$$
is it worth the effort?

Code is longer & more complex

$O(n \log n)$ vs $O(n^2)$ may hide $10x$ in constant?

How many points?

<table>
<thead>
<tr>
<th>n</th>
<th>Speedup: $n^2 / (10 \cdot n \log_2 n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.3</td>
</tr>
<tr>
<td>100</td>
<td>1.5</td>
</tr>
<tr>
<td>1,000</td>
<td>10</td>
</tr>
<tr>
<td>10,000</td>
<td>75</td>
</tr>
<tr>
<td>100,000</td>
<td>602</td>
</tr>
<tr>
<td>1,000,000</td>
<td>5,017</td>
</tr>
<tr>
<td>10,000,000</td>
<td>43,004</td>
</tr>
</tbody>
</table>
Going From Code to Recurrence
Carefully define what you’re counting, and write it down!

“Let $C(n)$ be the number of comparisons between sort keys used by MergeSort when sorting a list of length $n \geq 1$”

In code, clearly separate base case from recursive case, highlight recursive calls, and operations being counted.

Write Recurrence(s)
merge sort

MS(A: array[1..n]) returns array[1..n] {
  If(n=1) return A;
  New L:array[1:n/2] = MS(A[1..n/2]);
  New R:array[1:n/2] = MS(A[n/2+1..n]);
  Return(Merge(L,R));
}

Merge(A,B: array[1..n]) {
  New C: array[1..2n];
  a=1; b=1;
  For i = 1 to 2n {
    C[i] = “smaller of A[a], B[b] and a++ or b++”;
  }
  Return C;
}
the recurrence

\[ C(n) = \begin{cases} 
0 & \text{if } n = 1 \\
2C(n/2) + (n - 1) & \text{if } n > 1 
\end{cases} \]

Base case

Recursive calls

One compare per element added to merged list, except the last.

Total time: proportional to \( C(n) \)
(loops, copying data, parameter passing, etc.)
Carefully define what you’re counting, and write it down!

“Let $D(n)$ be the number of pairwise distance calculations in the Closest-Pair Algorithm when run on $n \geq 1$ points”

In code, clearly separate base case from recursive case, highlight recursive calls, and operations being counted.

Write Recurrence(s)
Closest Pair \(p_1, \ldots, p_n\) {
  if \(n \leq 1\) return \(\infty\)

  Compute separation line \(L\) such that half the points are on one side and half on the other side.

  \[\delta_1 = \text{Closest Pair(left half)}\]
  \[\delta_2 = \text{Closest Pair(right half)}\]
  \[\delta = \min(\delta_1, \delta_2)\]

  Delete all points further than \(\delta\) from separation line \(L\).

  Sort remaining points \(p[1] \ldots p[m]\) by \(y\)-coordinate.

  for \(i = 1 \ldots m\)
    \(k = 1\)
    while \(i+k \leq m \land p[i+k].y < p[i].y + \delta\)
      \(\delta = \min(\delta, \text{distance between } p[i] \text{ and } p[i+k])\)
      \(k++\)
  return \(\delta\).}
Analysis, I: Let $D(n)$ be the number of pairwise distance calculations in the Closest-Pair Algorithm when run on $n \geq 1$ points.

$$D(n) \leq \begin{cases} 0 & n = 1 \\ 2D(n/2) + 7n & n > 1 \end{cases} \Rightarrow D(n) = O(n \log n)$$

BUT – that’s only the number of distance calculations.

What if we counted comparisons?
Carefully define what you’re counting, and write it down!

“Let \( D(n) \) be the number of comparisons between coordinates/distances in the Closest-Pair Algorithm when run on \( n \geq 1 \) points”

In code, clearly separate \textit{base case} from \textit{recursive case}, highlight \textit{recursive calls}, and \textit{operations being counted}.

