CS 3100 Data Structures and Algorithms 2 Lecture 8: Divide and Conquer

Co-instructors: Robbie Hott and Ray Pettit Spring 2024

Readings in CLRS 4th edition:

• Section 4.1-4.4

Announcements

- PS3 due tomorrow
- PA2 coming soon
- Office hours
 - Prof Hott Office Hours: Mondays 11a-12p, Fridays 10-11a and 2-3p
 - Prof Pettit Office Hours: Mondays and Wednesdays 2:30-4:00p
 - TA office hours posted on our website
- Quizzes 1-2 coming February 29, 2024
 - Both quizzes taken the same day
 - If you have SDAC, please schedule for 1 exam (*not a quiz*)

Divide and Conquer

[CLRS Chapter 4]

Divide:

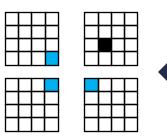
 Break the problem into multiple subproblems, each smaller instances of the original

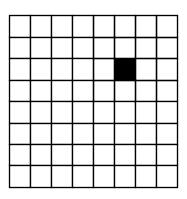
Conquer:

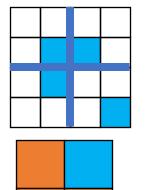
- If the suproblems are "large":
 - Solve each subproblem recursively
- If the subproblems are "small":
 - Solve them directly (base case)

Combine:

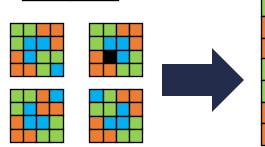
• Merge solutions to subproblems to obtain solution for original problem

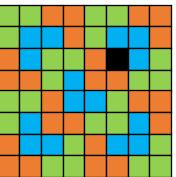






When is this an effective strategy?





Multiplication

Want to multiply large numbers together

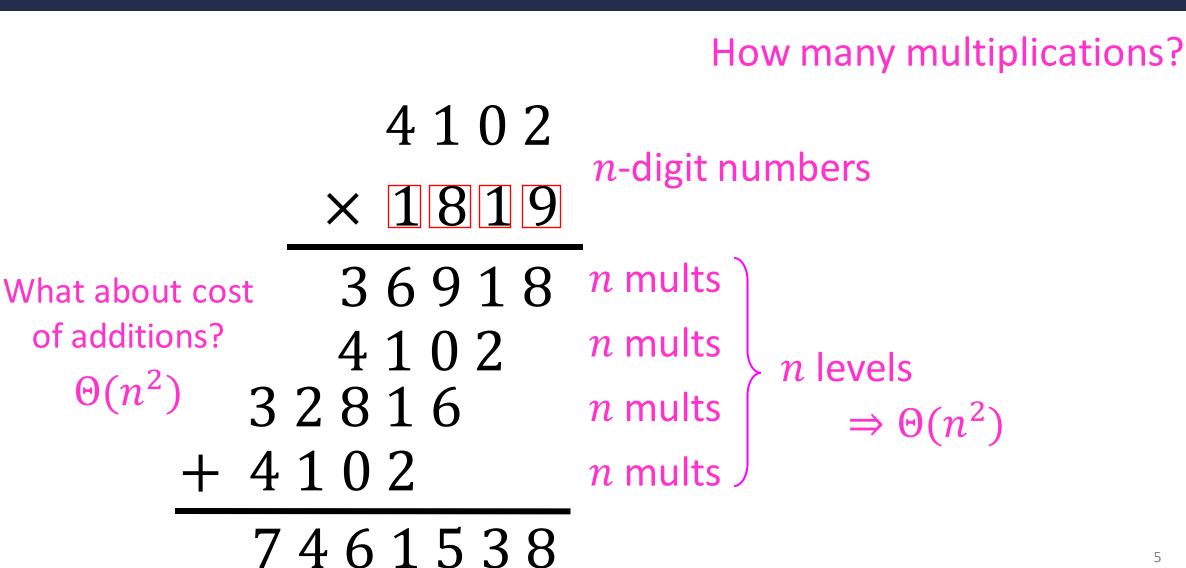
How do we measure input size?

What do we "count" for run time?

number of digits

number of <u>elementary</u> operations (single-digit multiplications)

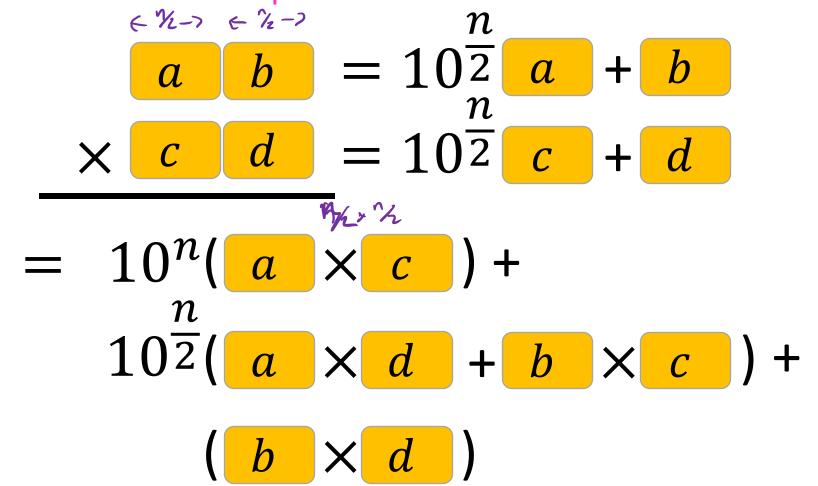
"Schoolbook" Multiplication



"Schoolbook" Multiplication

Can we do	How many multiplications?			
better?	4102	an altations		
	× 1819	<i>n</i> -digit numbers		
What about cost	36918	n mults		
of additions?	4102	n mults	> n levels	
$\Theta(n^2)$ 3	2816	<i>n</i> mults	$\Rightarrow \Theta(n^2)$	
+ 4102		n mults J		
7	461538			6

1. Break into smaller subproblems



Divide:

• Break *n*-digit numbers into four numbers of *n*/2 digits each (call them *a*, *b*, *c*, *d*)

Conquer:

- If n > 1:
 - Recursively compute *ac*, *ad*, *bc*, *bd*
- If n = 1: (i.e. one digit each)
 - Compute *ac*, *ad*, *bc*, *bd* directly (base case)

Combine:

• $10^n(ac) + 10^{n/2}(ad + bc) + bd$

For simplicity, assume that $n = 2^k$ is a power of 2

2. Use recurrence relation to express recursive running time

$$10^{n}(ac) + 10^{n/2}(ad + bc) + bd$$

Recursively solve

T(n)

2. Use recurrence relation to express recursive running time

$$10^{n}(ac) + 10^{n/2}(ad + bc) + bd$$

Recursively solve

$$T(n) = 4T\left(\frac{n}{2}\right)$$

Need to compute 4 multiplications, each of size n/2

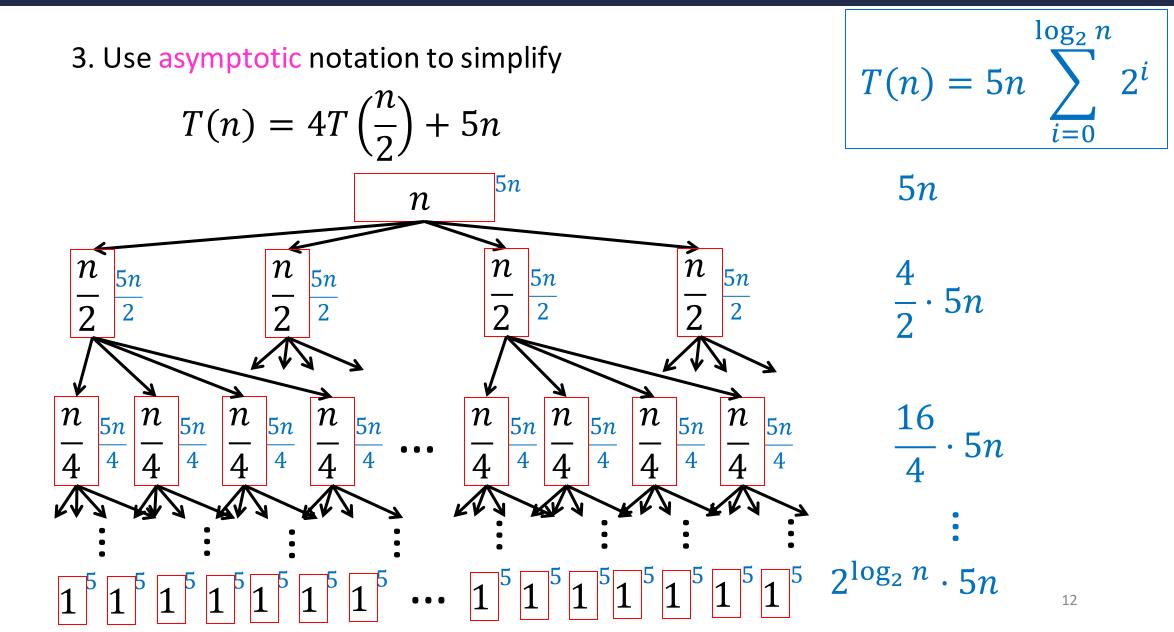
2. Use recurrence relation to express recursive running time

$$10^{n}(ac) + 10^{n/2}(ad + bc) + bd$$

Recursively solve

$$T(n) = 4T\left(\frac{n}{2}\right) + 5n$$

Need to compute 4 multiplications, each of size n/2 2 shifts and 3 additions on *n*-bit values



3. Use asymptotic notation to simplify

$$T(n) = 4T\left(\frac{n}{2}\right) + 5n$$

$$T(n) = 5n \sum_{i=0}^{\log_2 n} 2^i$$

$$T(n) = 5n \frac{2^{\log_2 n+1} - 1}{2 - 1}$$

$$T(n) = 5n(2n - 1) = \Theta(n^2)$$

No better than the schoolbook method!

1. Break into smaller subproblems

$$a \quad b = 10\frac{n}{2} \quad a + b$$

$$\times c \quad d = 10\frac{n}{2} \quad c + d$$

$$= 10^{n}(a \times c) + 10\frac{n}{2}(a \times d + b \times c) + 10\frac{n}{2}(a \times d \times c) + 10\frac{n}{2}(a \times c) + 10\frac{n}{2}($$

Recall: previous divideand-conquer recursively computed *ac*, *ad*, *bc*, *bd*

$$10^{n}(ac) + 10^{\frac{n}{2}}(ad + bc) + bd$$
This can be

Can't avoid these

This can be simplified!

$$(a+b)(c+d) =$$
$$ac + ad + bc + bd$$

$$\frac{ad+bc}{\mathsf{Two}} = \frac{(a+b)(c+d)-ac-bd}{ac-bd}$$

multiplications

One multiplication

2. Use recurrence relation to express recursive running time



$$10^{n}(ac) + 10^{n/2} ((a+b)(c+d) - ac - bd) + bd$$

Recursively solve

$$T(n) =$$

2. Use recurrence relation to express recursive running time



$$10^{n}(ac) + 10^{n/2}((a+b)(c+d) - ac - bd) + bd$$

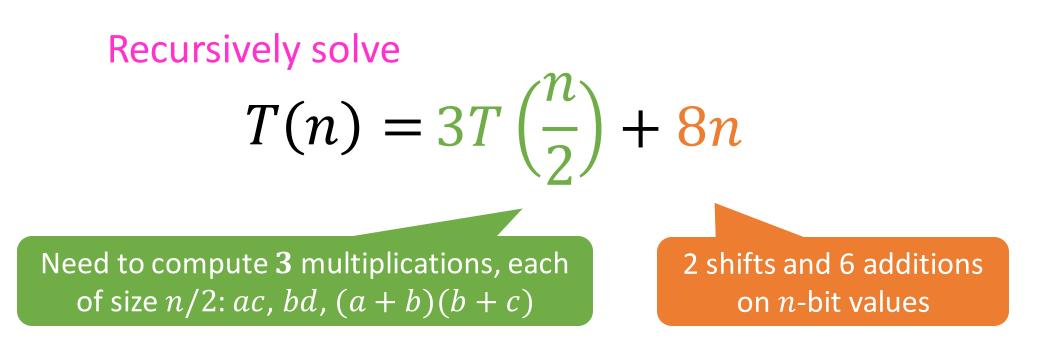
Recursively solve
$$T(n) = 3T\left(\frac{n}{2}\right)$$

