

Lecture 15

Asymptotics 2

CS61B, Spring 2024 @ UC Berkeley

Slides Credit: Josh Hug



If my function f(n) is	Doubling N	Adding 1 to N
Θ(1)	Doesn't affe	ect runtime
Θ(log n) (base independent)	Adds 1 to runtime	Affects runtime minimally
Θ(n)	Doubles runtime	Adds 1 to runtime
Θ(n ²)	Quadruples runtime	Adds n to runtime
Θ(2 ⁿ) (base dependent)	Squares runtime	Doubles runtime



Nested For Loops

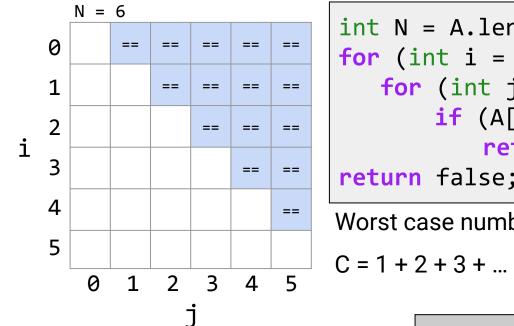
Lecture 15, CS61B, Spring 2024

Nested For Loops

There is No Magic Shortcut for Asymptotic Analysis Amortized Analysis Recursive Analysis Binary Search (Intuitive) Binary Search (Exact) (Bonus Video) Mergesort



Find the order of growth of the worst case runtime of dup1.



Worst case number of operations:

C = 1 + 2 + 3 + ... + (N - 3) + (N - 2) + (N - 1) = N(N-1)/2

operation	worst case count
==	Θ(N ²)



Find the order of growth of the worst case runtime of dup1.

Worst case number of operations:

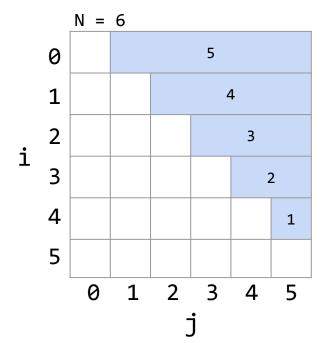
C = 1 + 2 + 3 + ... + (N - 3) + (N - 2) + (N - 1) = N(N-1)/2

operation	worst case count
Total	Θ(N ²)



Loops Example 1: Based on Exact Count

Find the order of growth of the worst case runtime of dup1.



Worst case number of operations:

C = 1 + 2 + 3 + ... + (N - 3) + (N - 2) + (N - 1) = N(N-1)/2

operation	worst case count
Total	Θ(N ²)



Loops Example 1: Based on Exact Count

Find the order of growth of the worst case runtime of dup1.

int N = A.length;
for (int i = 0; i < N; i += 1)
 // N-i-1 units of work
return false;</pre>

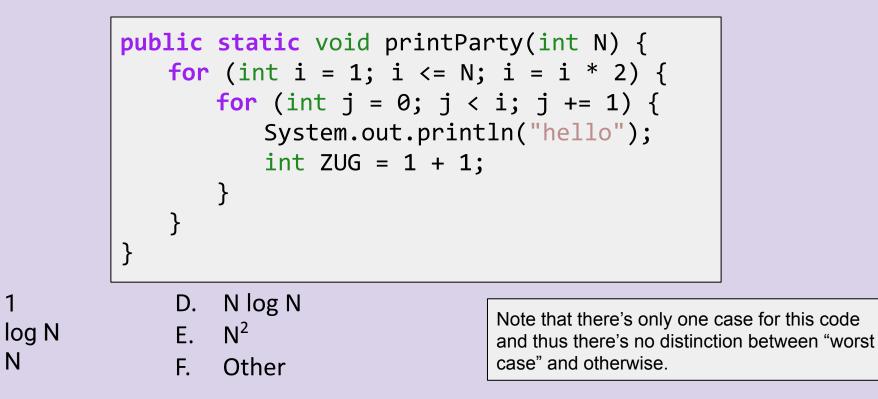
Worst case number of operations:

- Given by area of right triangle of side length N-1.
- Area is $\Theta(N^2)$.

operation	worst case count
Total	Θ(N ²)



Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$. By simple, we mean there should be no unnecessary multiplicative constants or additive terms.





Α.

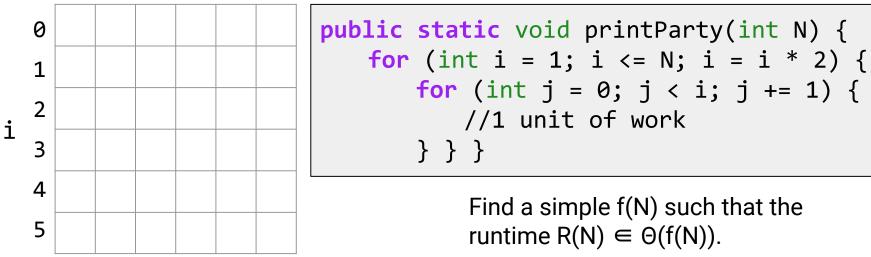
Β.

C.

Ν

0 1 2 3 4 5

Ť

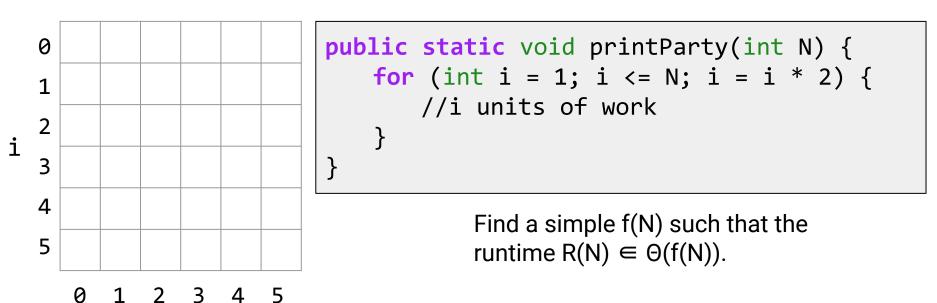


Cost model C(N):

