



CPTS 223 Advanced Data Structure C/C++

Why Advanced Data Structure

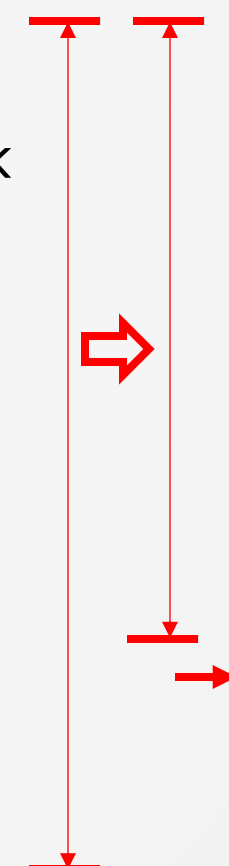
Introduction: why advanced data structure?

First impression in daily life

- Find a correct car wiper blade from a size chart book
- Target: Model **Kb**
- Approaches to locating Make K?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle? binary search
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section

- Which one is the best? Why?



A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Introduction: why advanced data structure?

First impression: case study

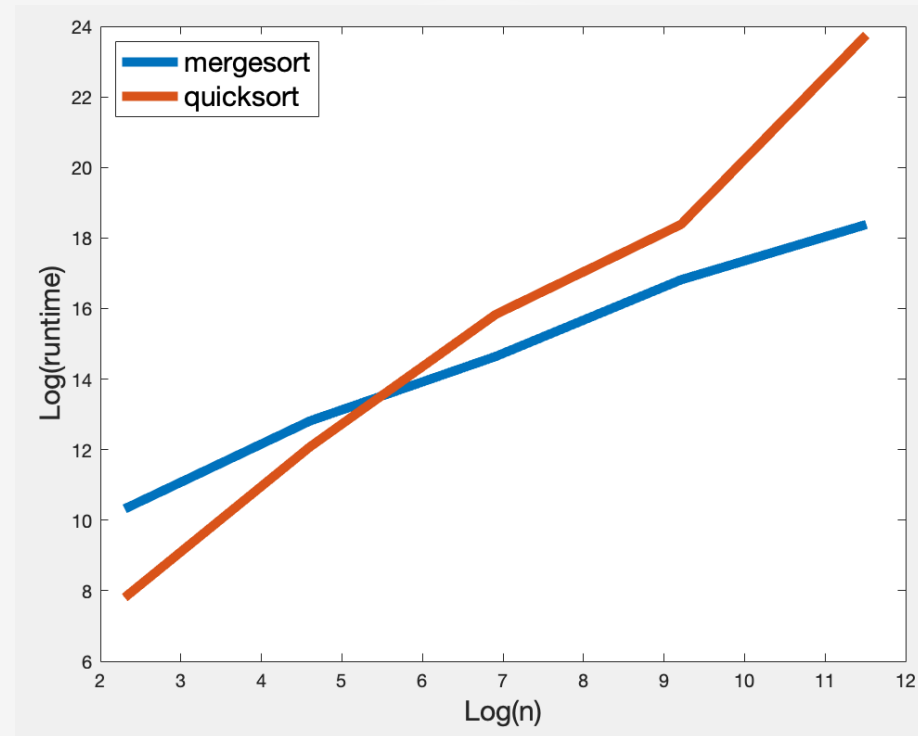
- Case study: comparing **sorting** algorithms:
 - Mergesort V.S. Quicksort (in terms of running time)

Running time (in ns)

n (input size)	Mergesort	Quicksort
10	30375	2458
100	367666	176750
1000	2280125	7493833
10000	20054042	96236458
100000	96236458	20707570875

Introduction: why advanced data structure?

First impression: case study



- Is it possible to have a **comparison benchmark**?

Why advanced data structure?