Need to compute **3** multiplications, each of size n/2: ac, bd, (a + b)(b + c)

2. Use recurrence relation to express recursive running time



$$10^{n}(ac) + 10^{n/2}((a+b)(c+d) - ac - bd) + bd$$



Divide:

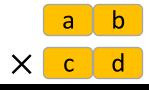
Break n-digit numbers into four numbers of ⁿ/₂ digits each (call them a, b, c, d)

Conquer:

- If n > 1:
 - Recursively compute ac, bd, (a + b)(c + d)
- If n = 1:
 - Compute ac, bd, (a + b)(c + d) directly (base case)

Combine:

• $10^{n}(ac) + 10^{n/2}((a+b)(c+d) - ac - bd) + bd$



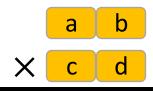
1. Recursively compute: ac, bd, (a + b)(c + d)2. $(ad + bc) \neq (a + b)(c + d) - ac - bd$ 3. Return $10^{n}(ac) + 10^{\frac{n}{2}}(ad + bc) + bd$

Pseudocode:

1. $x \leftarrow \text{Karatsuba}(a, c)$ 2. $y \leftarrow \text{Karatsuba}(a, d)$ 3. $z \leftarrow \text{Karatsuba}(a + b, c + d) - x - y$ $T(n) = 3T\left(\frac{n}{2}\right) + 8n$ 4. Return $10^n x + 10^{n/2} z + y$

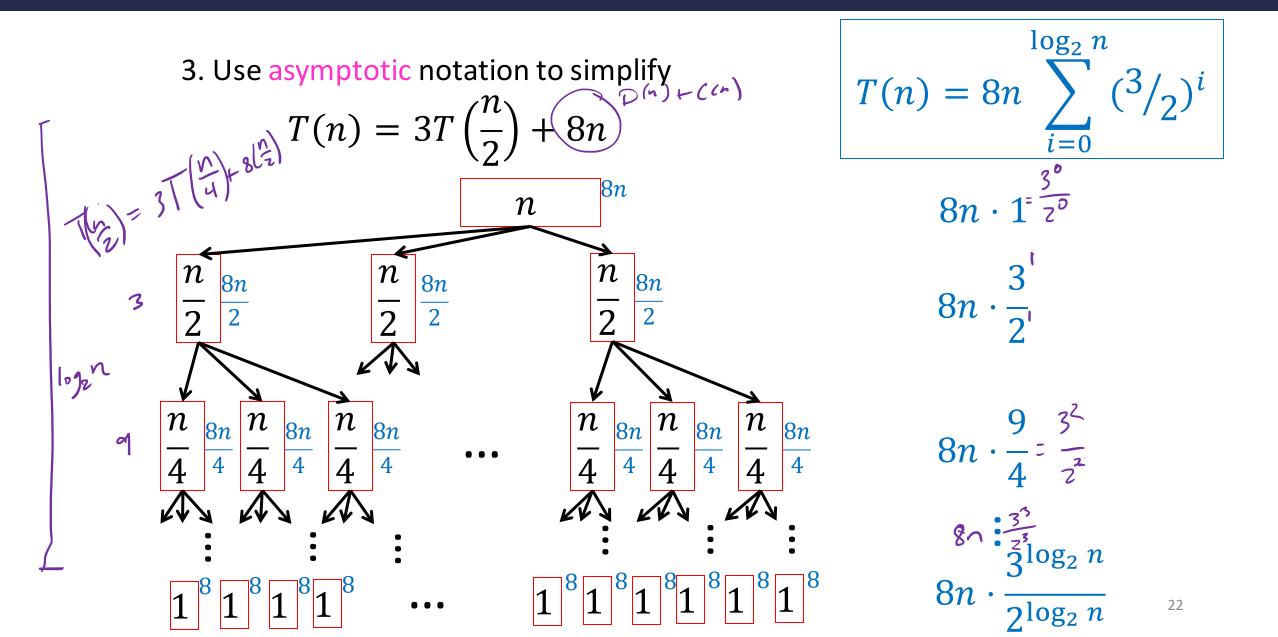
1. Recursively compute:
$$ac, bd, (a + b)(c + d)$$

2. $(ad + bc) = (a + b)(c + d) - ac - bd$
3. Return $10^{n}(ac) + 10^{\frac{n}{2}}(ad + bc) + bd$



Pseudocode:

1. $x \leftarrow \text{Karatsuba}(a, c)$ 2. $y \leftarrow \text{Karatsuba}(a, d)$ 3. $z \leftarrow \text{Karatsuba}(a + b, c + d) - x - y$ $T(n) = 3T\left(\frac{n}{2}\right) + 8n$ 4. Return $10^n x + 10^{n/2} z + y$



3. Use asymptotic notation to simplify $T(n) = 3T\left(\frac{n}{2}\right) + 8n$ $\log_2 n$ $T(n) = 8n \sum_{i=0}^{100} (3/2)^{i}$ $T(n) = 8n \frac{(3/2)^{\log_2 n+1} - 1}{3/2 - 1}$

Math, math, and more math...(on board, see lecture supplement)

 $3 = 2^{\log_2 3}$

 $T(n) = 8n \frac{(3/2)^{\log_2 n + 1} - 1}{(3/2)^{\log_2 n + 1}} = 16n \left(\frac{3}{2} \right)^{\log_2 n + 1} - 1 \right)$ $=16n\left(\frac{2^{1052^3}}{2}\right)^{1092^{n+1}}$ $= 16n \left(\left(2^{\log_2 3} - 1 \right)^{\log_2 m} + 1 - 1 \right)$ Megative Jbjr / bjz $-1) = 24 n^{\log_2 3}$ $= 16\eta \left(2^{10}\eta_{2}^{3}\eta_{2}^{9}\right) \cdot \left(2^{1}\eta_{2}^{9}\right) \cdot \left(2^{1}\eta_{2}$ (-1en)(2", $\frac{1}{2'2'} = \frac{1}{2}$ $\mathcal{C}\left(n^{1.32}\right) \approx \mathcal{O}\left(n^{1.585}\right)$ nlog2 S 3

$$T(n) = 8n \frac{(3/2)^{\log_2 n+1} - 1}{3/2 - 1}$$

How to simplify this (using asymptotic notation)?