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

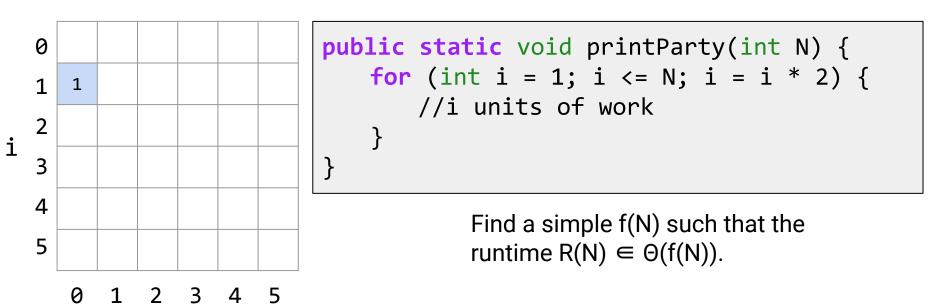
0

 \odot



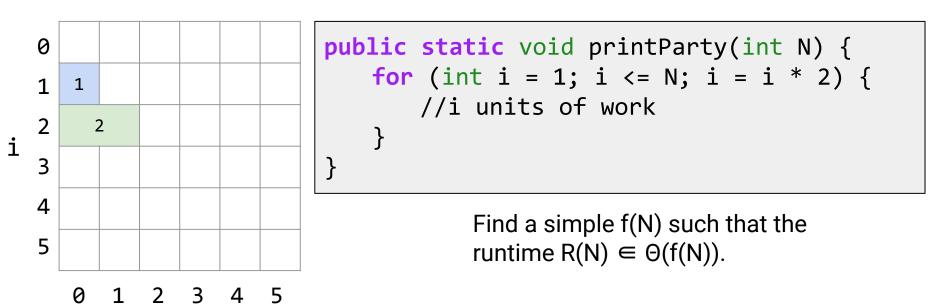
Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

 \odot

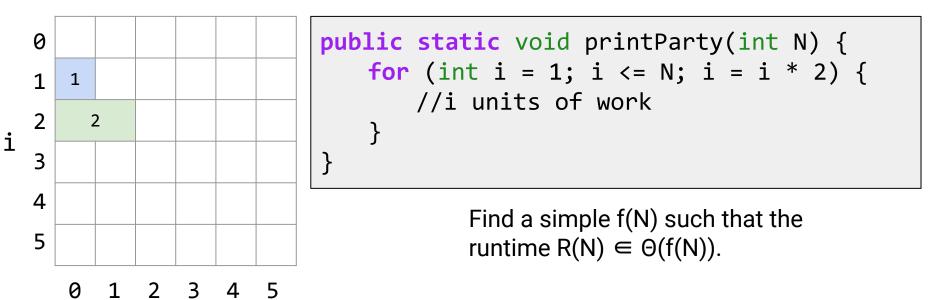


Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1																	

 \odot



Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3																

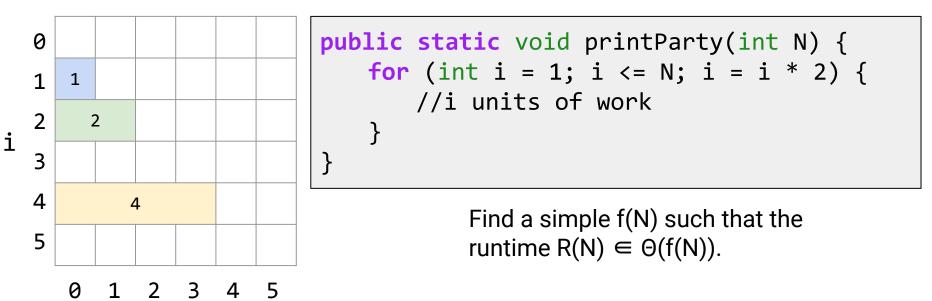


Work Cost model C(N):

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3															

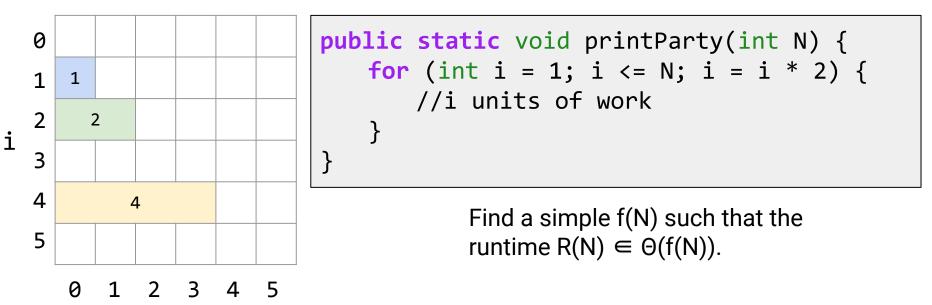
N=3 doesn't do anything extra

@0\$0



Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3	7														

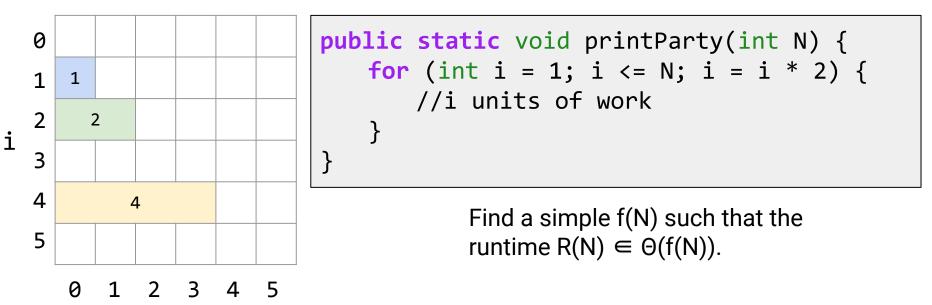
Work



Cost model C(N):

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3	7	7	7	7											
080				N	l=4,5,0	6,7 all	print	7 time	es									

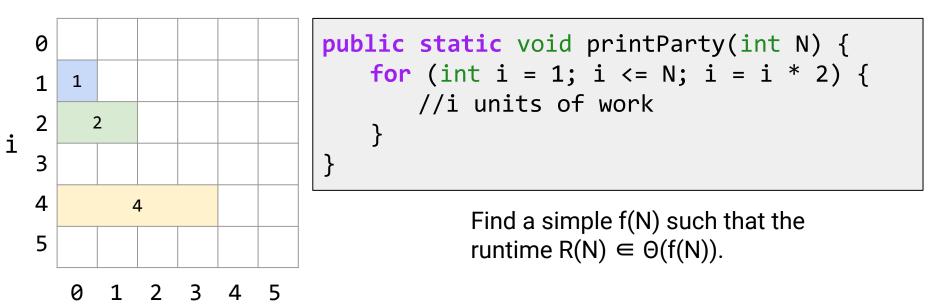
Work



Cost model C(N):

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3	7	7	7	7											
080				N	l=4,5,0	6,7 all	print	7 time	es									

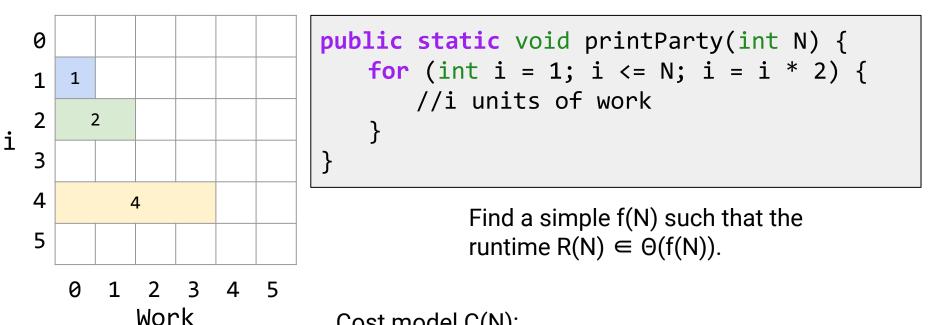
@080



Work Cost model C(N):

