Efficiently Searching

Giovanni Dicanio
AUTHOR, SOFTWARE ENGINEER

https://blogs.msmvps.com/gdicanio

Overview

Linear search

Binary search

Theory (slides) & practice (demos)

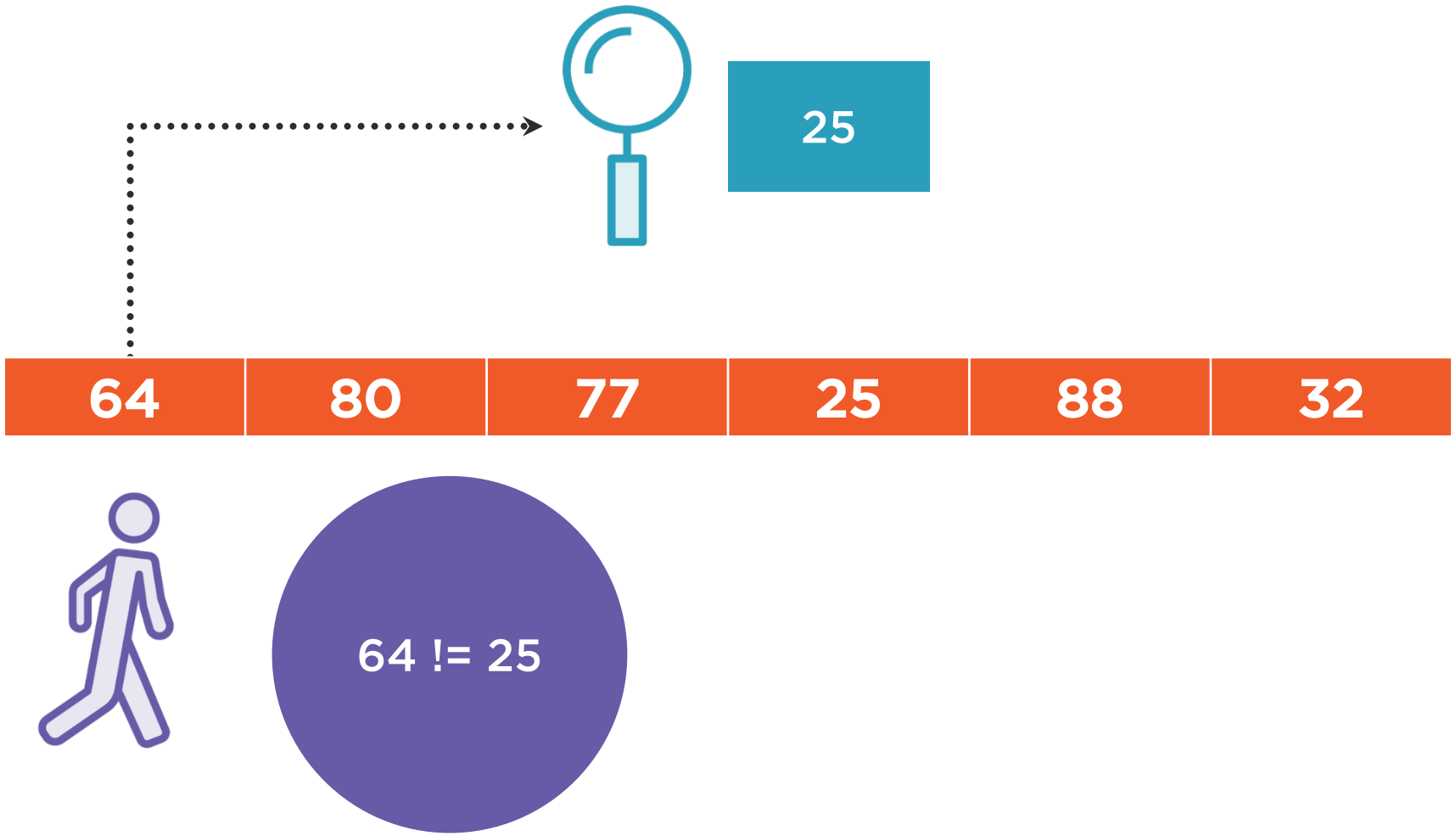
«Big-O» notation

Linear Search

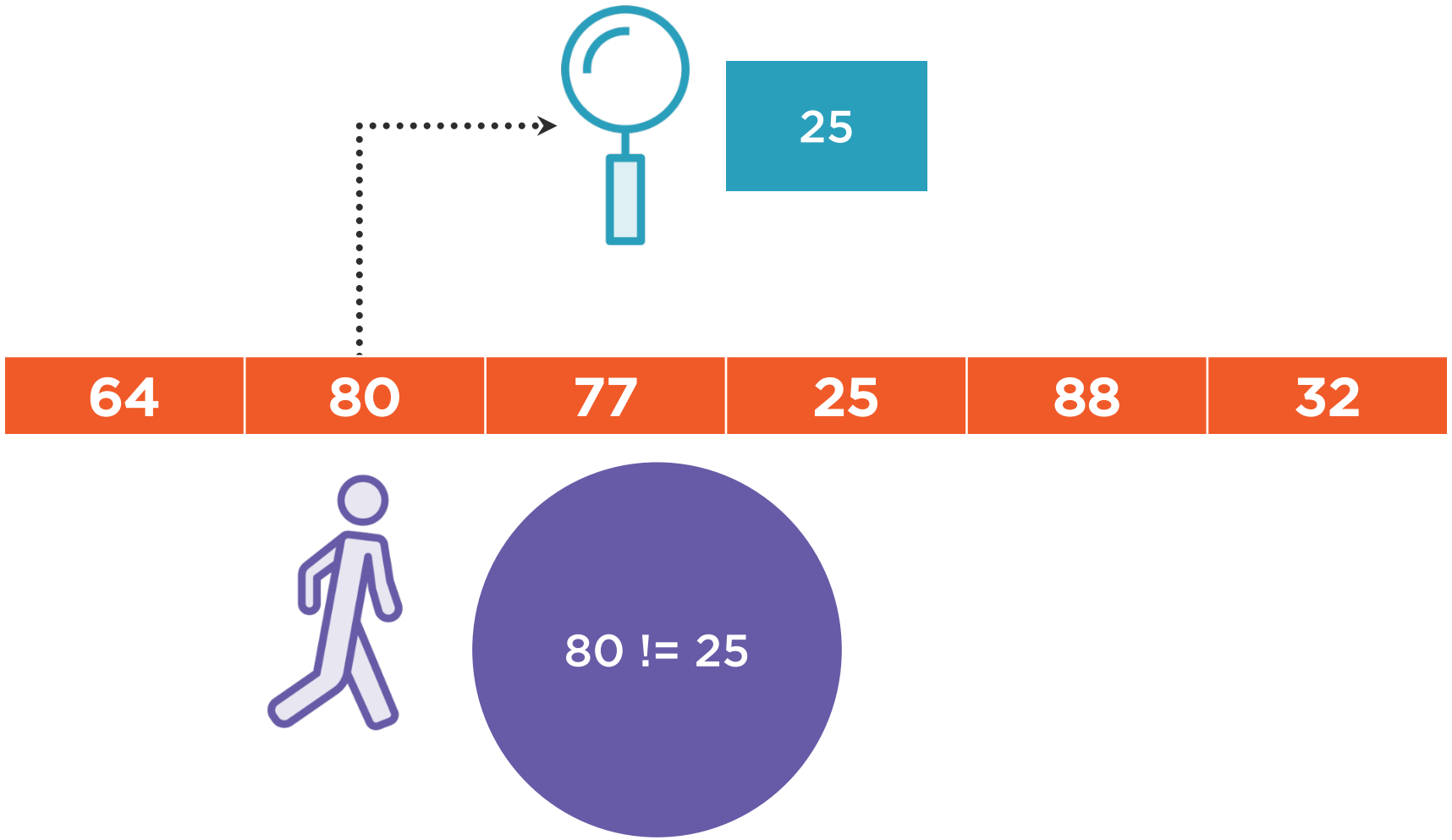
25

64	80	77	25	88	32

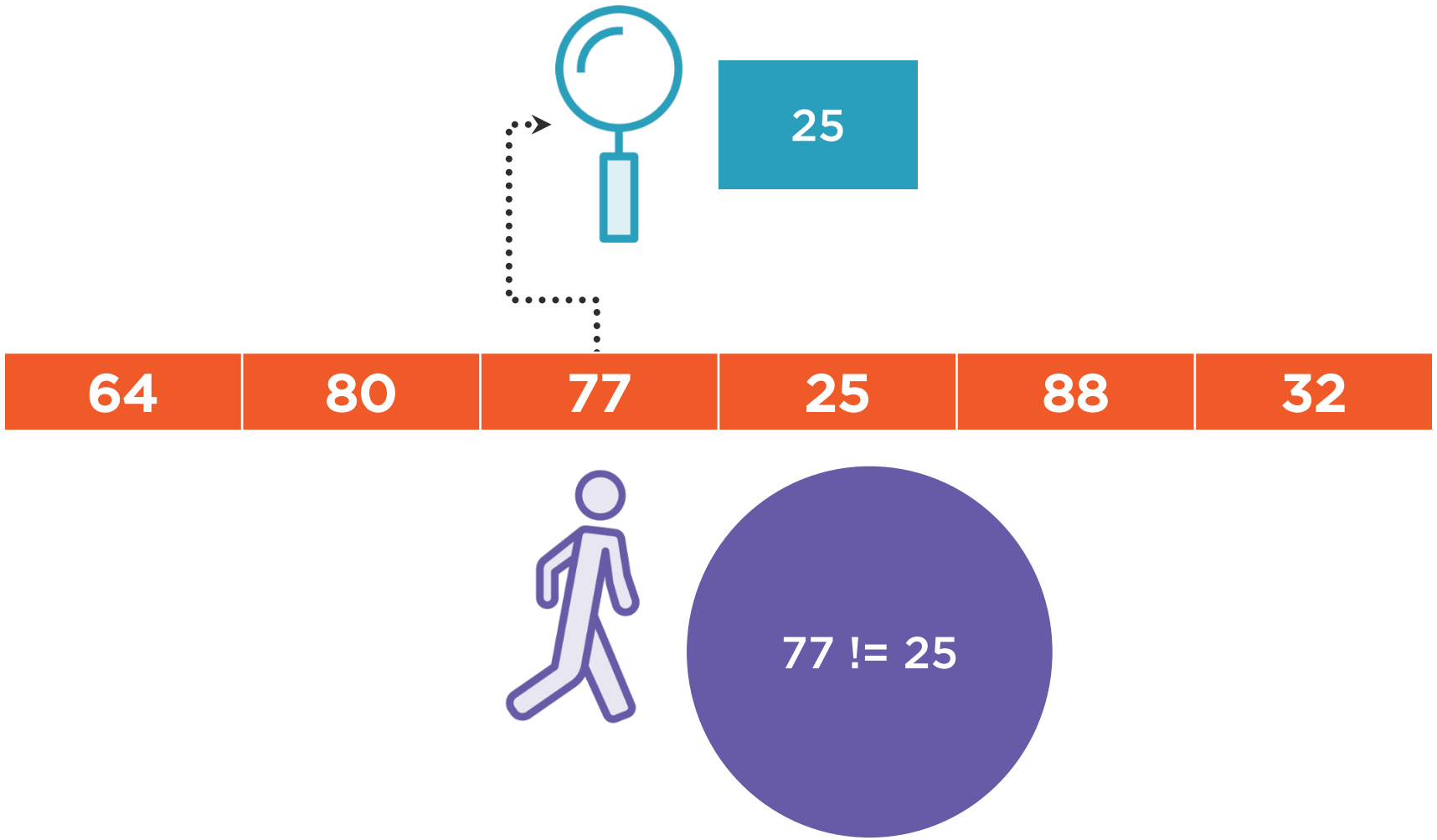
Linear Search



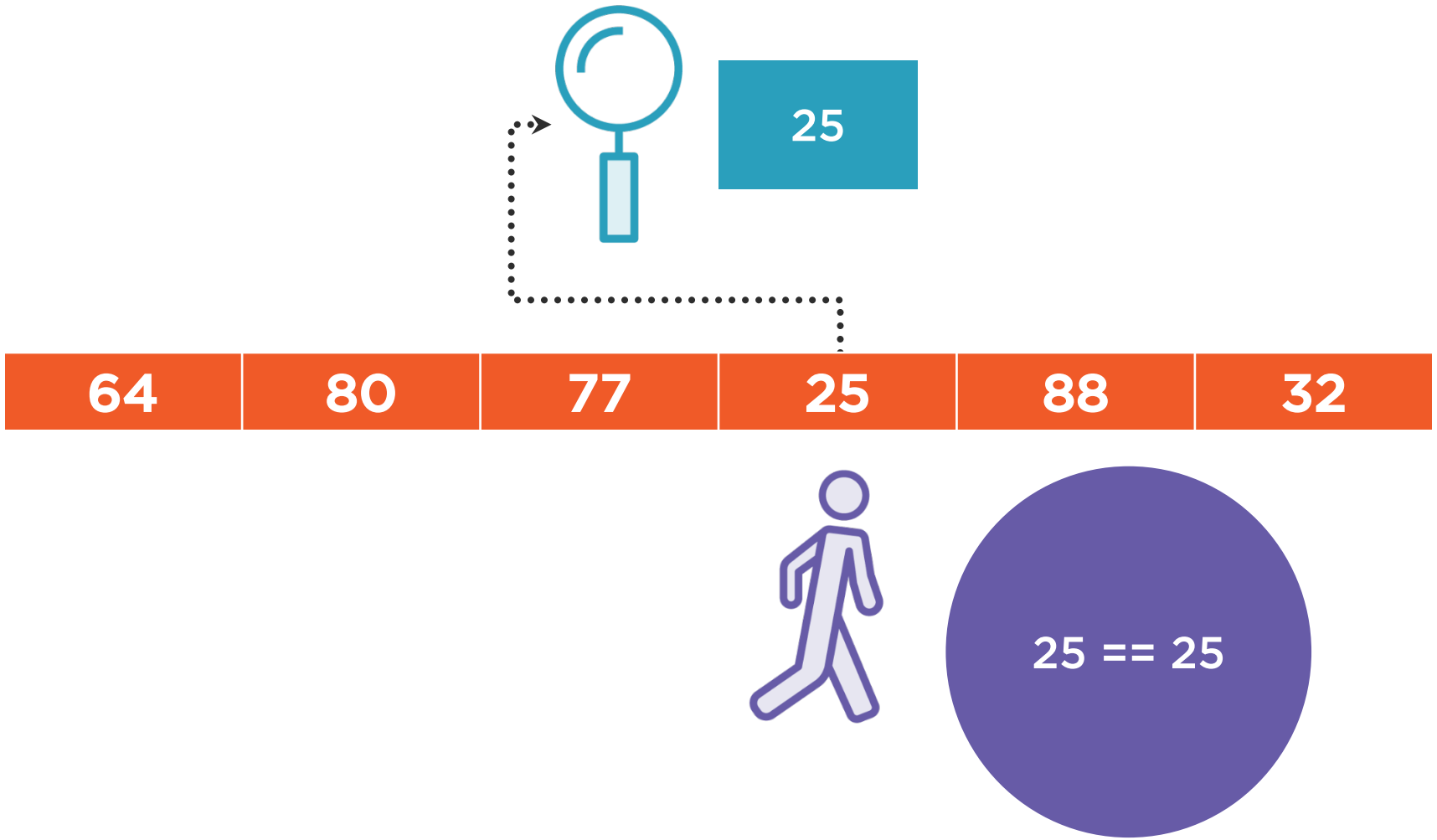