Drop constant multiples

$$T(n) = 8n \frac{(3/2)^{\log_2 n+1} - 1}{3/2 - 1}$$

$$= \Theta\left(n\left(\frac{3}{2}\right)^{\log_2 n+1} - 1\right)\right)$$

$$= \Theta\left(\frac{3}{2}n \cdot \left(\frac{3}{2}\right)^{\log_2 n} - n\right)$$

How to simplify this (using asymptotic notation)?

Drop constant multiples

Distribute terms

$$T(n) = 8n \frac{(3/2)^{\log_2 n+1} - 1}{3/2 - 1}$$

$$= \Theta\left(n\left(\frac{3}{2}\right)^{\log_2 n+1} - 1\right)\right)$$

$$= \Theta\left(\frac{3}{2}n \cdot \left(\frac{3}{2}\right)^{\log_2 n} - n\right)$$

$$= \Theta\left(n \cdot \left(\frac{3}{2}\right)^{\log_2 n}\right)$$

How to simplify this (using asymptotic notation)?

Drop constant multiples

Distribute terms

Drop constants and loworder terms

28

$$T(n) = \Theta\left(n \cdot \left(\frac{3}{2}\right)^{\log_2 n}\right)$$

How to simplify this (using asymptotic notation)?

Properties of logarithms:

$$2^{\log_2 n} = n$$

 $3^{\log_2 n} = 2^{\log_2(3^{\log_2 n})} = 2^{(\log_2 n)(\log_2 3)} = (2^{\log_2 n})^{\log_2 3} = n^{\log_2 3}$
 $2^{\log_2 n} = n$
 $\log a^b = b \log a$
 $2^{\log_2 n} = n$

$$T(n) = \Theta\left(n \cdot \left(\frac{3}{2}\right)^{\log_2 n}\right)$$
$$= \Theta\left(n \cdot \left(\frac{3^{\log_2 n}}{2^{\log_2 n}}\right)\right)$$
$$= \Theta\left(n \cdot \left(\frac{n^{\log_2 3}}{n}\right)\right)$$
$$= \Theta(n^{\log_2 3}) \approx \Theta(n^{1.585})$$

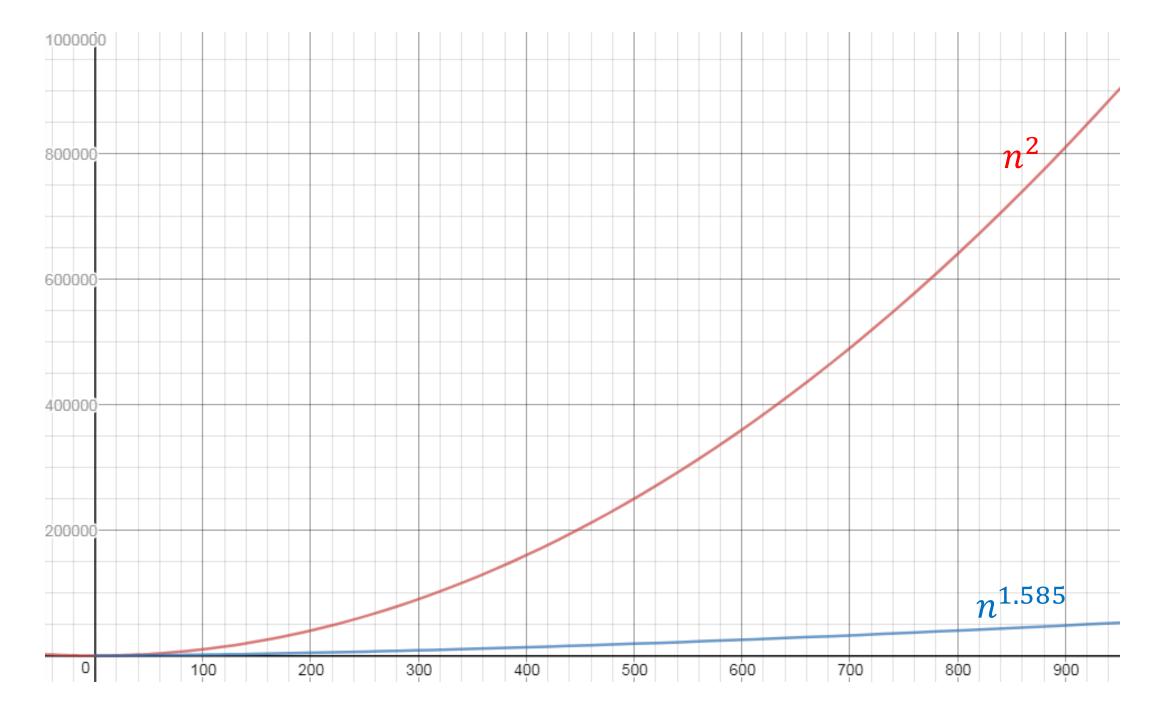
How to simplify this (using asymptotic notation)?

$$2^{\log_2 n} = n$$

 $3^{\log_2 n} = n^{\log_2 3}$

1......

Strictly better than schoolbook method!



Analyzing Divide and Conquer

- 1. Break into smaller subproblems
- 2. Use recurrence relation to express recursive running time
- 3. Use asymptotic notation to simplify

Divide: D(n) time

Conquer: Recurse on smaller problems of size s_1, \ldots, s_k

Combine: C(n) time

Recurrence:

•
$$T(n) = D(n) + \sum_{i \in [k]} T(s_i) + C(n)$$

Recurrence Solving Techniques





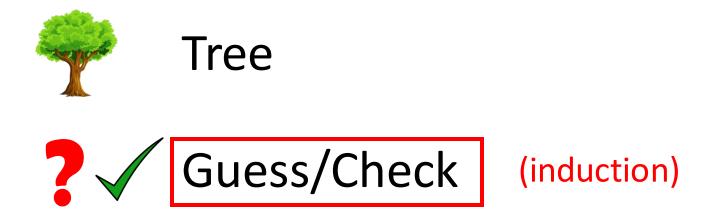
"Cookbook" MAGIC!