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3	7	7	7	7	15	15	15	15	15	15	15	15			

These N all print 15 times



Cost model C(N):

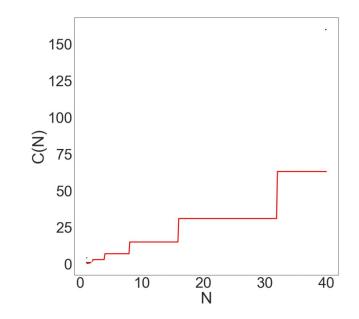
Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3	7	7	7	7	15	15	15	15	15	15	15	15	31	31	31

C(N) = 1 + 2 + 4 + ... + N, if N is a power of 2



Loops Example 2: Prelude to Attempt #2

We're trying to find the order of growth of C(N):



Cost model C(N):

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3	7	7	7	7	15	15	15	15	15	15	15	15	31	31	31

C(N) = 1 + 2 + 4 + ... + N, if N is a power of 2



Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

A.1D.N log NB.log NE. N^2 C.NF.Other

public static void printParty(int N) {
 for (int i = 1; i<=N; i = i * 2) {
 for (int j = 0; j < i; j += 1) {
 System.out.println("hello");
 int ZUG = 1 + 1;
 }
}</pre>

Cost model C(N):

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C(N)	1	3	3	7	7	7	7	15	15	15	15	15	15	15	15	31	31	31

C(N) = 1 + 2 + 4 + ... + N, if N is a power of 2



Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

	N	C(N)	0.5 N	2N
	1	1	0.5	2
	4	1 + 2 + 4 = 7	2	8
	7	1 + 2 + 4 = 7	3.5	14
	8	1 + 2 + 4 + 8 = 15	4	16
	27	1 + 2 + 4 + 8 + 16 = 31	13.5	54
	185	+ 64 + 128 = 255	92.5	370
	715	+ 256 + 512 = 1023	357.5	1430
<u>.</u>	\$0 NC 54			

Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

р

Ν	C(N)	0.5 N	2N
1	1	0.5	2
4	7	2	8
7	7	3.5	14
8	15	4	16
27	31	13.5	54
185	255	92.5	370
715	1023	357.5	1430

<pre>public static void printParty(int n) {</pre>
<pre>for (int i = 1; i<=n; i = i * 2) {</pre>
<pre>for (int j = 0; j < i; j += 1) {</pre>
<pre>System.out.println("hello");</pre>
int ZUG = 1 + 1;

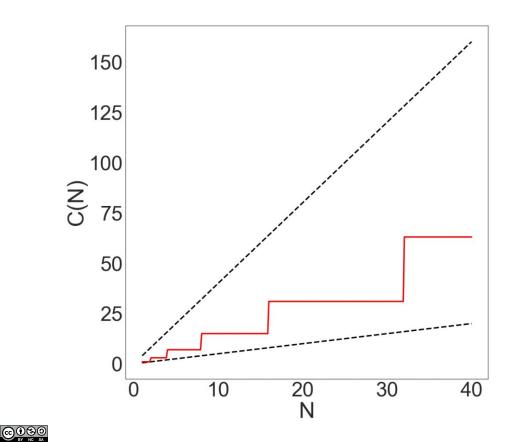
Cost model C(N), println("hello") calls:

• $R(N) = \Theta(1 + 2 + 4 + 8 + ... + N)$ if N is power of 2.

- A. 1 B. log N C. N
- D. N log N
- E. N²
- F. Something else

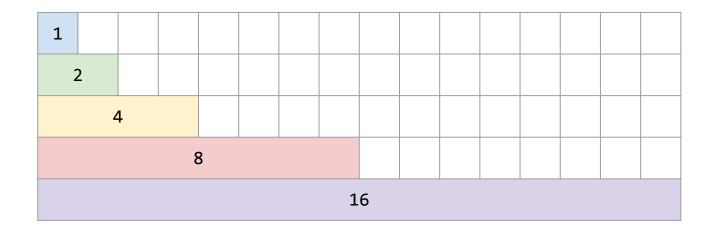
Loops Example 2 [attempt #3] (no yellkey)

Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

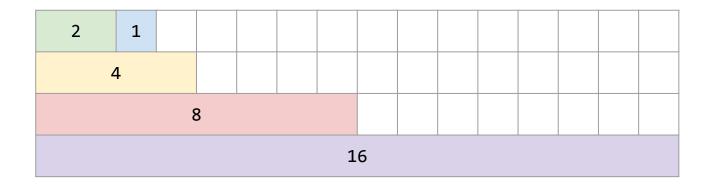


$$R(N) = \Theta(1 + 2 + 4 + 8 + ... + N)$$

= $\Theta(N)$
A. 1 D. N log N
B. log N E. N²
C. N F. Something else









4	2	1							
٤	8								
	1						 		



• 1 + 2 + 4 + ... + N = 2N - 1

 $R(N) = \Theta(1 + 2 + 4 + 8 + ... + N)$ = $\Theta(N)$ A. 1 D. N log N

8	4	2	1	
1	6			



There is No Magic Shortcut for Asymptotic Analysis

Lecture 15, CS61B, Spring 2024

Nested For Loops

There is No Magic Shortcut for Asymptotic Analysis

Amortized Analysis Recursive Analysis Binary Search (Intuitive) Binary Search (Exact) (Bonus Video) Mergesort



Repeat After Me...

There is no magic shortcut for asymptotic analysis problems

- Runtime analysis often requires careful thought.
- CS70 and especially CS170 will cover this in much more detail.
- This is not a math class, though we'll expect you to know these:

 $1 + 2 + 3 + ... + Q = Q(Q+1)/2 = \Theta(Q^2)$ ← Sum of First Natural Numbers (Link) Ο $○ 1^{k}+2^{k}+3^{k}+...+Q^{k}$ (k >=0) = Θ(Q^{k+1}) ← Generalization of ^ ○ $1 + 2 + 4 + 8 + ... + 2^{Q} = 2(2^{Q}) - 1 = Θ(2^{Q}) \leftarrow \text{Sum of First Powers of 2 (Link)}$

- \circ k⁰+k¹+k²+ ... + k^Q (k > 1) = Θ(k^Q) ← Generalization of ^

```
public static void printParty(int n) {
  for (int i = 1; i <= n; i = i * 2) {</pre>
    for (int j = 0; j < i; j += 1) {</pre>
      System.out.println("hello");
      int ZUG = 1 + 1;
```



Repeat After Me...