An analysis

$N=10000, M=100: \text{Prob}=1-(1-1/N)^{100} = 0.01$
 $N=10000, M=1000: \text{Prob}=1-(1-1/N)^{1000} = 0.0952$
 $N=10000, M=10000: \text{Prob}=1-(1-1/N)^{10000} = 0.6421$

$\text{Prob}(\text{find } K \text{ within } M \text{ picks}) = 1 - (1 - 1/N)^M$

$\text{Prob}(\text{find } K \text{ within } 10 \text{ picks}) = 1 - (1 - 1/N)^{10}$

$\text{Prob}(\text{not } K \text{ for } 10 \text{ picks}) = (1 - 1/N)^{10}$

- Find a correct car wiper blade from a size chart book
- Target: Model **Kb**
- Approaches to locating Make **K**?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle?
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section
 - Which one is the best? Why?

Random pick a row?

$\text{Prob}(\text{find } K) = 1/N$

$\text{Prob}(\text{not } K) = 1 - 1/N$

N: # rows

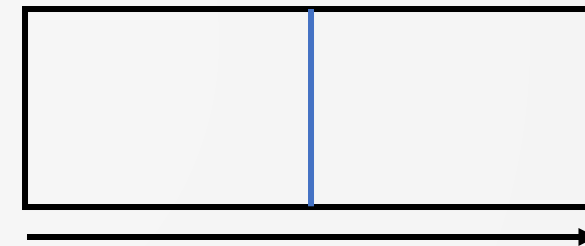
A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis

- Find a correct car wiper blade from a size chart book
- Target: Model **Kb**
- Approaches to locating Make **K**?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle? **binary search**
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section



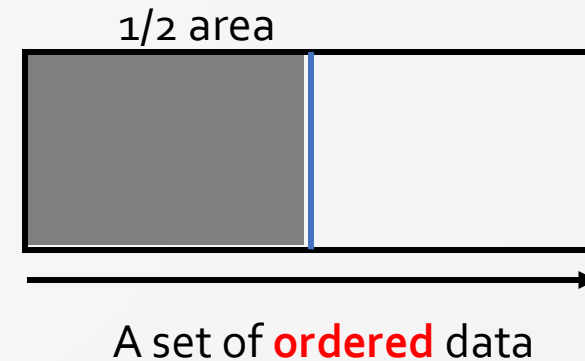
A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis

- Find a correct car wiper blade from a size chart book
- Target: Model **Kb**
- Approaches to locating Make **K**?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle? **binary search**
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section



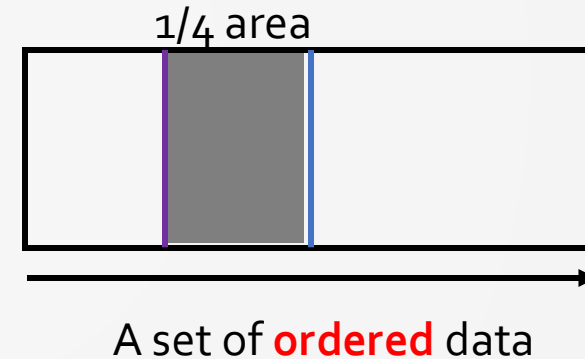
A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis

- Find a correct car wiper blade from a size chart book
- Target: Model **Kb**
- Approaches to locating Make **K**?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle? **binary search**
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section



A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis

- Find a correct car wiper blade from a size chart book
- Target: Model **Kb**
- Approaches to locating Make **K**?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle? binary search
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section



A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis

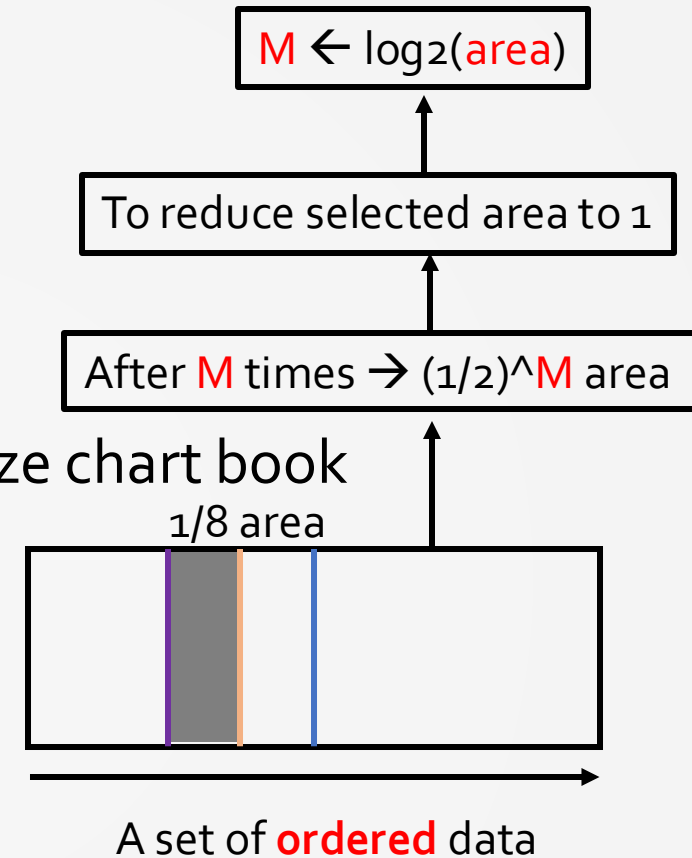
- Find a correct car wiper blade from a size chart book