Write Recurrence(s)
Closest Pair \((p_1, \ldots, p_n)\) {
    if\((n \leq 1)\) return \(\infty\)
    Compute separation line \(L\) such that half the points are on one side and half on the other side.
    \[
    \delta_1 = \text{Closest Pair(left half)}
    \]
    \[
    \delta_2 = \text{Closest Pair(right half)}
    \]
    \[
    \delta = \min(\delta_1, \delta_2)
    \]
    Delete all points further than \(\delta\) from separation line \(L\)
    Sort remaining points \(p[1]\ldots p[m]\) by \(y\)-coordinate.
    for \(i = 1..m\)
        \(k = 1\)
        while \(i + k \leq m \land p[i+k].y < p[i].y + \delta\)
            \(\delta = \min(\delta, \text{distance between } p[i] \text{ and } p[i+k])\)
            \(k++\)
    return \(\delta\).
}
Analysis, II: Let $C(n)$ be the number of comparisons of coordinates/distances in the Closest-Pair Algorithm when run on $n \geq 1$ points

$$C(n) \leq \begin{cases} 
0 & n = 1 \\
2C(n/2) + k_4 n \log n & n > 1
\end{cases} \Rightarrow C(n) = O(n \log^2 n)$$

for some $k_4 \leq k_1 + k_2 + k_3 + 7$

Q. Can we achieve time $O(n \log n)$?

A. Yes. Don't sort points from scratch each time.
   Sort by $x$ at top level only.
   Each recursive call returns $\delta$ and list of all points sorted by $y$
   Sort by merging two pre-sorted lists.

$$T(n) \leq 2T(n/2) + O(n) \Rightarrow T(n) = O(n \log n)$$
Integer Multiplication
Add. Given two n-bit integers $a$ and $b$, compute $a + b$.

$O(n)$ bit operations.
**integer arithmetic**

**Add.** Given two n-bit integers $a$ and $b$, compute $a + b$.

**O(n) bit operations.**

**Multiply.** Given two n-bit integers $a$ and $b$, compute $a \times b$.

The “grade school” method:

**Θ(n²) bit operations.**
**divide & conquer multiplication: warmup**

To multiply two 2-digit integers:

Multiply four 1-digit integers.

Add, shift some 2-digit integers to obtain result.

\[
x = 10 \cdot x_1 + x_0 \\
y = 10 \cdot y_1 + y_0 \\
xy = (10 \cdot x_1 + x_0)(10 \cdot y_1 + y_0) \\
= 100 \cdot x_1y_1 + 10 \cdot (x_1y_0 + x_0y_1) + x_0y_0
\]

Same idea works for *long* integers – can split them into 4 half-sized ints
divide & conquer multiplication: warmup

To multiply two \( n \)-bit integers:

Multiply four \( \frac{1}{2}n \)-bit integers.

Add two \( \frac{1}{2}n \)-bit integers, and shift to obtain result.

\[
x = 2^{n/2} \cdot x_1 + x_0 \\
y = 2^{n/2} \cdot y_1 + y_0 \\
xy = (2^{n/2} \cdot x_1 + x_0)(2^{n/2} \cdot y_1 + y_0) \\
   = 2^n \cdot x_1 y_1 + 2^{n/2} \cdot (x_1 y_0 + x_0 y_1) + x_0 y_0
\]

\[
T(n) = 4T(n/2) + \Theta(n) \quad \Rightarrow \quad T(n) = \Theta(n^2)
\]

\[\quad \uparrow \]

assumes \( n \) is a power of 2
key trick: 2 multiplies for the price of 1:

\[
\begin{align*}
    x &= 2^{n/2} \cdot x_1 + x_0 \\
    y &= 2^{n/2} \cdot y_1 + y_0 \\
    xy &= (2^{n/2} \cdot x_1 + x_0)(2^{n/2} \cdot y_1 + y_0) \\
    &\quad = 2^n \cdot x_1 y_1 + 2^{n/2} \cdot (x_1 y_0 + x_0 y_1) + x_0 y_0
\end{align*}
\]