Linear Search



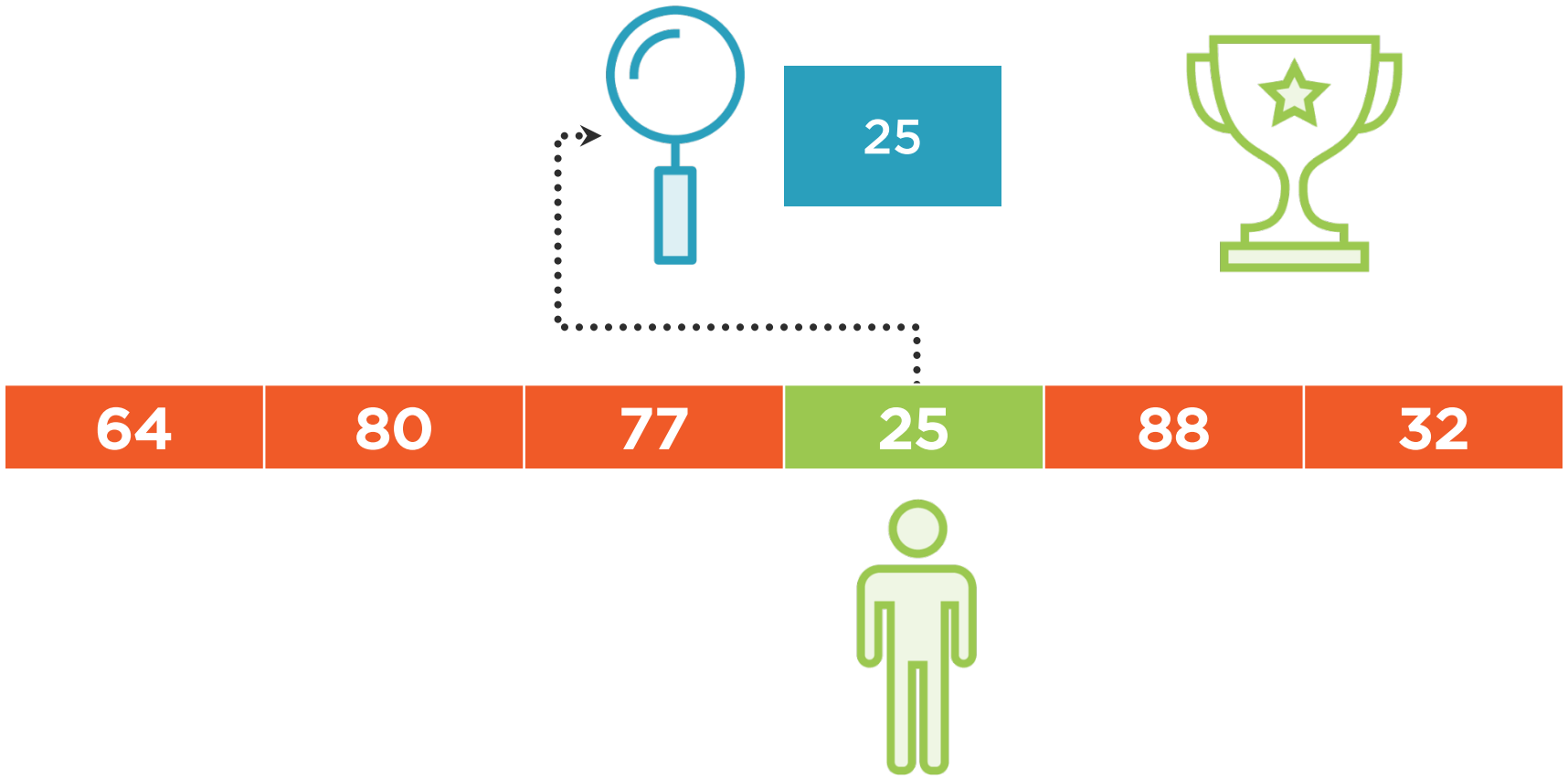
Linear Search



Linear Search



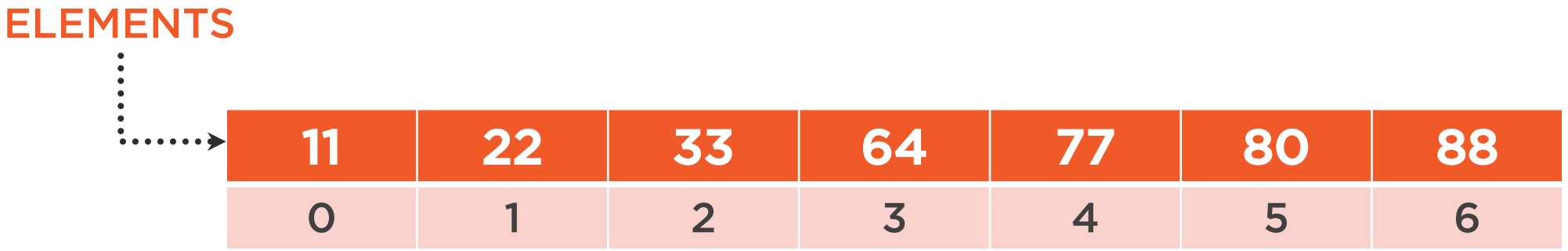
Linear Search



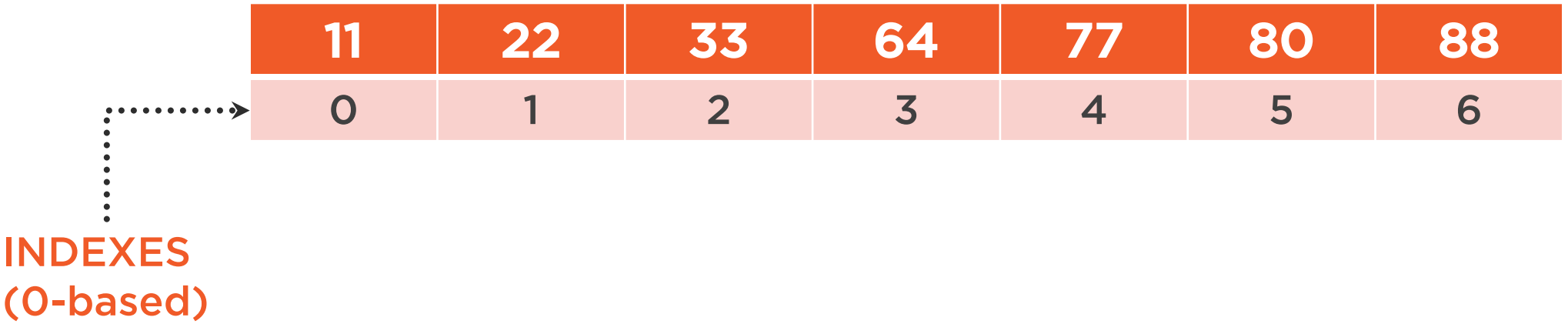