- Target: Model **Kb**

- Approaches to locating Make **K**?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle? **binary search**
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section

- Which one is the best? Why?



A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis

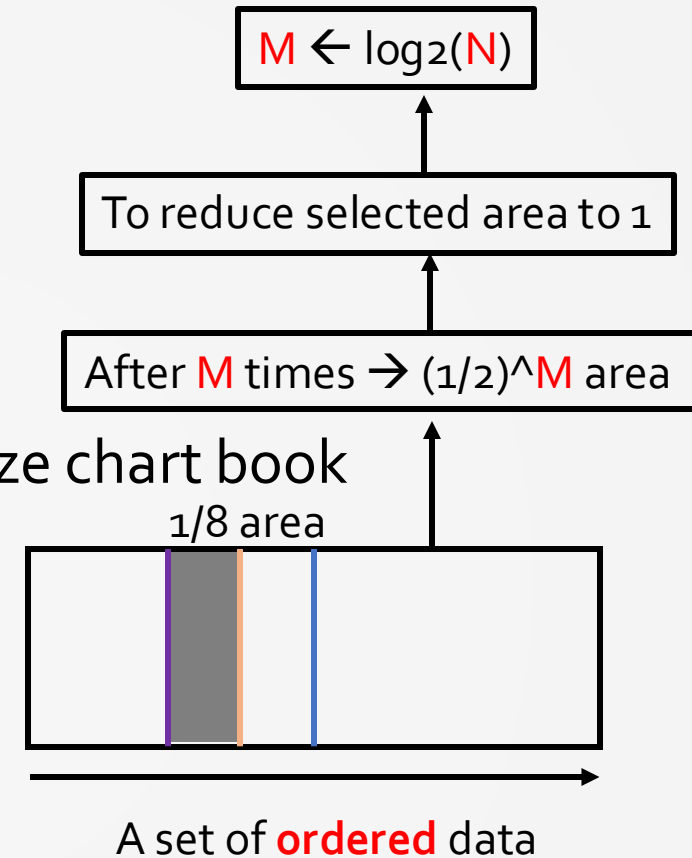
- Find a correct car wiper blade from a size chart book

- Target: Model **Kb**

- Approaches to locating Make **K**?

- Difference? {
- Random pick a row?
 - From 1st to last row?
 - From last to 1st row?
 - Query Makes from middle? binary search
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section

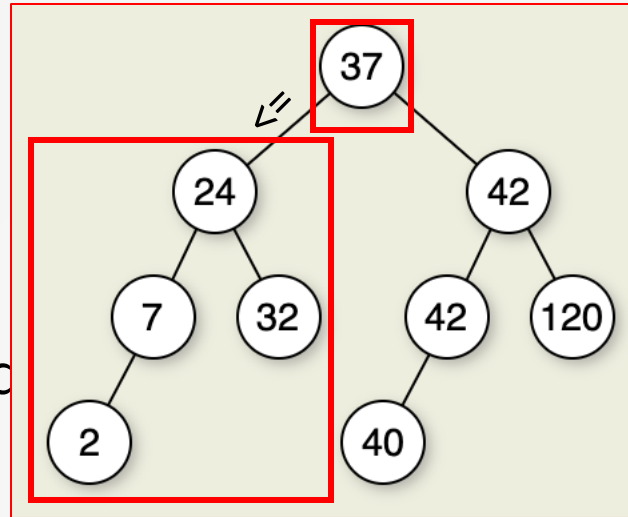
- Which one is the best? Why?