Well, ok, 4 for 3 is more accurate…

\[
\begin{align*}
    \alpha &= x_1 + x_0 \\
    \beta &= y_1 + y_0 \\
    \alpha \beta &= (x_1 + x_0)(y_1 + y_0) \\
    &\quad = x_1 y_1 + (x_1 y_0 + x_0 y_1) + x_0 y_0 \\
    (x_1 y_0 + x_0 y_1) &= \alpha \beta - x_1 y_1 - x_0 y_0
\end{align*}
\]
Karatsuba multiplication

To multiply two \( n \)-bit integers:

Add two \( \frac{1}{2}n \) bit integers.

Multiply three \( \frac{1}{2}n \)-bit integers.

Add, subtract, and shift \( \frac{1}{2}n \)-bit integers to obtain result.

\[
\begin{align*}
x &= 2^{n/2} \cdot x_1 + x_0 \\
y &= 2^{n/2} \cdot y_1 + y_0 \\
xy &= 2^n \cdot x_1y_1 + 2^{n/2} \cdot (x_1y_0 + x_0y_1) + x_0y_0 \\
&= 2^n \cdot x_1y_1 + 2^{n/2} \cdot (x_1 + x_0)(y_1 + y_0) - x_1y_1 - x_0y_0 + x_0y_0 \\
&= 2^n \cdot x_1y_1 + 2^{n/2} \cdot (x_1 + x_0)(y_1 + y_0) - x_1y_1 - x_0y_0 + x_0y_0
\end{align*}
\]

Theorem. [Karatsuba-Ofman, 1962] Can multiply two \( n \)-digit integers in \( O(n^{1.585}) \) bit operations.

\[
T(n) \leq T\left(\left\lfloor \frac{n}{2} \right\rfloor \right) + T\left(\left\lfloor \frac{n}{2} \right\rfloor \right) + T\left(1 + \left\lfloor \frac{n}{2} \right\rfloor \right) + \Theta(n)
\]

Sloppy version: \( T(n) \leq 3T(n/2) + O(n) \)

\( \Rightarrow T(n) = O(n^{\log_2 3}) = O(n^{1.585}) \)
Karatsuba multiplication

Theorem. [Karatsuba-Ofman, 1962] Can multiply two $n$-digit integers in $O(n^{1.585})$ bit operations.

\[
T(n) \leq \underbrace{T(\lfloor n/2 \rfloor)}_{\text{recursive calls}} + \underbrace{T(\lceil n/2 \rceil)}_{\text{recursive calls}} + \underbrace{T(1+\lceil n/2 \rceil)}_{\Theta(n)} + \Theta(n)
\]

Sloppy version: $T(n) \leq 3T(n/2) + O(n)$

$\Rightarrow T(n) = O(n^{\log_2 3}) = O(n^{1.585})$
Karatsuba multiplication

Theorem. [Karatsuba-Ofman, 1962] Can multiply two $n$-digit integers in $O(n^{1.585})$ bit operations.

$$T(n) \leq T\left(\left\lfloor n/2 \right\rfloor \right) + T\left(\left\lceil n/2 \right\rceil \right) + T\left(1 + \left\lfloor n/2 \right\rfloor \right) + \Theta(n)$$

Rewritten term: Sloppy version: $T(n) \leq 3T(n/2) + O(n)$

$$\Rightarrow T(n) = O(n^{\log_2 3}) = O(n^{1.585})$$

$n \rightarrow 2^{\log_2 n}$
Karatsuba multiplication

Theorem. [Karatsuba-Ofman, 1962] Can multiply two n-digit integers in $O(n^{1.585})$ bit operations.

$$T(n) \leq T\left(\left\lfloor \frac{n}{2} \right\rfloor \right) + T\left(\left\lfloor \frac{n}{2} \right\rfloor \right) + T\left(1 + \left\lfloor \frac{n}{2} \right\rfloor \right) + \Theta(n)$$