Binary Search

11	22	33	64	77	80	88
0	1	2	3	4	5	6

Binary Search



Binary Search



Binary Search

11	22	33	64	77	80	88
0	1	2	3	4	5	6

Binary Search

11	22	33	64	77	80	88
0	1	2	3	4	5	6

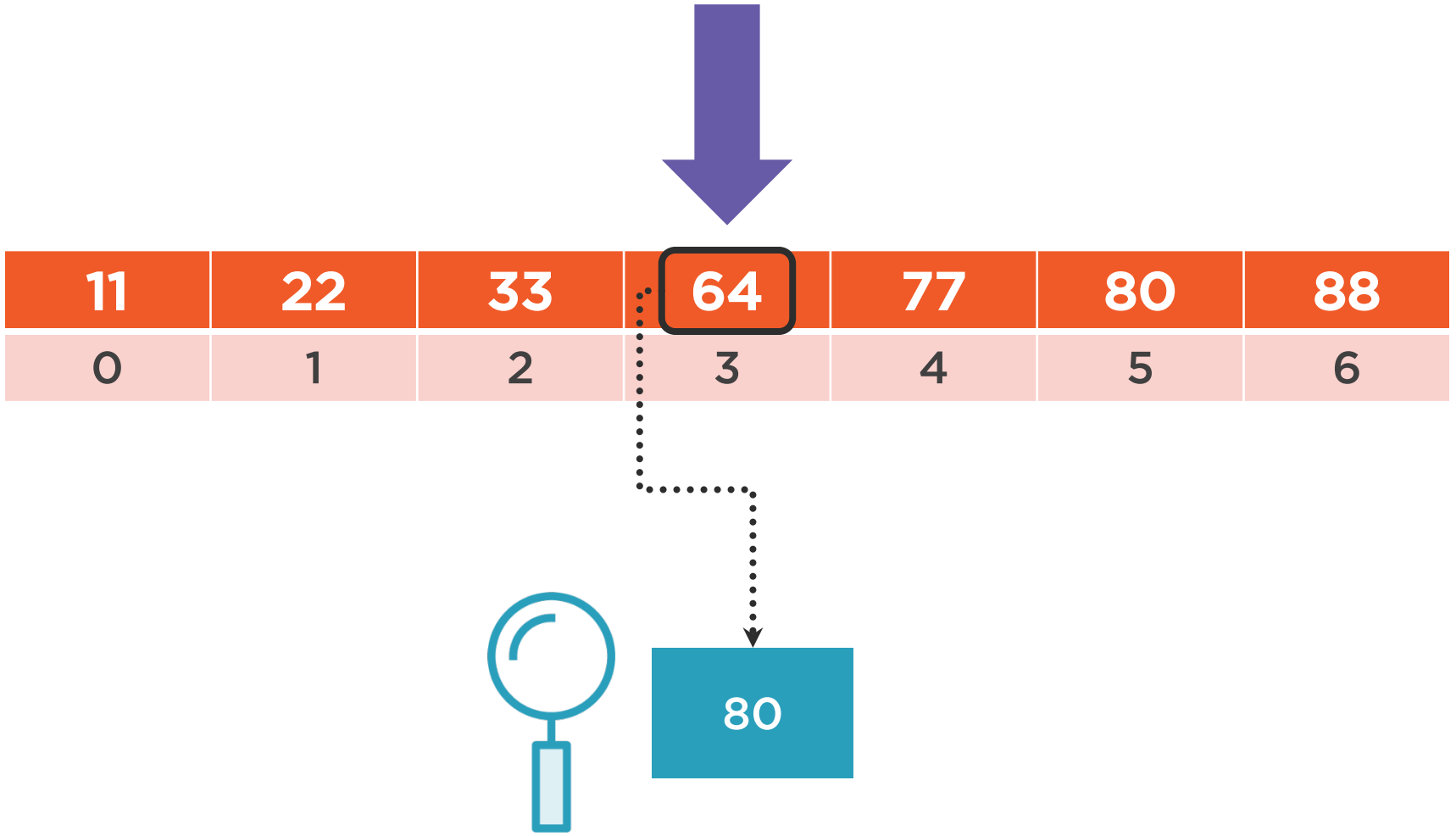
80

Binary Search

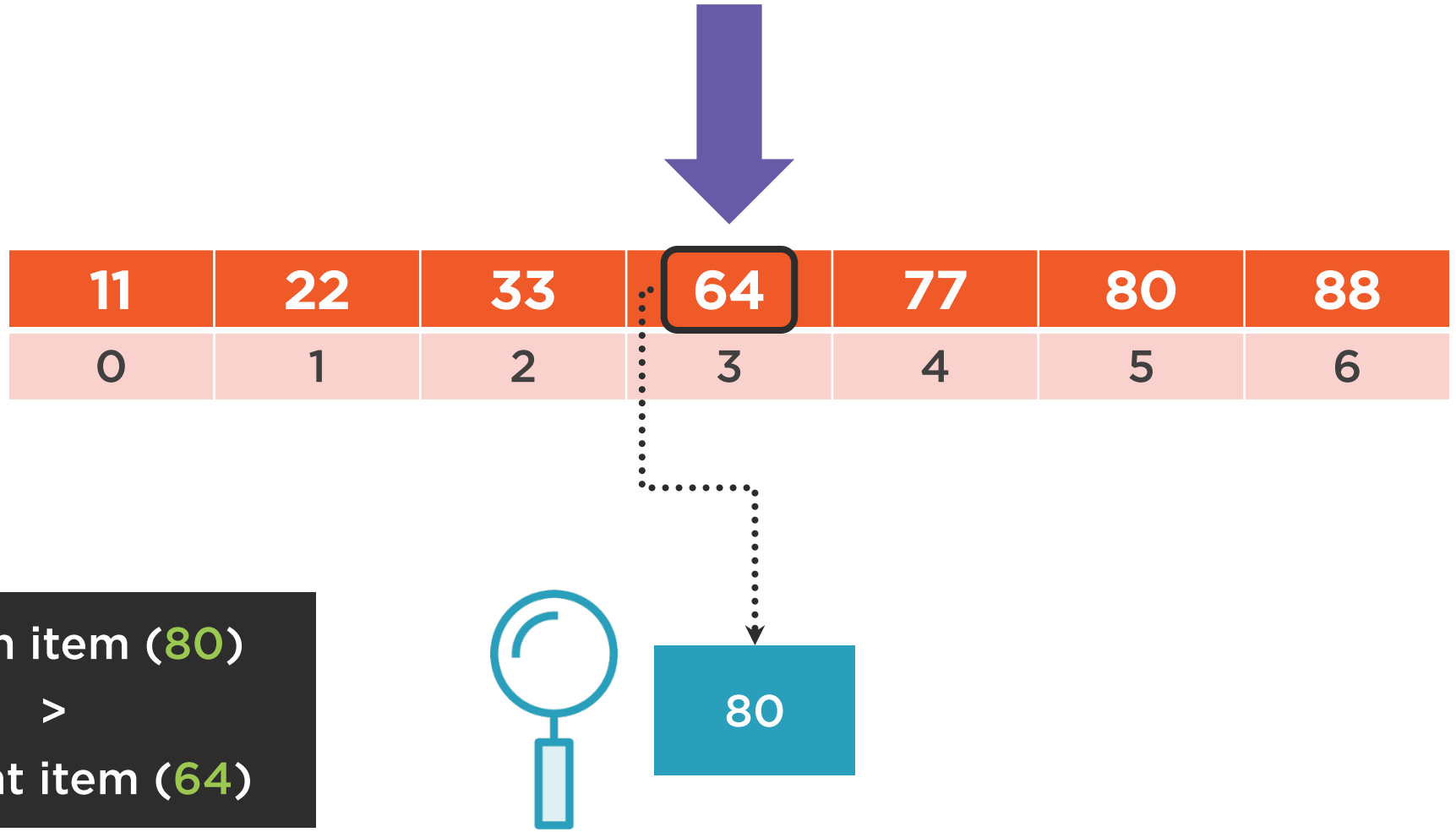
11	22	33	64	77	80	88
0	1	2	3	4	5	6

80

Binary Search



Binary Search

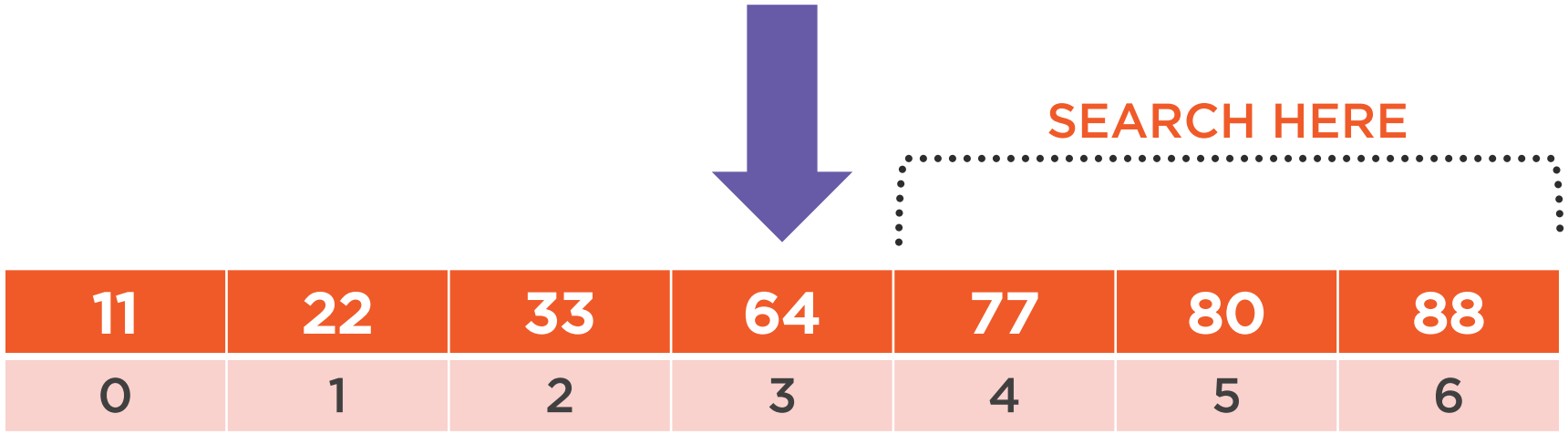