A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis



- Find a correct car wiper blade
- Target: Model **Kb**
- Approaches to locating Make **K**?

Difference?

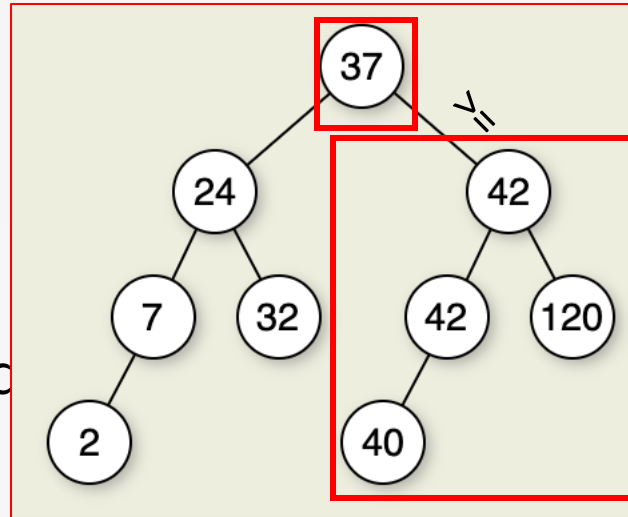
- Random pick a row?
- From 1st to last row?
- From last to 1st row?
- Query Makes from middle? binary search
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section

- Which one is the best? Why?

	A	B
Make	Model	
A	Aa	
A	Ab	
A	Ac	
B	Ba	
B	Bb	
B	Bc	
.		
.		
.		
J	Jc	
K	Ka	
K	Kb	
K	Kc	
L	La	
.		
.		
.		
Z	Za	
Z	Zb	
Z	Zc	

Why advanced data structure?

An analysis



- Find a correct car wiper blade
- Target: Model **Kb**
- Approaches to locating Make **K**?

Difference?

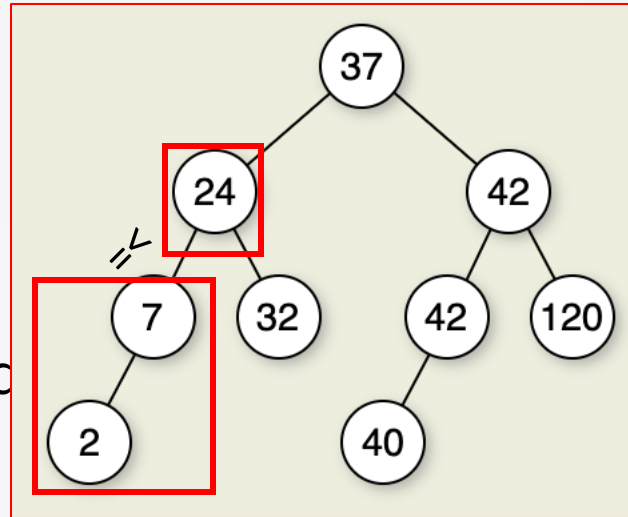
- Random pick a row?
- From 1st to last row?
- From last to 1st row?
- Query Makes from middle? binary search
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section

- Which one is the best? Why?

	A	B
Make	Model	
A	Aa	
A	Ab	
A	Ac	
B	Ba	
B	Bb	
B	Bc	
.		
.		
.		
J	Jc	
K	Ka	
K	Kb	
K	Kc	
L	La	
.		
.		
.		
Z	Za	
Z	Zb	
Z	Zc	

Why advanced data structure?

An analysis



- Find a correct car wiper blade
- Target: Model **Kb**
- Approaches to locating Make **K**?

Difference?