Substitution

substitute in to simplify

Recurrence Solving Techniques





"Cookbook"



Substitution

Guess and Check Blueprint

Show: T(n) = O(g(n))

Consider: $g_*(n) = c \cdot g(n)$ for some constant c

 $g_{*}(n) \in O(g(n))$

Goal: show $\exists n_0$ such that $\forall n > n_0$, $T(n) \leq g_*(n)$

• (definition of big-O)

Technique: Induction

- Base cases:
 - Show $T(1) \le g_*(1)$ (sometimes, may need to consider <u>additional</u> base cases)
- Hypothesis:
 - $\forall n \leq x_0, T(n) \leq g_*(n)$
- Inductive step:

• Show that
$$T(x_0 + 1) \le g_*(x_0 + 1)$$

Need to ensure that in inductive step, can either appeal to a <u>base</u> <u>case</u> or to the <u>inductive hypothesis</u>

Mergesort Guess and Check

$$T(n) = 2T(n/2) + n \in \Theta(n \log n) \quad g(n)$$

$$g_{x}(n) = n \log_{2} n \quad T(n) \in O(g_{x}(n)) \approx O(g_{y}(n))$$

$$\exists n \ge 1 \quad g_{2}(1) = n \log_{2} n \quad T(n) \in O(g_{y}(n))$$

$$Hyri \quad \forall n \le x_{0} \quad T(n) \le n \log_{2} n \quad \le x_{0}$$

$$Hyri \quad \forall n \le x_{0} \quad T(n) \le n \log_{2} n \quad \le x_{0}$$

$$T(x_{0}+1) = 2T \left(\frac{x_{0}+1}{2}\right) + (x_{0}+1)$$

$$= (x_{0}+1) \left(\log_{2}(x_{0}+1) - 1\right) + x_{0}+1$$

$$= (x_{0}+1) \log_{2}(x_{0}+1) - (x_{0}+1) + (x_{0}+1)$$

$$T(n) \in O(n \log n)$$

Karatsuba Analysis using Guess and Check

$$T(n) = 3T(n/2) + 8n \quad \in \mathcal{O}(n^{1/4})$$

Goal:
$$T(n) \le 3000 \ n^{1.6} = O(n^{1.6})$$

Base case: $T(1) = 8 \le 3000$

Hypothesis:

 $\forall n \le x_0, \ T(n) \le 3000 n^{1.6}$

Inductive step:

Show $T(x_0 + 1) \le 3000(x_0 + 1)^{1.6}$

Karatsuba Guess and Check (Loose)

$$T(n) = 3T(n/2) + 8n$$
Hypothesis: $\forall n \le x_0$: $T(n) \le 3000n^{1.6}$
Show: $T(x_0 + 1) \le 3000(x_0 + 1)^{1.6}$

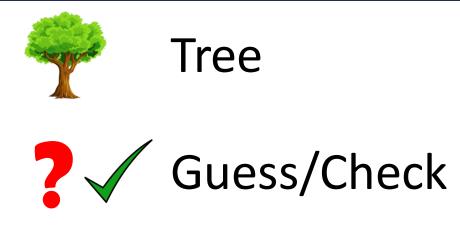
$$T(x_0 + 1) = 3T\left(\frac{x_0 + 1}{2}\right) + 8(x_0 + 1)$$
Recurrence definition
$$\le 3\left(3000\left(\frac{x_0 + 1}{2}\right)^{1.6}\right) + 8(x_0 + 1)$$
Inductive hypothesis

Karatsuba Guess and Check (Loose)

Karatsuba Guess and Check (Loose)

$$\begin{split} T(x_0+1) &= 3T\left(\frac{x_0+1}{2}\right) + 8(x_0+1) & \text{Recurrence definition} \\ &\leq 3\left(3000\left(\frac{x_0+1}{2}\right)^{1.6}\right) + 8(x_0+1) & \text{Inductive hypothesis} \\ &\leq 3\left(3000\left(\frac{x_0+1}{2}\right)^{1.6}\right) + 8(x_0+1)^{1.6} & \forall x \ge 0; x^{1.6} \ge x \\ &= \left(\frac{9000}{2^{1.6}} + 8\right)(x_0+1)^{1.6} & \text{Distributive property} \\ &\leq 3000(x_0+1)^{1.6} & \frac{9000}{2^{1.6}} + 8 \le 3000 \\ & \text{Show:} T(x_0+1) \le 3000(x_0+1)^{1.6} & 40 \end{split}$$

Recurrence Solving Techniques









Substitution

Observation

Divide: D(n) time

Conquer: Recurse on smaller problems of size s_1, \ldots, s_k

Combine: C(n) time

Recurrence:

• $T(n) = D(n) + \sum_{i \in [k]} T(s_i) + C(n)$

Many divide and conquer algorithms have recurrences are of form:

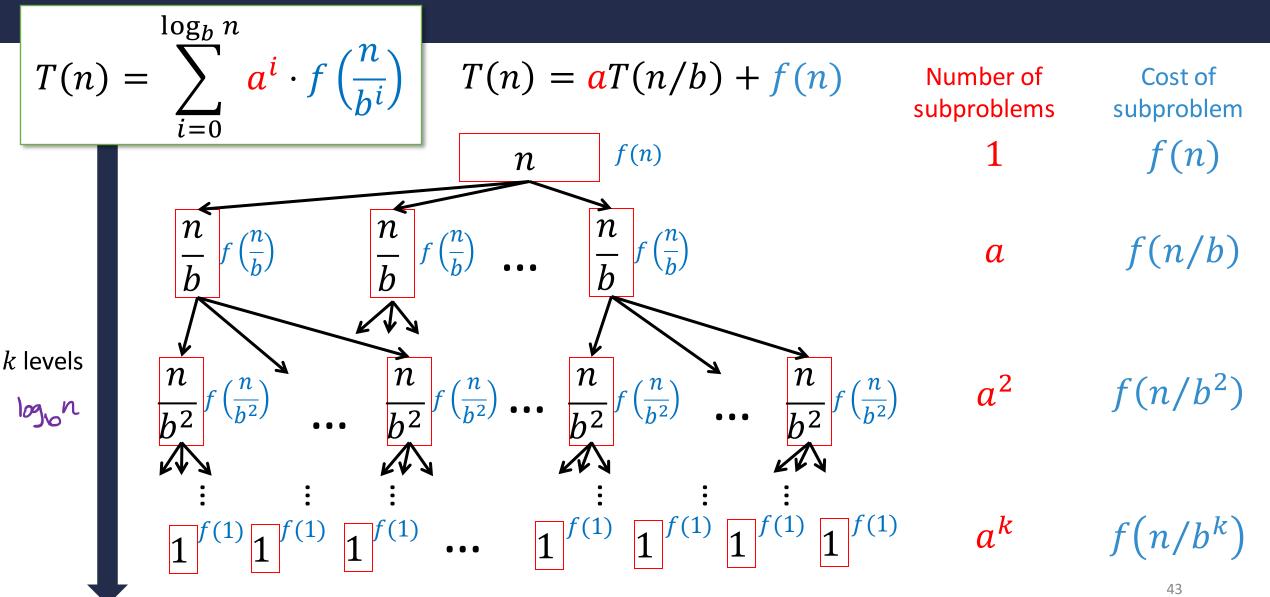
• $T(n) = a \cdot T(n/b) + f(n)$ a and b are constants

Mergesort: T(n) = 2T(n/2) + n

Divide and Conquer Multiplication: T(n) = 4T(n/2) + 5n

Karatsuba Multiplication: T(n) = 3T(n/2) + 8n

General Recurrence



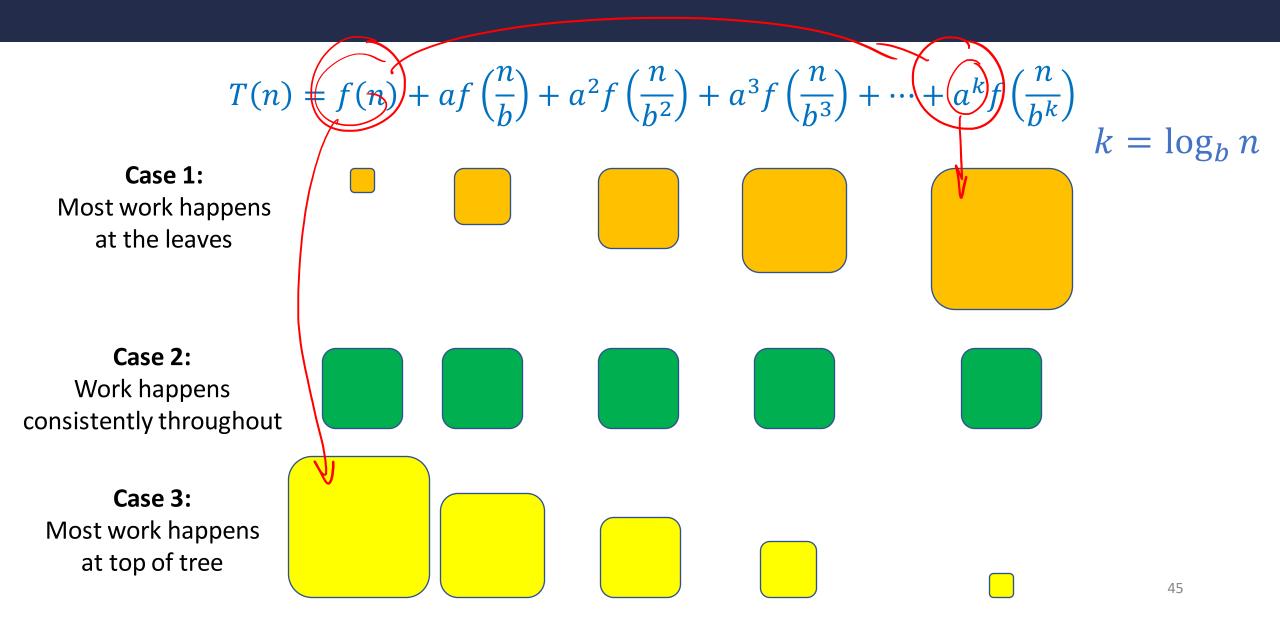
General Recurrence

$$k = \log_b n$$

An aside:

$$a^{\log_b n} = \left(b^{\log_b n} \right)^{\log_b n} = \left(b^{\log_b n} \right)^{\log_b n} = n^{\log_b n}$$

Three Cases



Master Theorem

$$T(n) = aT(n/b) + f(n)$$

$$\delta = \log_b a$$

	Requirement on <i>f</i>	Implication
Case 1	$f(n) \in O(n^{\delta-\varepsilon})$ for some constant $\varepsilon > 0$	$T(n) \in \Theta(n^{\delta})$

Master Theorem

$$T(n) = aT(n/b) + f(n)$$

$$\delta = \log_b a$$

	Requirement on <i>f</i>	Implication
Case 1	$f(n) \in O(n^{\delta-\varepsilon})$ for some constant $\varepsilon > 0$	$T(n) \in \Theta(n^{\delta})$
Case 2	$f(n) \in \Theta(n^{\delta})$	$T(n) \in \Theta(n^{\delta} \log n)$

Master Theorem

$$T(n) = aT(n/b) + f(n)$$

$$\delta = \log_b a$$

	Requirement on <i>f</i>	Implication
Case 1	$f(n) \in O(n^{\delta-\varepsilon})$ for some constant $\varepsilon > 0$	$T(n) \in \Theta(n^{\delta})$
Case 2	$f(n) \in \Theta(n^{\delta})$	$T(n) \in \Theta(n^{\delta} \log n)$
Case 3	$f(n) \in \Omega(n^{\delta + \varepsilon}) \text{ for some constant } \varepsilon > 0$ AND $af\left(\frac{n}{b}\right) \leq cf(n) \text{ for constant } c < 1 \text{ and}$ sufficiently large n	$T(n) \in \Theta(f(n))$