Search item (80)

>

Current item (64)

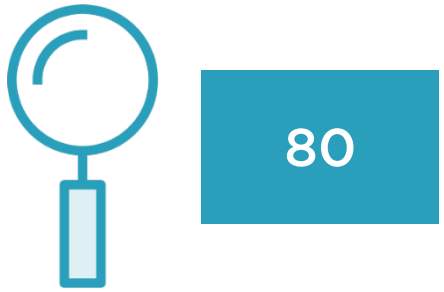
Binary Search



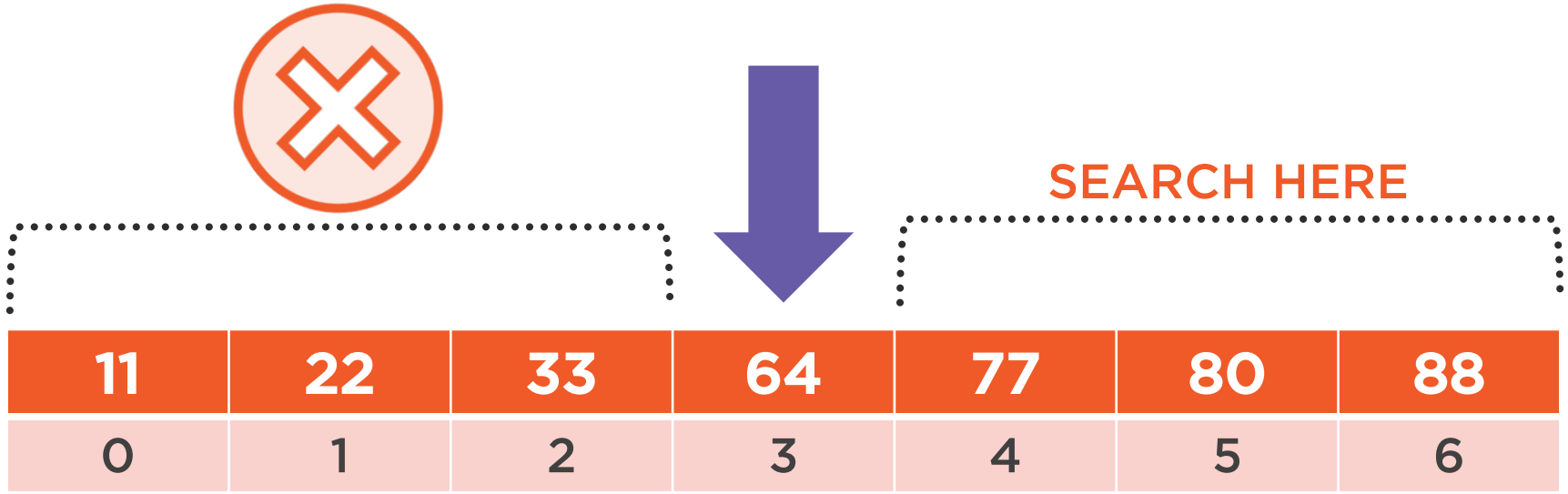
Search item (80)

>

Current item (64)