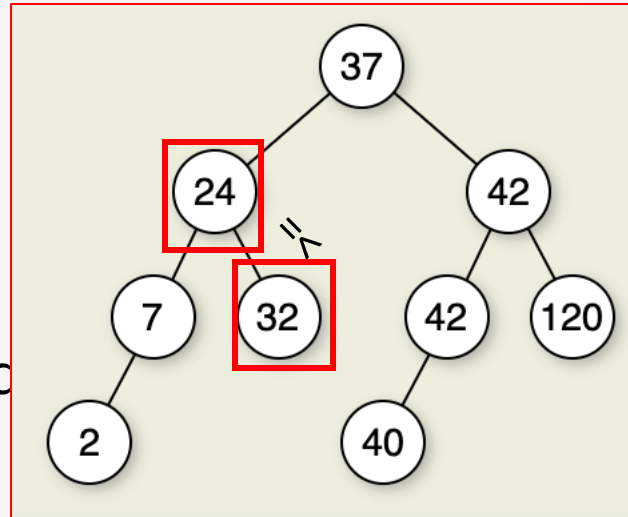
- Random pick a row?
- From 1st to last row?
- From last to 1st row?
- Query Makes from middle? **binary search**
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section

- Which one is the best? Why?

A	B
Make	Model
A	Aa
A	Ab
A	Ac
B	Ba
B	Bb
B	Bc
.	
.	
.	
J	Jc
K	Ka
K	Kb
K	Kc
L	La
.	
.	
.	
Z	Za
Z	Zb
Z	Zc

Why advanced data structure?

An analysis



- Find a correct car wiper blade
- Target: Model **Kb**
- Approaches to locating Make **K**?

Difference?

- Random pick a row?
- From 1st to last row?
- From last to 1st row?
- Query Makes from middle? binary search
 - If queried **Make** < **K** → query next one in latter section
 - If queried **Make** > **K** → query next one in former section

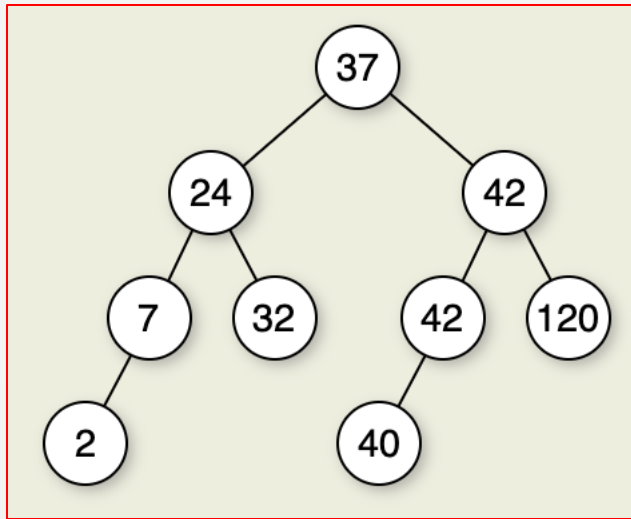
- Which one is the best? Why?

	A	B
Make	Model	
A	Aa	
A	Ab	
A	Ac	
B	Ba	
B	Bb	
B	Bc	
.		
.		
.		
J	Jc	
K	Ka	
K	Kb	
K	Kc	
L	La	
.		
.		
.		
Z	Za	
Z	Zb	
Z	Zc	

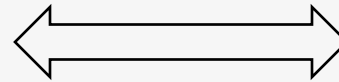
Why advanced data structure?