Recursive calls: add, subtract, shift

Sloppy version: $T(n) \leq 3T(n/2) + O(n)$

$\Rightarrow T(n) = O(n^{\log_2 3}) = O(n^{1.585})$
Naïve: $\Theta(n^2)$

Karatsuba: $\Theta(n^{1.59\ldots})$

Amusing exercise: generalize Karatsuba to do 5 size $n/3$ subproblems $\rightarrow \Theta(n^{1.46\ldots})$

Best known: $\Theta(n \log n \log \log n)$

"Fast Fourier Transform"

but mostly unused in practice (unless you need really big numbers - a billion digits of $\pi$, say)

High precision arithmetic IS important for crypto
Another D&C Example: Multiplying Polynomials

Similar ideas apply to polynomial multiplication

We’ll describe the basic ideas by multiplying polynomials rather than integers.
In fact, it’s somewhat simpler: no carries!
Notes on Polynomials

These are just formal sequences of coefficients so when we show something multiplied by $x^k$ it just means shifted $k$ places to the left – basically no work

Usual Polynomial Multiplication:

\[
\begin{array}{c}
3x^2 + 2x + 2 \\
\times \quad x^2 - 3x + 1 \\
\hline
3x^2 + 2x + 2 \\
-9x^3 - 6x^2 - 6x \\
3x^4 + 2x^3 + 2x^2 \\
3x^4 - 7x^3 - x^2 - 4x + 2
\end{array}
\]
Polynomial Multiplication

Given:

Degree \( m-1 \) polynomials \( P \) and \( Q \)

\[
P = a_0 + a_1 x + a_2 x^2 + \ldots + a_{m-2} x^{m-2} + a_{m-1} x^{m-1}
\]

\[
Q = b_0 + b_1 x + b_2 x^2 + \ldots + b_{m-2} x^{m-2} + b_{m-1} x^{m-1}
\]

Compute:

Degree \( 2m-2 \) Polynomial \( P \times Q \)

\[
P \times Q = a_0 b_0 + (a_0 b_1 + a_1 b_0) x + (a_0 b_2 + a_1 b_1 + a_2 b_0) x^2
+ \ldots + (a_{m-2} b_{m-1} + a_{m-1} b_{m-2}) x^{2m-3} + a_{m-1} b_{m-1} x^{2m-2}
\]

Obvious Algorithm:

Compute all \( a_i b_j \) and collect terms

\( \Theta (m^2) \) time
Naïve Divide and Conquer

Assume \( m=2k \)

\[
P = (a_0 + a_1 x + a_2 x^2 + \ldots + a_{k-2} x^{k-2} + a_{k-1} x^{k-1}) + \\
(a_k + a_{k+1} x + \ldots + a_{m-2} x^{k-2} + a_{m-1} x^{k-1}) x^k
\]

\[
= P_0 + P_1 x^k \\
= Q_0 + Q_1 x^k
\]

\[
P Q = (P_0 + P_1 x^k)(Q_0 + Q_1 x^k)
\]

\[
= P_0 Q_0 + (P_1 Q_0 + P_0 Q_1) x^k + P_1 Q_1 x^{2k}
\]

4 sub-problems of size \( k=m/2 \) plus linear combining

\[
T(m)=4T(m/2)+cm
\]

Solution \( T(m) = \Theta(m^2) \)
Karatsuba’s Algorithm

A better way to compute terms

Compute

\[ P_0 Q_0 \]
\[ P_1 Q_1 \]
\[ (P_0 + P_1)(Q_0 + Q_1) \] which is \[ P_0 Q_0 + P_1 Q_0 + P_0 Q_1 + P_1 Q_1 \]