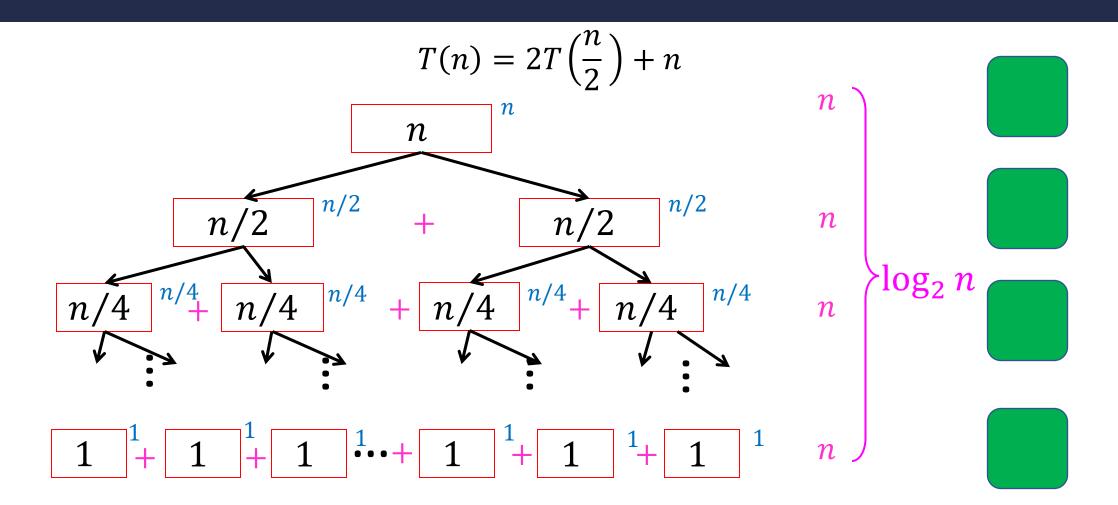
$$T(n) = aT\left(\frac{n}{b}\right) + f(n)$$
Case 1: if $f(n) = O(n^{\log_b a - \varepsilon})$ for some constant $\varepsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$
Case 2: if $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a} \log n)$
Case 3: if $f(n) = \Omega(n^{\log_b a + \varepsilon})$ for some constant $\varepsilon > 0$, and if $af\left(\frac{n}{b}\right) \le cf(n)$ for some constant $c < 1$ and all sufficiently large n , then $T(n) = \Theta(f(n))$

$$T(n) = 2T\left(\frac{n}{2}\right) + n$$

Case 2

$$\Theta(n^{\log_2 2} \log n) = \Theta(n \log n)$$

Tree method



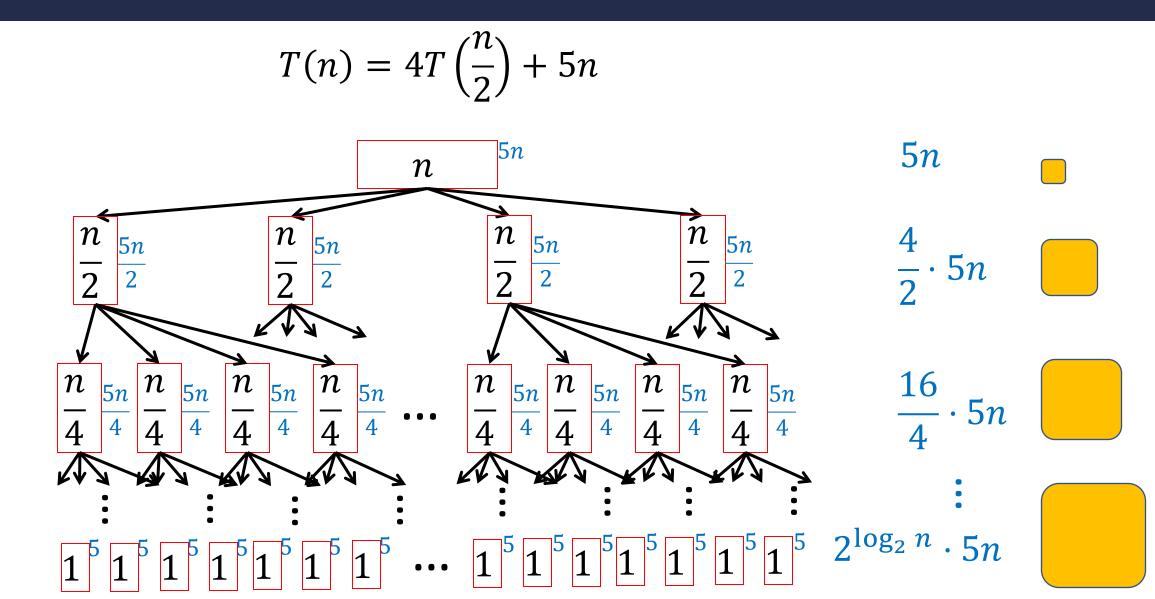
$$T(n) = aT\left(\frac{n}{b}\right) + f(n)$$
Case 1: if $f(n) = O(n^{\log_b a - \varepsilon})$ for some constant $\varepsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$
Case 2: if $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a} \log n)$
Case 3: if $f(n) = \Omega(n^{\log_b a + \varepsilon})$ for some constant $\varepsilon > 0$, and if $af\left(\frac{n}{b}\right) \le cf(n)$ for some constant $c < 1$ and all sufficiently large n , then $T(n) = \Theta(f(n))$

$$T(n) = 4T\left(\frac{n}{2}\right) + 5n$$

Case 1

$$\Theta(n^{\log_2 4}) = \Theta(n^2)$$

Tree method

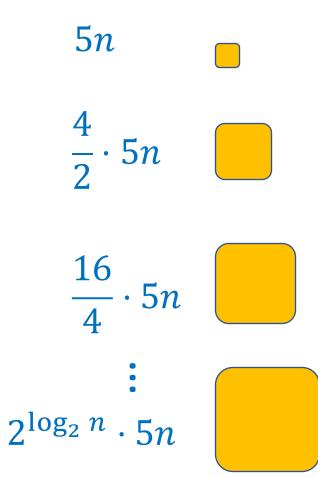


Tree method

$$T(n) = 4T\left(\frac{n}{2}\right) + 5n$$

Cost is <u>increasing</u> with the recursion depth (due to large number of subproblems)