There is no magic shortcut for asymptotic analysis problems

- Runtime analysis often requires careful thought.
- CS70 and especially CS170 will cover this in much more detail.
- This is not a math class, though we'll expect you to know these:
 - 1+2+3+...+Q = Q(Q+1)/2 = Θ(Q²) ← Sum of First Natural Numbers (Link)
 1^k+2^k+3^k+...+Q^k (k >=0) = Θ(Q^{k+1}) ← Generalization of ^
 1+2+4+8+...+N = 2(N) 1 = Θ(N) ← Sum of First Powers of 2 (Link)
 k⁰+k¹+k²+...+N (k > 1) = Θ(N) ← Generalization of ^



Repeat After Me...

There is no magic shortcut for asymptotic analysis problems (well... usually)

- Runtime analysis often requires careful thought.
- CS70 and especially CS170 will cover this in much more detail.
- This is not a math class, though we'll expect you to know these:
 - $\begin{array}{ll} \circ & 1^{k} + 2^{k} + 3^{k} + \dots + Q^{k} & (k \ge 0) & = \Theta(Q^{k+1}) \\ \circ & k^{0} + k^{1} + k^{2} + \dots + k^{Q} & (k \ge 1) & = \Theta(k^{Q}) \end{array}$
- Strategies:
 - Find exact sum.
 - Write out examples.
 - Draw pictures.

QR decomposition runtime, from "Numerical Linear Algebra" by Trefethen.

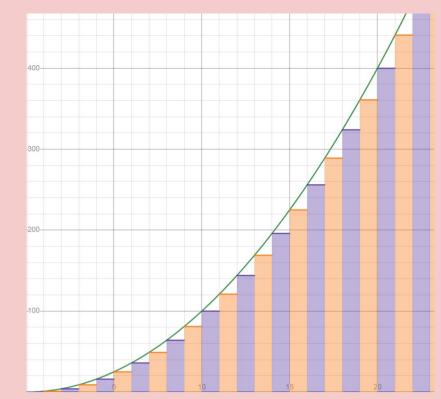


Generalizing the two cases you need to know (Out of Scope)

- $1^{k}+2^{k}+3^{k}+...+Q^{k}$ $(k \ge 0) = \Theta(Q^{k+1})$ $k^{0}+k^{1}+k^{2}+...+k^{Q}$ $(k \ge 1) = \Theta(k^{Q})$

f(0)+f(1)+f(2)+...+f(N) (f continuous)

$$\int_0^N f(x) dx \approx$$





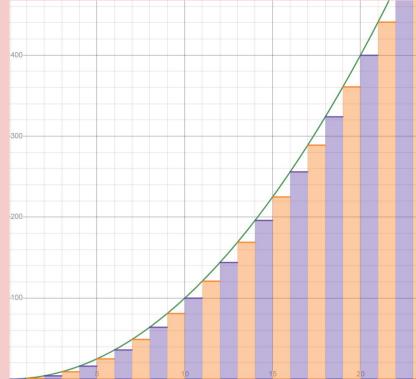
Generalizing the two cases you need to know (Out of Scope)

- $1^{k}+2^{k}+3^{k}+...+Q^{k}$ $(k \ge 0) = \Theta(Q^{k+1})$ $k^{0}+k^{1}+k^{2}+...+k^{Q}$ $(k \ge 1) = \Theta(k^{Q})$
- $f(0)+f(1)+f(2)+...+f(N) = \Theta(integral of f(x) from 0 to N)$

Some shortcuts exist

But they tend to be limited in scope

$$\int_0^N f(x) dx \approx$$





Amortized Analysis

Lecture 15, CS61B, Spring 2024

Nested For Loops There is No Magic Shortcut for Asymptotic Analysis **Amortized Analysis Recursive Analysis** Binary Search (Intuitive) Binary Search (Exact) (Bonus Video) Mergesort



Let's play with this function a bit

```
public static void printParty(int n) {
   for (int i = 1; i <= n; i = i * 2) {</pre>
       //i units of work
    }
}
```



Add $\Theta(N)$ work: Total runtime is still $\Theta(N)$

```
public static void printParty(int n) {
    for (int i = 1; i <= n; i = i * 2) {</pre>
       //i units of work
    for (int i = 1; i <= n; i++) {</pre>
       //1 unit of work
    }
}
```



Surely no function does this, right?

Combine the for loops. No asymptotic change in work done

```
public static void printParty(int n) {
   for (int i = 1; i <= n; i++) {</pre>
       //1 unit of work
       if(i is a power of 2) {
            //i units of work
        }
}
```



Put things in separate functions. Changes an O(1) thing to more O(1) things, so no runtime difference.

```
public static void printParty(int n) {
   for (int i = 1; i <= n; i++) {</pre>
       foo();
    } }
public static void foo() {
   if(i is a power of 2) {
            bar(i);
   // 1 unit of work
}
public static void bar(i) {
   //i units of work
}
```



Rename functions and define what we do in the commented code. No change in runtime

```
public void addMany(int n) {
   for (int i = 0; i < n; i++) {</pre>
       addLast(1);
    } }
public void addLast(int value) {
   if(this.length == arr.length) {
       resize(this.length * 2);
   } //Happens every time length is 2^{k}
   //addLast code takes 1 unit of work
}
public void resize(int i) {
   //resizing takes i units of work
}
```



When we discussed ArrayLists, we handwaved why geometric resizing is better than linear resizing. Now that we know asymptotics, we can finally prove this.

```
public void addLast(int x) {
    if (size == items.length) {
        resize(size + RFACTOR);
    }
    items[size] = x;
    size += 1;
}
```

After N addLasts, runtime is $\Theta(N)$

```
public void addLast(int x) {
    if (size == items.length) {
        resize(size * RFACTOR);
    }
    items[size] = x;
    size += 1;
}
```



Why geometric resizing is faster

Even though the worst-case resize is still $\Theta(N)$, they happen so infrequently with geometric resizing that we get $\Theta(1)$ runtime on average *regardless of how we order List operations*.

```
public void addLast(int x) {
    if (size == items.length) {
        resize(size + RFACTOR);
    }
    items[size] = x;
    size += 1;
}
```

Each addLast takes on average Θ (1) time

```
Each addLast takes on
average Θ(N) time
public void addLast(int x) {
   if (size == items.length) {
      resize(size * RFACTOR);
   }
   items[size] = x;
```

```
size += 1;
```



This is known as Amortized Runtime

- Any single operation may take longer, but if we use it over many operations, we're guaranteed to have a better average performance
- So amortized runtime gives a better estimate of how much time it takes to use something in practice
- Disjoint sets also used amortized runtime; WQU with path compression still has $\Theta(\log(n))$ runtime in the worst case, but $\Theta(\alpha(n))$ amortized runtime.

addLast is $\Theta(1)$ amortized

```
public void addLast(int x) {
    if (size == items.length) {
        resize(size * RFACTOR);
    }
    items[size] = x;
    size += 1;
}
```