Then

\[ P_0 Q_1 + P_1 Q_0 = (P_0 + P_1)(Q_0 + Q_1) - P_0 Q_0 - P_1 Q_1 \]

3 sub-problems of size \( m/2 \) plus \( O(m) \) work

\[ T(m) = 3 \ T(m/2) + cm \]

\[ T(m) = O(m^{\alpha}) \] where \( \alpha = \log_2 3 = 1.585... \)
Karatsuba: Details

PolyMul(P, Q):

// P, Q are length m = 2k vectors, with P[i], Q[i] being
// the coefficient of x^i in polynomials P, Q respectively.
if (m==1) return (P[0]*Q[0]);

Let Pzero be elements 0..k-1 of P; Pone be elements k..m-1
Qzero, Qone : similar

Prod1 = PolyMul(Pzero, Qzero);  // result is a (2k-1)-vector
Prod2 = PolyMul(Pone, Qone);   // ditto
Pzo = Pzero + Pone;            // add corresponding elements
Qzo = Qzero + Qone;            // ditto
Prod3 = PolyMul(Pzo, Qzo);     // another (2k-1)-vector
Mid = Prod3 – Prod1 – Prod2;   // subtract corr. elements
R = Prod1 + Shift(Mid, m/2) + Shift(Prod2,m) // a (2m-1)-vector
Return( R );
Polynomials

Naïve: $\Theta(n^2)$

Karatsuba: $\Theta(n^{1.585...})$

Best known: $\Theta(n \log n)$

"Fast Fourier Transform"

Integers

Similar, but some ugly details re: carries, etc.

gives $\Theta(n \log n \log \log n)$,

but mostly unused in practice
d & c summary

Idea:

“Two halves are better than a whole”
if the base algorithm has super-linear complexity.

“If a little's good, then more's better”
repeat above, recursively

Applications: Many.

Binary Search, Merge Sort, (Quicksort), Closest points, Integer multiply,…
Recurrences

Above: Where they come from, how to find them

Next: how to solve them
Mergesort: (recursively) sort 2 half-lists, then merge results.

\[ T(n) = 2T(n/2)+cn, \quad n \geq 2 \]

\[ T(1) = 0 \]

Solution: \( \Theta(n \log n) \) (details later)
Solve: \[ T(1) = c \]
\[ T(n) = 2 \cdot T(n/2) + cn \]

<table>
<thead>
<tr>
<th>Level</th>
<th>Num</th>
<th>Size</th>
<th>Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1 = 2^0</td>
<td>n</td>
<td>cn</td>
</tr>
<tr>
<td>1</td>
<td>2 = 2^1</td>
<td>n/2</td>
<td>2cn/2</td>
</tr>
<tr>
<td>2</td>
<td>4 = 2^2</td>
<td>n/4</td>
<td>4cn/4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>i</td>
<td>2^i</td>
<td>n/2^i</td>
<td>2^i cn/2^i</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>k-1</td>
<td>2^{k-1}</td>
<td>n/2^{k-1}</td>
<td>2^{k-1} cn/2^{k-1}</td>
</tr>
<tr>
<td>k</td>
<td>2^k</td>
<td>n/2^k = 1</td>
<td>2^k T(1)</td>
</tr>
</tbody>
</table>

n = 2^k ; k = \log_2 n

Total Work: \[ c \cdot n \cdot (1+\log_2 n) \] (add last col)
Solve:  
\[ T(1) = c \]
\[ T(n) = 4 \cdot T(n/2) + cn \]

\[ n = 2^k \Rightarrow k = \log_2{n} \]

<table>
<thead>
<tr>
<th>Level</th>
<th>Num</th>
<th>Size</th>
<th>Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>4^0</td>
<td>cn</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>4^1</td>
<td>4cn/2</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>4^2</td>
<td>16cn/4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>i</td>
<td>4^i</td>
<td>n/2^i</td>
<td>4^i \cdot cn/2^i</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>k-1</td>
<td>4^{k-1}</td>
<td>n/2^{k-1}</td>
<td>4^{k-1} \cdot cn/2^{k-1}</td>
</tr>
<tr>
<td>k</td>
<td>4^k</td>
<td>n/2^k = 1</td>
<td>4^k \cdot T(1)</td>
</tr>
</tbody>
</table>