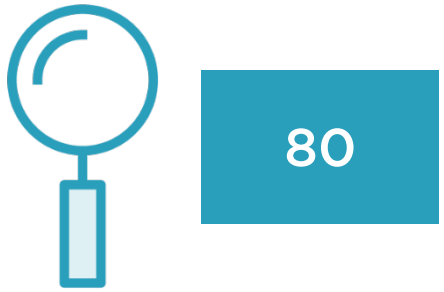
Binary Search



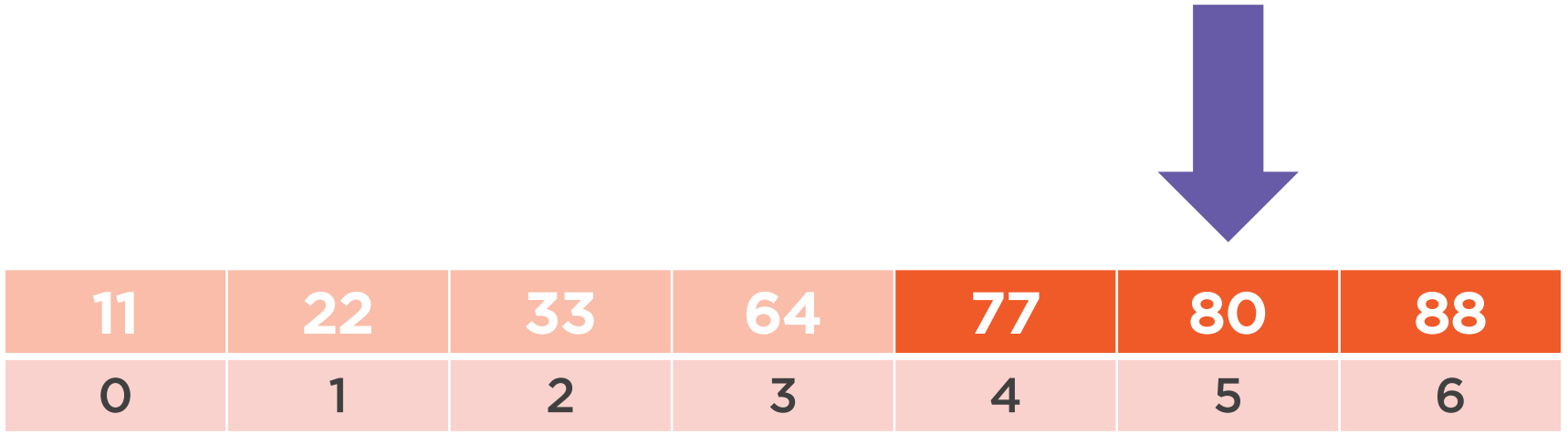
Search item (80)

>

Current item (64)