Recursive Analysis

Lecture 15, CS61B, Spring 2024

Nested For Loops There is No Magic Shortcut for Asymptotic Analysis Amortized Analysis **Recursive Analysis** Binary Search (Intuitive)

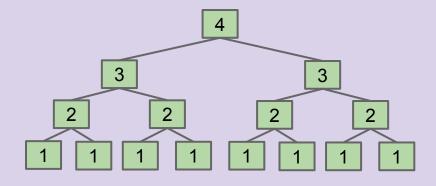
Binary Search (Exact) (Bonus Video)

Mergesort

```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

Using your intuition, give the order of growth of the runtime of this code as a function of N?

- A. 1
- B. log N
- C. N
- D. N²
- E. 2^N

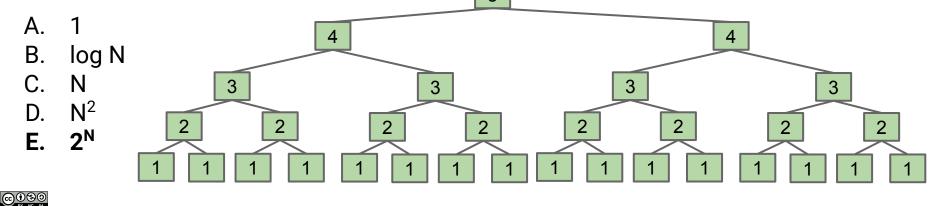




```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

2^N: Every time we increase N by 1, we double the work!

Using your intuition, give the order of growth of the runtime of this code as a function of N?

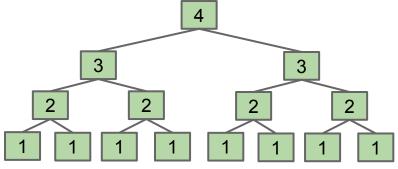


```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

Another approach: Count number of calls to f3, given by C(N).

Each function call does a constant amount of work (not counting recursive calls), so $C(N) \in \Theta(R(N))$

- C(1) = 1
- C(2) = 1 + 2
- C(3) = 1 + 2 + 4





Recursion, Approach 2: Exact Counting: http://yellkey.com/happy

Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

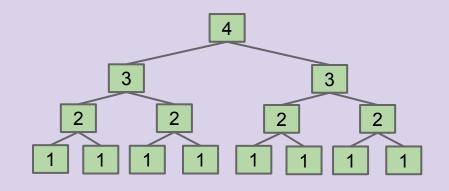
```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

Another approach: Count number of calls to f3, given by C(N).

- C(3) = 1 + 2 + 4
- C(N) = 1 + 2 + 4 + ... + ???

What is the final term of the sum?

A.ND. 2^{N-1} B. 2^{N} E. 2^{N-1} C. 2^{N} 2^{N-1} 2^{N-1}





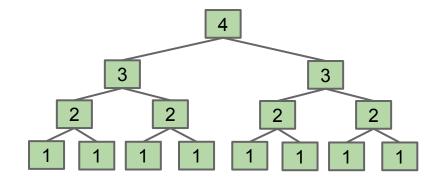
```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

Another approach: Count number of calls to f3, given by C(N).

- C(3) = 1 + 2 + 4
- C(N) = 1 + 2 + 4 + ... + ???

What is the final term of the sum?

D.
$$2^{N-1}$$



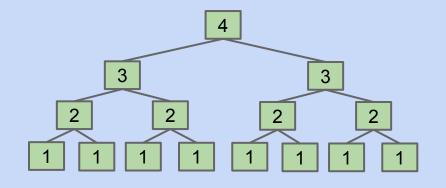


```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

Another approach: Count number of calls to f3, given by C(N).

• $C(N) = 1 + 2 + 4 + ... + 2^{N-1}$

Give a simple expression for C(N).





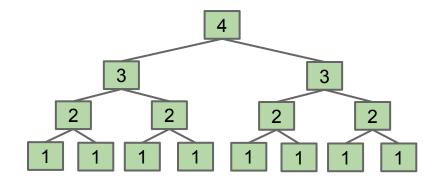
```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

Another approach: Count number of calls to f3, given by C(N).

• $C(N) = 1 + 2 + 4 + ... + 2^{N-1}$

Give a simple expression for C(N).

- $C(N) = 2(2^{N-1}) 1$
- $C(N) = 2^{N}-1$





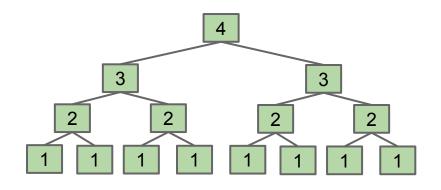
```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1);
}</pre>
```

Another approach: Count number of calls to f3, given by C(N).

- $C(N) = 1 + 2 + 4 + ... + 2^{N-1}$
- Solving, we get $C(N) = 2^{N} 1$

Since work during each call is constant:

• $R(N) = \Theta(2^N)$



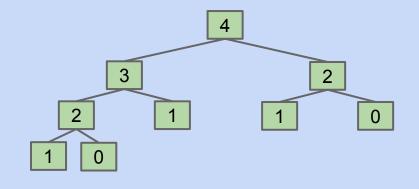


```
public static int fib(int n) {
    if (n <= 1)
        return 1;
    return fib(n-1) + fib(n-2);
}</pre>
```

What happens if we make this tiny change?

- C(0) = 1
- C(1) = 1
- C(2) = 1+1+1
- C(3) = 1+1+2+1
- ?????

Give a simple expression for C(N).



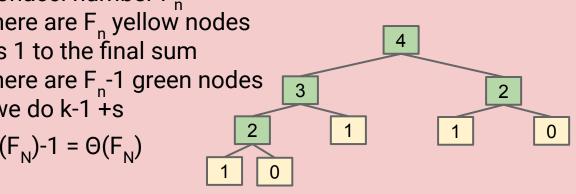


```
public static int fib(int n) {
    if (n <= 1)
        return 1;
    return fib(n-1) + fib(n-2);
}</pre>
```

For this one, we'll have to be a bit more creative:

- fib(n) returns the nth Fibonacci number F_n
- In the tree on the right, there are F_n yellow nodes
 - Why? Each leaf adds 1 to the final sum
- In the tree on the right, there are F_n -1 green nodes
 - Why? To sum k 1s, we do k-1 +s

C(N) = # yellow + # green = $2(F_N)$ -1 = $\Theta(F_N)$





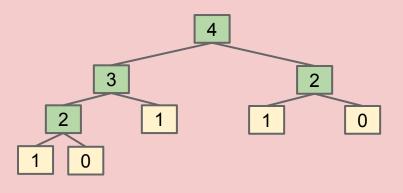
```
public static int fib(int n) {
    if (n <= 1)
        return 1;
    return fib(n-1) + fib(n-2);
}</pre>
```

$$\begin{split} & C(N) = \# \text{ yellow} + \# \text{ green} = 2(F_N) - 1 = \Theta(F_N) \\ & \text{If you do enough math, you find that } F_N \Subset \Theta(\phi^N) \\ & \text{Where } \phi = (1 + \text{sqrt}(5))/2 \approx 1.618 \\ & \text{Each function call does 1 unit of work} \end{split}$$

So $R(N) = \Theta(1.618^{N})$

In conclusion: There is no Magic Shortcut for

Asymptotic Analysis



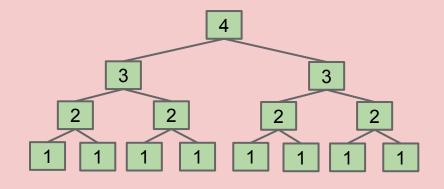
Recursion, Approach 3: Recurrence Relations (Out of Scope for 61B)

Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1)
}</pre>
```

Third approach: Count number of calls to f3, given by a "recurrence relation" for C(N).