Total Work:  
\[ T(n) = \sum_{i=0}^{k} 4^i \cdot cn / 2^i = O(n^2) \]

\[ 4^k = (2^2)^k = (2^k)^2 = n^2 \]
Solve: \[ T(1) = c \]
\[ T(n) = 3 \cdot T(n/2) + cn \]

Level | Num  | Size   | Work     |
------|------|--------|----------|
0     | 1 = 3^0 | n      | cn       |
1     | 3 = 3^1 | n/2    | 3cn/2    |
2     | 9 = 3^2 | n/4    | 9cn/4    |
...   | ...    | ...    | ...      |
i     | 3^i    | n/2^i  | 3^i \cdot cn/2^i |
...   | ...    | ...    | ...      |
k-1   | 3^{k-1}| n/2^{k-1}| 3^{k-1} \cdot cn/2^{k-1} |
k     | 3^k    | n/2^k  = 1 | 3^k \cdot T(1) |

Total Work: \[ T(n) = \sum_{i=0}^{k} 3^i \cdot cn / 2^i \]
Theorem:

\[ 1 + x + x^2 + x^3 + \ldots + x^k = \frac{x^{k+1} - 1}{x - 1} \]

proof:

\[ y = 1 + x + x^2 + x^3 + \ldots + x^k \]

\[ xy = x + x^2 + x^3 + \ldots + x^k + x^{k+1} \]

\[ xy - y = x^{k+1} - 1 \]

\[ y(x - 1) = x^{k+1} - 1 \]

\[ y = \frac{x^{k+1} - 1}{x - 1} \]
Solve:

\[ T(1) = c \]

\[ T(n) = 3 \, T(n/2) + cn \] (cont.)

\[
T(n) = \sum_{i=0}^{k} 3^i \frac{cn}{2^i}
\]

\[
= cn \sum_{i=0}^{k} \frac{3^i}{2^i}
\]

\[
= cn \sum_{i=0}^{k} \left(\frac{3}{2}\right)^i
\]

\[
= cn \frac{\left(\frac{3}{2}\right)^{k+1} - 1}{\left(\frac{3}{2}\right) - 1}
\]
Solve:

\[ T(1) = c \]
\[ T(n) = 3 \cdot T(n/2) + cn \]  \hspace{1cm} \text{(cont.)}

\[
\begin{align*}
  cn & \left( \frac{3}{2} \right)^{k+1} - 1 \\
  \frac{\left( \frac{3}{2} \right)^{k+1}}{\left( \frac{3}{2} \right) - 1} & = 2cn \left( \left( \frac{3}{2} \right)^{k+1} - 1 \right) \\
  < 2cn \left( \frac{3}{2} \right)^{k+1} \\
  = 3cn \left( \frac{3}{2} \right)^{k} \\
  = 3cn \frac{3^{k}}{2^{k}}
\end{align*}
\]
Solve:

\[ T(1) = c \]
\[ T(n) = 3 \cdot T(n/2) + cn \quad \text{(cont.)} \]

\[
3cn \frac{3^k}{2^k} = 3cn \frac{3^{\log_2 n}}{2^{\log_2 n}} \\
= 3cn \frac{3^{\log_2 n}}{n} \\
= 3c3^{\log_2 n} \\
= 3c (n^{\log_2 3}) \\
= O(n^{1.585...})
\]

\[
a^{\log_b n} \\
= \left(b^{\log_b a}\right)^{\log_b n} \\
= \left(b^{\log_b n}\right)^{\log_b a} \\
= n^{\log_b a}
\]
divide and conquer – master recurrence

\[ T(n) = aT(n/b) + cn^k \]

for \( n > b \) then

- \( a > b^k \implies T(n) = \Theta(n^\log_b a) \) [many subprobs \( \rightarrow \) leaves dominate]
- \( a < b^k \implies T(n) = \Theta(n^k) \) [few subprobs \( \rightarrow \) top level dominates]
- \( a = b^k \implies T(n) = \Theta(n^k \log n) \) [balanced \( \rightarrow \) all log \( n \) levels contribute]