Most of the work happening in the leaves



$$T(n) = aT\left(\frac{n}{b}\right) + f(n)$$
Case 1: if $f(n) = O(n^{\log_b a - \varepsilon})$ for some constant $\varepsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$
Case 2: if $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a} \log n)$
Case 3: if $f(n) = \Omega(n^{\log_b a + \varepsilon})$ for some constant $\varepsilon > 0$, and if $af\left(\frac{n}{b}\right) \le cf(n)$ for some constant $c < 1$ and all sufficiently large n , then $T(n) = \Theta(f(n))$

$$T(n) = 3T\left(\frac{n}{2}\right) + 8n$$

Case 1

$$\Theta(n^{\log_2 3}) \approx \Theta(n^{1.585})$$

Karatsuba

$$T(n) = 3T\left(\frac{n}{2}\right) + 8n$$

$$n = 3T\left(\frac{n}{2}\right) + 8n$$

$$8 \cdot 1n$$

$$\frac{n}{2} \frac{8n}{2} - \frac{n}{2} \frac{8n}{2} - \frac{n}{2} \frac{8n}{2} - \frac{2}{2} - \frac{8}{2} \cdot 3n$$

$$\frac{n}{4} \frac{8n}{4} \frac{n}{4} \frac{8n}{4} - \frac{n}{4} \frac{8n}{4} \frac{n}{4} \frac{8n}{4} - \frac{8}{4} - \frac{8}{4}$$

$$T(n) = aT\left(\frac{n}{b}\right) + f(n)$$
Case 1: if $f(n) = O(n^{\log_b a - \varepsilon})$ for some constant $\varepsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$
Case 2: if $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a} \log n)$
Case 3: if $f(n) = \Omega(n^{\log_b a + \varepsilon})$ for some constant $\varepsilon > 0$, and if $af\left(\frac{n}{b}\right) \le cf(n)$ for some constant $c < 1$ and all sufficiently large n , then $T(n) = \Theta(f(n))$

$$T(n) = 2T\left(\frac{n}{2}\right) + 15n^3$$

Case 3

$$T(n) = aT\left(\frac{n}{b}\right) + f(n)$$
Case 1: if $f(n) = O(n^{\log_b a - \varepsilon})$ for some constant $\varepsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$
Case 2: if $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a} \log n)$
Case 3: if $f(n) = \Omega(n^{\log_b a + \varepsilon})$ for some constant $\varepsilon > 0$, and if $af\left(\frac{n}{b}\right) \le cf(n)$ for some constant $c < 1$ and all sufficiently large n , then $T(n) = \Theta(f(n))$

$$T(n) = 2T\left(\frac{n}{2}\right) + 15n^3$$

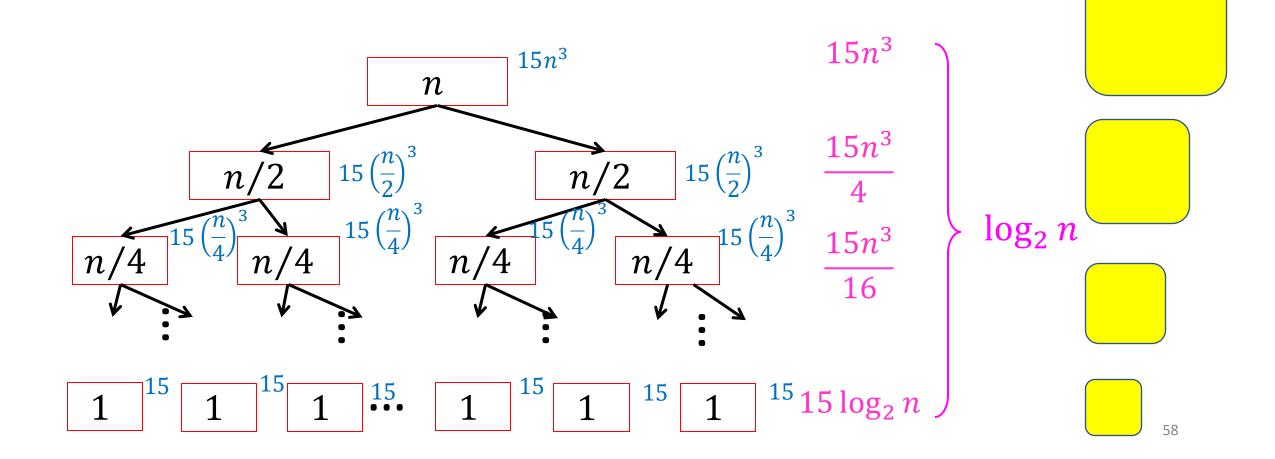
Case 3
 $\Theta(n^3)$

Important: For Case 3, need to additionally check that $2f(n/2) \le cf(n)$ for constant c < 1 and sufficiently large n

$$2f(n/2) = 30(n/2)^3 = \frac{30}{8}n^3 \le \frac{1}{4}(15n^3)$$

Master Theorem Example 4 (Visually)

 $T(n) = 2T(n/2) + 15n^3$

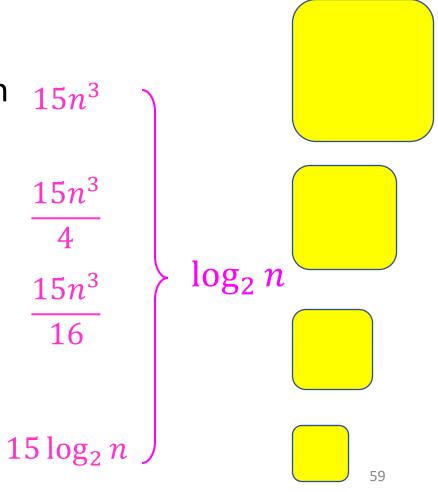


Master Theorem Example 4 (Visually)

$$T(n) = 2T(n/2) + 15n^3$$

Cost is <u>decreasing</u> with the recursion depth $15n^{3}$ (due to high *non-recursive* cost)

Most of the work happening at the top



 $15n^{3}$

4

 $15n^{3}$

16