- C(1) = 1
- C(N) = 2C(N-1) + 1





Recursion, Approach 3: Recurrence Relations (Out of Scope for 61B)

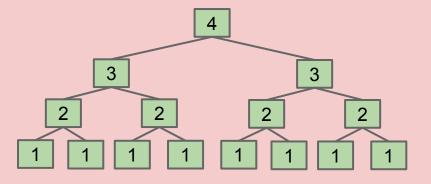
Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1)
}</pre>
```

Third approach: Count number of calls to f3, given by a "recurrence relation" for C(N).

- C(1) = 1
- C(N) = 2C(N-1) + 1

More technical to solve. Won't do this in our course. See next slide for solution.





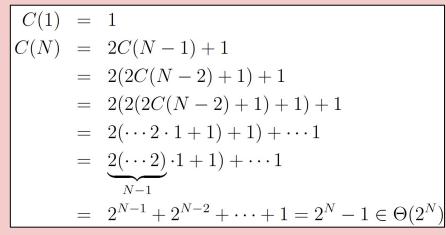
Recursion, Approach 3: Recurrence Relations (Out of Scope for 61B)

Find a simple f(N) such that the runtime $R(N) \in \Theta(f(N))$.

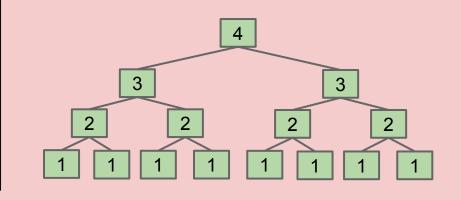
```
public static int f3(int n) {
    if (n <= 1)
        return 1;
    return f3(n-1) + f3(n-1)
}</pre>
```

This approach not covered in class. Provided for those of you who want to see a recurrence relation solution.

Third approach: Count number of calls to f3, given by a "recurrence relation" for C(N).



@080



Binary Search (Intuitive)

Lecture 15, CS61B, Spring 2024

Nested For Loops There is No Magic Shortcut for Asymptotic Analysis Amortized Analysis Recursive Analysis

Binary Search (Intuitive)

Binary Search (Exact) (Bonus Video) Mergesort



Binary Search (demo: <u>https://goo.gl/3VvJNw</u>)

Trivial to implement?

- Idea published in 1946.
- First correct implementation in 1962.
 Bug in Java's binary search discovered in 2006.
 <u>http://goo.gl/gQI0FN</u>

See Jon Bentley's book Programming Pearls.

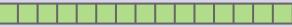
```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

Binary Search (Intuitive): http://yellkey.com/beyond

```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
}
```

Goal: Find runtime in terms of N = hi - lo + 1 [i.e. # of items being considered]

- Intuitively, what is the order of growth of the worst case runtime?
 - A. 1
 - B. $\log_2 N$
 - C. N
 - D. $N \log_2 N$
 - E. 2^N



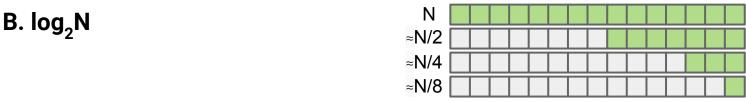


Binary Search (Intuitive)

```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
}
```

Goal: Find runtime in terms of N = hi - lo + 1 [i.e. # of items being considered]

• Intuitively, what is the order of growth of the worst case runtime?



Why? Problem size halves over and over until it gets down to 1.

If C is number of calls to binarySearch, solve for $1 = N/2^{C} \rightarrow C = \log_{2}(N)$

Log Time Is Really Terribly Fast

In practice, logarithmic time algorithms have almost constant runtimes.

• Even for incredibly huge datasets, practically equivalent to constant time.

Ν	log ₂ N	Typical runtime (seconds)
100	6.6	1 nanosecond
100,000	16.6	2.5 nanoseconds
100,000,000	26.5	4 nanoseconds
100,000,000,000	36.5	5.5 nanoseconds
100,000,000,000,000	46.5	7 nanoseconds



Binary Search Exact (Bonus)

Lecture 15, CS61B, Spring 2024

Nested For Loops There is No Magic Shortcut for Asymptotic Analysis Amortized Analysis **Recursive Analysis Binary Search Intuitive Binary Search Exact (Bonus)** Mergesort



This section is available as a pre-recorded video.

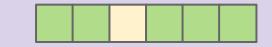
It's not "out of scope" since it's just another example problem using the same techniques used throughout the lecture.



```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

Each call does constant work (w/o recursive calls)

N=6



• What is C(6), number of total calls for N = 6?

C.
$$\log_2(6) = 2.568$$



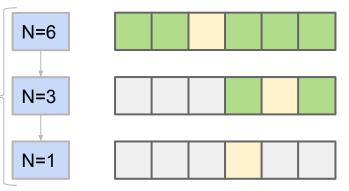
```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

3 calls

Each call does constant work (w/o recursive calls)

- What is C(6), number of total calls for N = 6?
 - B. 3

Three total calls, where N = 6, N = 3, and N = 1.





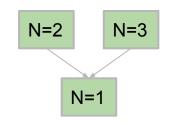
```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13
C(N)	1					3							



```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

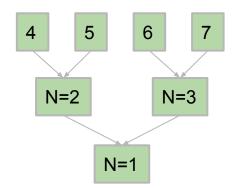
Ν	1	2	3	4	5	6	7	8	9	10	11	12	13
C(N)	1	2	2			3							





```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

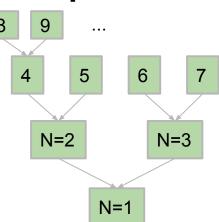
Ν	1	2	3	4	5	6	7	8	9	10	11	12	13
C(N)	1	2	2	3	3	3	3						





```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13
C(N)	1	2	2	3	3	3	3	4	4	4	4	4	4

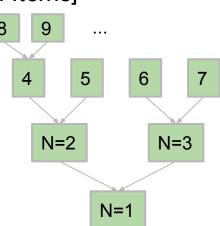




```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

Ν	1	2	3	4	5	6	7	8	9	10	11	12	13
C(N)	1	2	2	3	3	3	3	4	4	4	4	4	4

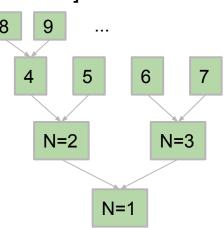
$$C(N) = Llog_{2}(N) \rfloor + 1$$





```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

- Number of binarySearch calls.
- $C(N) = Llog_2(N) \rfloor + 1$
- Since each call takes constant time, $R(N) = \Theta(Llog_2(N)J)$
 - This f(N) is way too complicated. Let's simplify.