An analysis

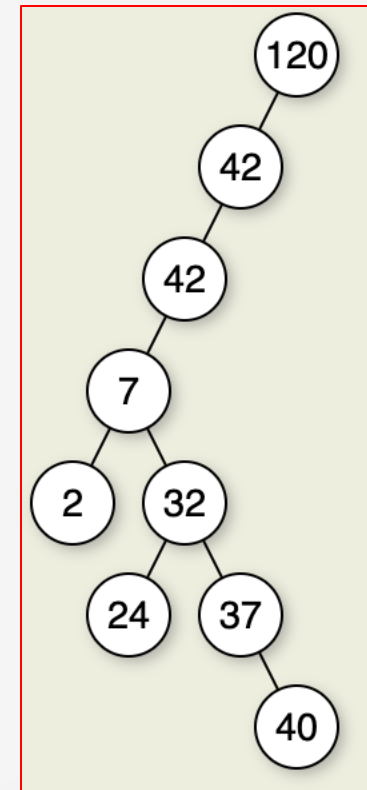
A good shape of BST



$\log_2(N)$



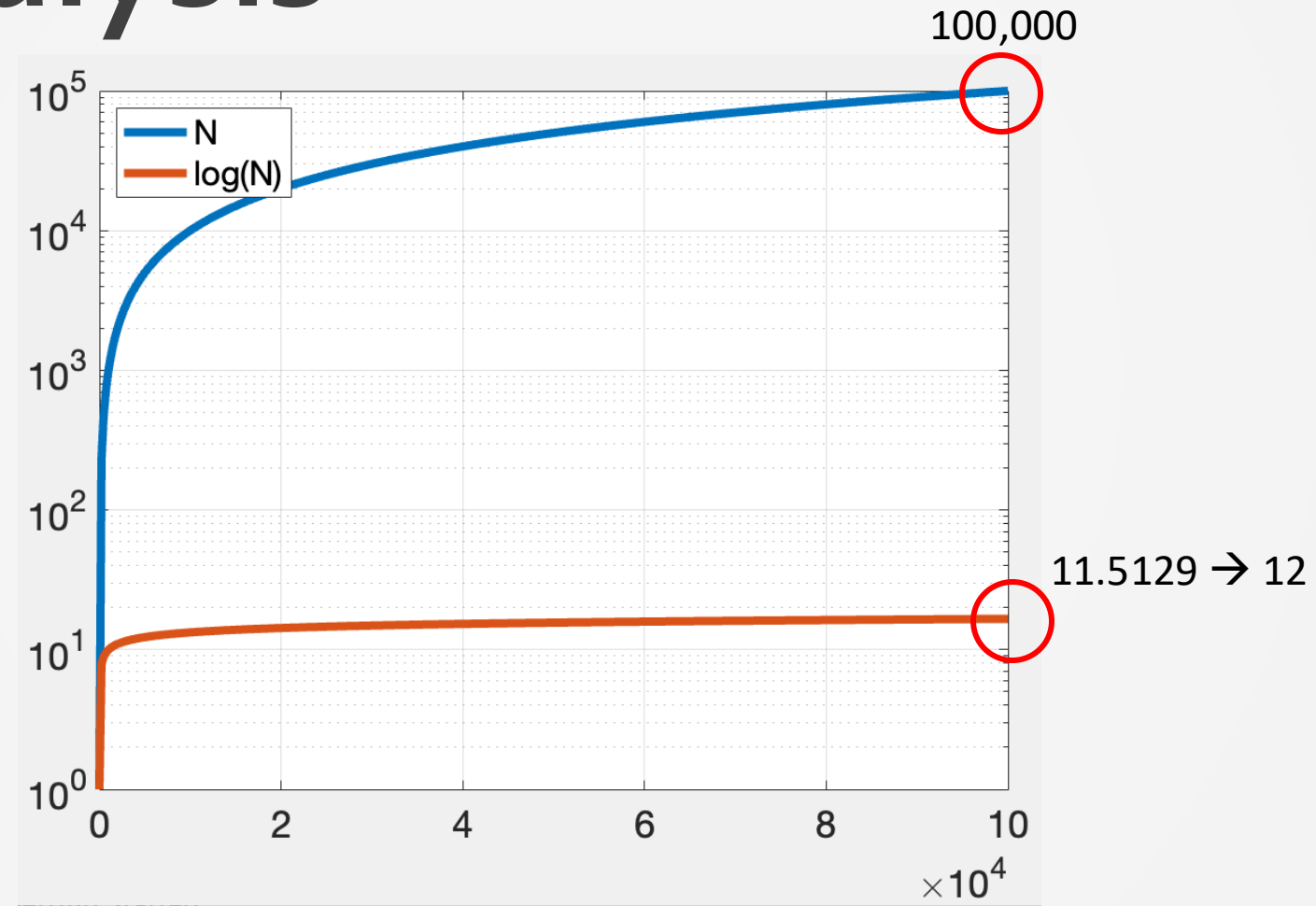
A bad shape of BST



N

Why advanced data structure?

An analysis



$N=100,000$

Why data structure

- Data organization, management, and storage format
- Efficient access and modification
- Consider and analyze:
 - Data values
 - Relationships among them
 - Functions or operations that can be applied