Fine print:
- \( a \geq 1; b > 1; c, d, k \geq 0; T(1) = d; n = b^t \) for some \( t > 0 \);
- \( a, b, k, t \) integers. True even if it is \( \lceil n/b \rceil \) instead of \( n/b \).
master recurrence: proof sketch

Expanding recurrence as in earlier examples, to get

\[ T(n) = n^h \ (d + c \ S) \]

where \( h = \log_b(a) \) (tree height) and \( S = \sum_{j=1}^{\log_b n} x^j \), where \( x = b^k / a \).

If \( c = 0 \) the sum \( S \) is irrelevant, and \( T(n) = O(n^h) \): all the work happens in the base cases, of which there are \( n^h \), one for each leaf in the recursion tree.

If \( c > 0 \), then the sum matters, and splits into 3 cases (like previous slide):

- if \( x < 1 \), then \( S < x/(1-x) = O(1) \). [\( S \) is just the first \( \log n \) terms of the infinite series with that sum].
- if \( x = 1 \), then \( S = \log_b(n) = O(\log n) \). [all terms in the sum are \( 1 \) and there are that many terms].
- if \( x > 1 \), then \( S = x \cdot (x^{1+\log_b(n)} - 1)/(x-1) \). After some algebra, \( n^h \cdot S = O(n^k) \).
another d&c example: fast exponentiation

Power(a,n)

Input: integer $n$ and number $a$

Output: $a^n$

Obvious algorithm

$n - 1$ multiplications

Observation:

if $n$ is even, $n = 2m$, then $a^n = a^m \cdot a^m$
divide & conquer algorithm

Power(a,n)
  if n = 0 then return(1)
  if n = 1 then return(a)
  x ← Power(a, ⌊n/2⌋)
  x ← x • x
  if n is odd then
    x ← a • x
  return(x)
Let $M(n)$ be number of multiplies

Worst-case recurrence:

$$M(n) = \begin{cases} 
0 & n \leq 1 \\
M\left(\left\lfloor \frac{n}{2} \right\rfloor \right) + 2 & n > 1 
\end{cases}$$

By master theorem

$$M(n) = \Theta(\log n) \quad (a=1, b=2, k=0)$$

More precise analysis:

$$M(n) = \left\lfloor \log_2 n \right\rfloor + \text{(\# of 1's in n's binary representation)} - 1$$

Time is $\Theta(M(n))$ if numbers $< $ word size, else also depends on length, multiply algorithm
Instead of $a^n$ want $a^n \mod N$

$$a^{i+j} \mod N = ((a^i \mod N) \cdot (a^j \mod N)) \mod N$$

same algorithm applies with each $x \cdot y$ replaced by

$$((x \mod N) \cdot (y \mod N)) \mod N$$

In RSA cryptosystem (widely used for security)

need $a^n \mod N$ where $a$, $n$, $N$ each typically have 1024 bits

Power: at most $2^{1024}$ multiplies of 1024 bit numbers

relatively easy for modern machines

Naive algorithm: $2^{1024}$ multiplies
Idea:

“Two halves are better than a whole”
if the base algorithm has super-linear complexity.

“If a little's good, then more's better”
repeat above, recursively

Analysis: recursion tree or Master Recurrence

Applications: Many.

Binary Search, Merge Sort, (Quicksort), Closest points, Integer multiply, exponentiation,…