Goal: Simplify $\Theta(Llog_2(N)J)$

For proof: See online textbook exercises.

- Three handy properties to help us simplify:
 - $Lf(N) = \Theta(f(N))$ [the floor of f has same order of growth as f]
 - $\Gamma f(N) = \Theta(f(N))$ [the ceiling of f has same order of growth as f]
 - $\circ \log_{P}(N) = \Theta(\log_{Q}(N))$ [logarithm base does not affect order of growth]

 $L\log_2(N) \rfloor = \Theta(\log N)$

Since base is irrelevant, we omit from our big theta expression. We also omit the parenthesis around N for aesthetic reasons.

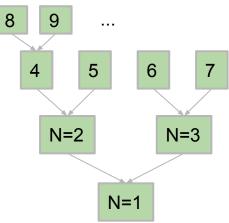


```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

Goal: Find worst case runtime in terms of N = hi - lo + 1 [i.e. # of items]

- Number of binarySearch calls.
- $C(N) = Llog_2(N) \downarrow +1 = \Theta(log N)$
- Since each call takes constant time, R(N) = Θ(log N)

... and we're done!





```
static int binarySearch(String[] sorted, String x, int lo, int hi)
    if (lo > hi) return -1;
    int m = (lo + hi) / 2;
    int cmp = x.compareTo(sorted[m]);
    if (cmp < 0) return binarySearch(sorted, x, lo, m - 1);
    else if (cmp > 0) return binarySearch(sorted, x, m + 1, hi);
    else return m;
```

Approach: Measure number of string comparisons for N = hi - lo + 1.

- C(0) = 0
- C(1) = 1
- C(N) = 1 + C((N-1)/2)

Can show that $C(N) = \Theta(\log N)$. Beyond scope of class, so won't solve in slides.



Mergesort

Lecture 15, CS61B, Spring 2024

Nested For Loops There is No Magic Shortcut for Asymptotic Analysis Amortized Analysis **Recursive Analysis Binary Search Intuitive** Binary Search Exact (Bonus) Mergesort

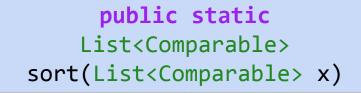


Sorting

Along with matrix multiplication, sorting is one of the problems that pops up most often in asymptotic analysis.

- Given a list of Comparables, return them in sorted order
 - Assumes the comparison method is transitive, reflexive, symmetric (see Lecture 12), and runs in $\Theta(1)$ time.

x = List.of({"he", "is", "the", "agoyatis", "of", "mr.", "conchis"})



x = { "agoyatis", "conchis", "he", "is", "mr.", "of", "the" }



Mergesort is a recursive way to sort a list:

- Split the list into two parts
- Sort the two lists individually
- Merge the two lists together

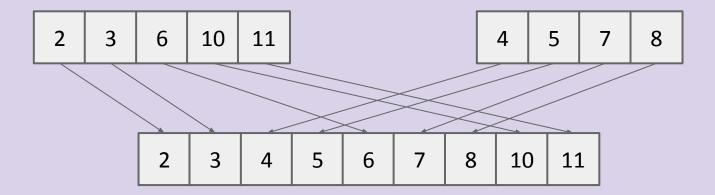
```
static List<Comparable> sort(List<Comparable> x)
List<Comparable> firsthalf = sort(x.sublist(0, x.size()/2));
List<Comparable> secondhalf = sort(x.sublist(x.size()/2, x.size()));
return merge(firsthalf, secondhalf);
```

The Merge Operation

Given two sorted arrays, the merge operation combines them into a single sorted array by successively copying the smallest item from the two arrays into a target array.

Merging Demo (Link)

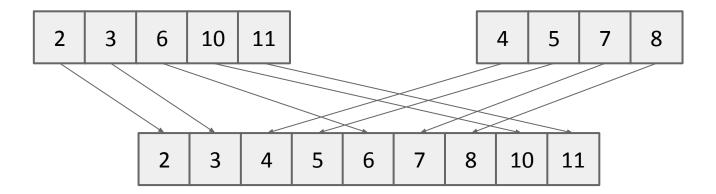




How does the runtime of merge grow with N, the total number of items?

- A. $\Theta(1)$ C. $\Theta(N)$
- B. $\Theta(\log N)$ D. $\Theta(N^2)$





How does the runtime of merge grow with N, the total number of items?

C. $\Theta(N)$. Why? $\Theta(1)$ time per element in the merged list, and the merged list has exactly N items



```
static List<Comparable> sort(List<Comparable> x)
  List<Comparable> firsthalf = sort(x.sublist(0, x.size()/2));
  List<Comparable> secondhalf = sort(x.sublist(x.size()/2, x.size()));
  return merge(firsthalf, secondhalf);
```

Since we don't care about the list itself, let's simplify our code a bit

- Our runtime should be in terms of x.size(), so let's let int n = x.size()
- merge takes Θ(n) time, so let's replace that with "n units of work"



Determining Mergesort Runtime

```
static void sortRuntime(int n)
    sortRuntime(n/2);
    sortRuntime(n/2);
    //n units of work
}
```

Since we don't care about the list itself, let's simplify our code a bit

- Our runtime should be in terms of x.size(), so let's let int n = x.size()
- merge takes Θ(n) time, so let's replace that with "n units of work"

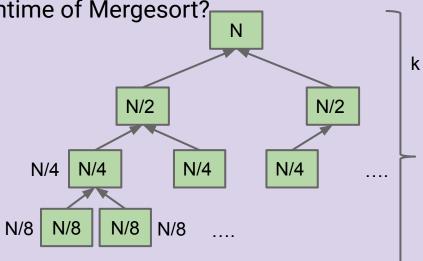


Example 5: Mergesort Order of Growth, yellkey.com/consider

```
static void sortRuntime(int n)
    sortRuntime(n/2);
    sortRuntime(n/2);
    //n units of work
}
```

For an array of size N, what is the worst case runtime of Mergesort?

- A. Θ(1)
- B. $\Theta(\log N)$
- C. $\Theta(N)$
- D. $\Theta(N \log N)$
- E. $\Theta(N^2)$





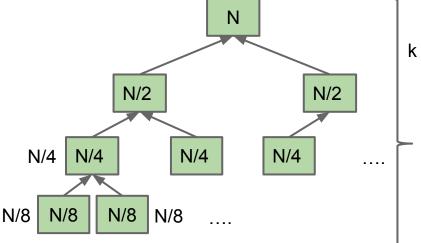
Example 5: Mergesort Order of Growth

Mergesort has worst case runtime = $\Theta(N \log N)$.

- Every level has N units of work.
 - Top level takes N units of work.
 - Next level takes N/2 + N/2 = N units of work.
 - One more level down: N/4 + N/4 + N/4 + N/4 = N.
- Thus, total runtime is Nk, where k is the number of levels.
 - How many levels? Goes until we get to size 1.
 - $k = \log_2(N)$.
- Overall runtime is $\Theta(N \log N)$.

Exact count explanation is tedious.

• Omitted here. See textbook exercises.





C(N): Number of calls to mergesort + number of array writes.

$$C(N) = \begin{cases} 1 & : N < 2\\ 2C(N/2) + N & : N \ge 2 \end{cases}$$

$$C(N) = 2(2C(N/4) + N/2) + N$$

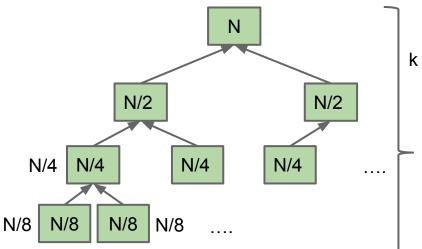
$$= 4C(N/4) + N + N$$

$$= 8C(N/8) + N + N + N$$

$$= N \cdot 1 + \underbrace{N + N + \dots + N}_{k = \lg N}$$

$$= N + N \lg N \in \Theta(N \lg N)$$

Only works for N=2^k. Can be generalized at the expense of some tedium by separately finding Big O and Big Omega bounds (see next lecture).





Using Sorting as a Tool

Recall from Lecture 13 the dup functions, which checked if a **sorted** array contained any duplicates.

• dup1 took $\Theta(N^2)$ runtime, while dup2 took $\Theta(N)$ runtime

```
public static boolean dup1(int[] A) {
   for (int i = 0; i < A.length; i += 1) {</pre>
     for (int j = i + 1; j < A.length; j += 1) {</pre>
                                                                               dup2
       if (A[i] == A[j]) {
                                 public static boolean dup2(int[] A) {
          return true;
                                   for (int i = 0; i < A.length - 1; i += 1) {</pre>
                                      if (A[i] == A[i + 1]) {
                                        return true;
   return false;
}
                                   return false;
dup1
                                 }
```



Using Sorting as a Tool

 $\bigcirc 0$

- dup1 still works normally, but it takes $\Theta(N^2)$ runtime
- dup2 no longer works... Can we fix it?

```
public static boolean dup1(int[] A) {
   for (int i = 0; i < A.length; i += 1) {</pre>
     for (int j = i + 1; j < A.length; j += 1) {</pre>
                                                                              dup2
       if (A[i] == A[j]) {
                                 public static boolean dup2(int[] A) {
          return true;
                                   for (int i = 0; i < A.length - 1; i += 1) {</pre>
                                      if (A[i] == A[i + 1]) {
                                        return true;
   return false;
                                   return false;
dup1
                                 }
```

- Solution: Sort A first!
- What's our new runtime?
 - Sorting took $\Theta(N \log N)$ time, the rest of dup2 took $\Theta(N)$ time

```
public static boolean dup1(int[] A) {
     for (int i = 0; i < A.length; i += 1) {</pre>
                                                                                 dup2
       for (int j = i + 1; j < A.length; j += 1) {</pre>
         if (A[i] == A[j]) {
                                    public static boolean dup2(int[] A) {
             return true;
                                      A = A.sort()
                                      for (int i = 0; i < A.length - 1; i += 1) {</pre>
                                        if (A[i] == A[i + 1]) \{
                                           return true;
     return false;
   }
                                      return false;
  dup1
@080
```

Using Sorting as a Tool

- dup1 still works normally, but it takes $\Theta(N^2)$ runtime
- If we use sort as a black box, we can modify dup2 so it runs in Θ(N log N) time!
- Can we do better? Yes, we can get Θ(N)... once we get to hashing

```
public static boolean dup1(int[] A) {
     for (int i = 0; i < A.length; i += 1) {</pre>
                                                                                 dup2
       for (int j = i + 1; j < A.length; j += 1) {</pre>
         if (A[i] == A[j]) {
                                    public static boolean dup2(int[] A) {
             return true;
                                      A = A.sort()
                                      for (int i = 0; i < A.length - 1; i += 1) {</pre>
                                        if (A[i] == A[i + 1]) {
                                          return true;
     return false;
   }
                                      return false;
  dup1
@080
```

Linear vs. Linearithmic (N log N) vs. Quadratic

N log N is basically as good as N, and is vastly better than N^2 .

• For N = 1,000,000, the log N is only 20.

	n	$n \log_2 n$	n ²	n ³	1.5 ⁿ	2 ⁿ	n!
n = 10	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	4 sec
n = 30	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	18 min	10 ²⁵ years
n = 50	< 1 sec	< 1 sec	< 1 sec	< 1 sec	11 min	36 years	very long
n = 100	< 1 sec	< 1 sec	< 1 sec	1 sec	12,892 years	1017 years	very long
<i>n</i> = 1,000	< 1 sec	< 1 sec	1 sec	18 min	very long	very long	very long
n = 10,000	< 1 sec	< 1 sec	2 min	12 days	very long	very long	very long
n = 100,000	< 1 sec	2 sec	3 hours	32 years	very long	very long	very long
n = 1,000,000	1 sec	20 sec	12 days	31,710 years	very long	very long	very long

Table 2.1 The running times (rounded up) of different algorithms on inputs of increasing size, for a processor performing a million high-level instructions per second. In cases where the running time exceeds 10²⁵ years, we simply record the algorithm as taking a very long time.

0000

(from Algorithm Design: Tardos, Kleinberg)

Summary

Theoretical analysis of algorithm performance requires *careful thought*.

- There are **no magic shortcuts** for analyzing code.
- In our course, it's OK to do exact counting or intuitive analysis.
 - Know how to sum $1^{k}+2^{k}+...+N^{k}$ and $k^{0}+k^{1}+...+k^{N}$.
 - We won't be writing mathematical proofs in this class.
- Many runtime problems you'll do in this class resemble one of the five problems from today. See textbook, study guide, and discussion for more practice.
- This topic has one of the highest skill ceilings of all topics in the course, and is a modern research topic

Different solutions to the same problem, e.g. sorting, may have different runtimes.

- N² vs. N log N is an enormous difference.
- Going from N log N to N is nice, but not a radical change.

Once you prove runtime for one problem, you may be able to use it in other problems to speed things up!