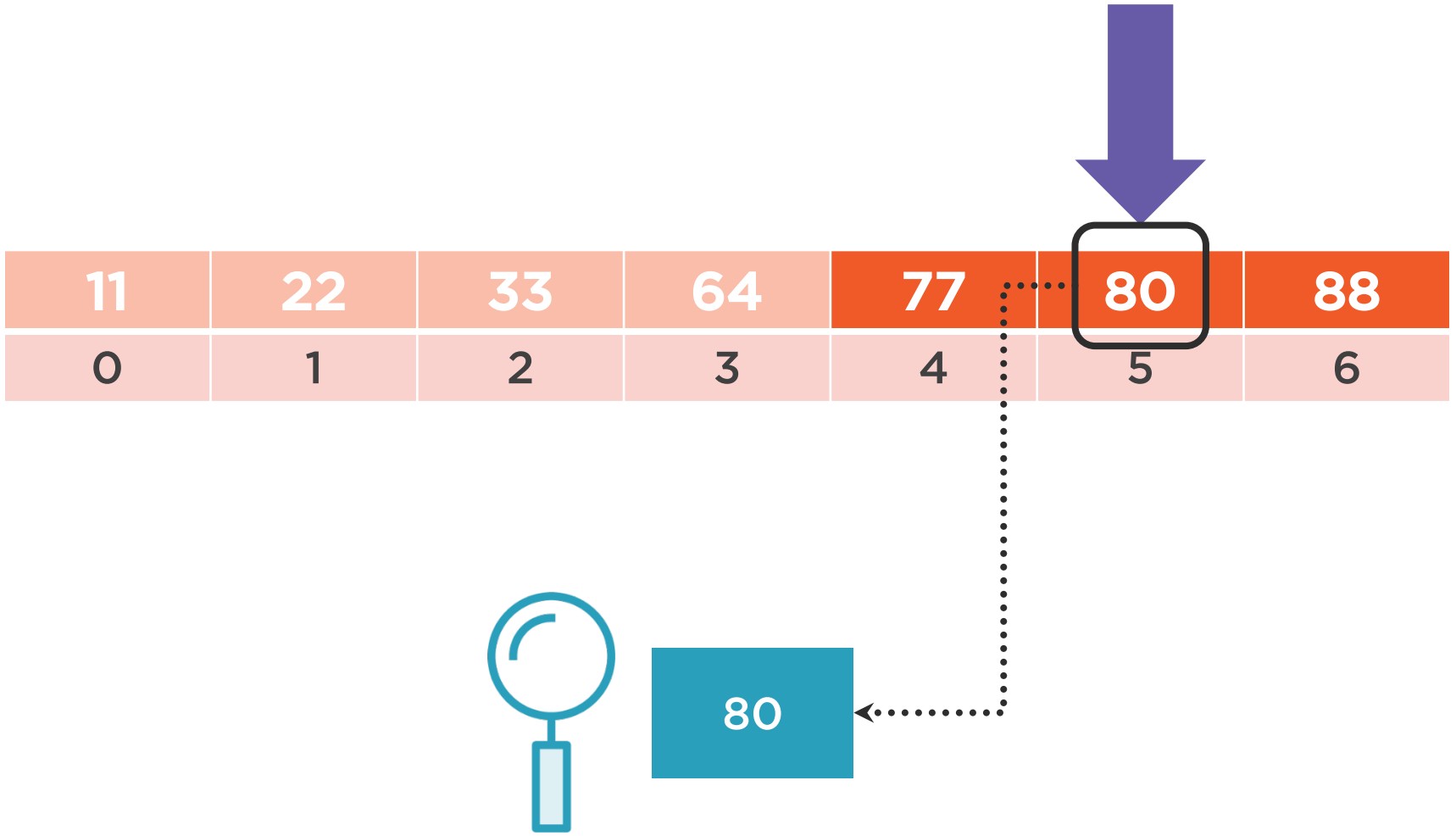
Binary Search



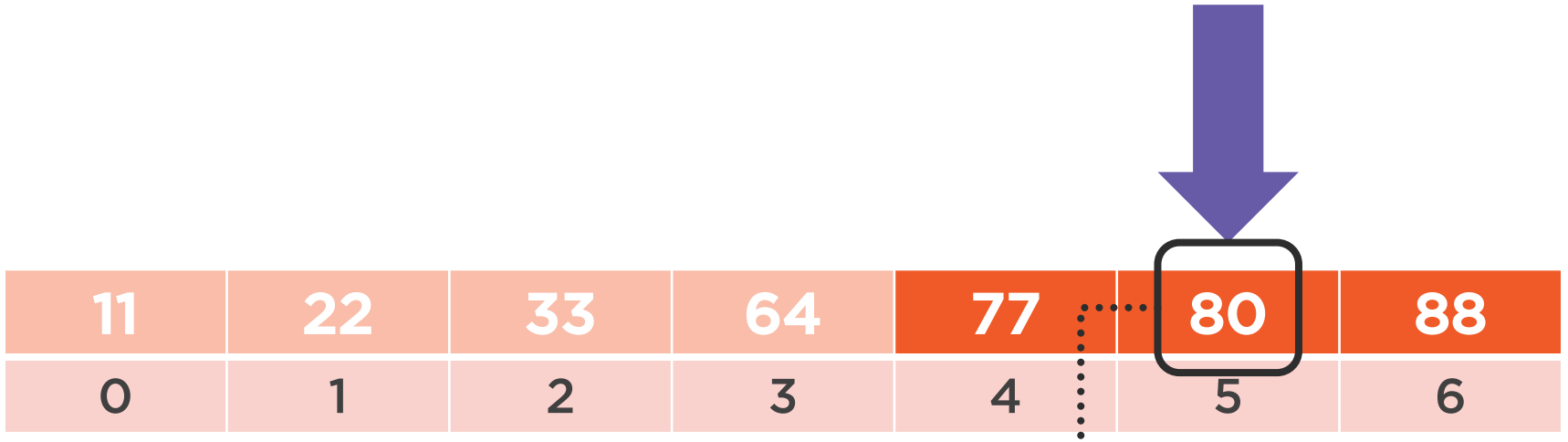
80

Repeat the process in the *right half*

Binary Search



Binary Search

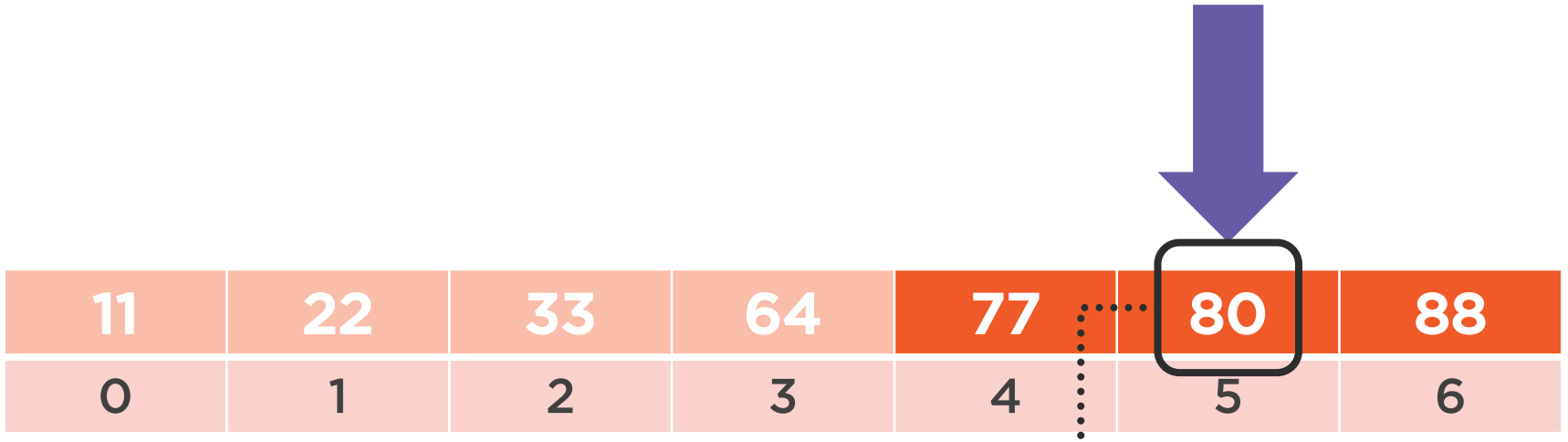


Search item (80)

==

Current item (80)

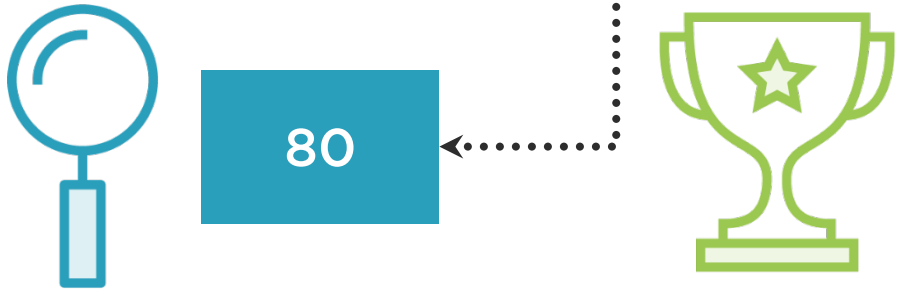
Binary Search



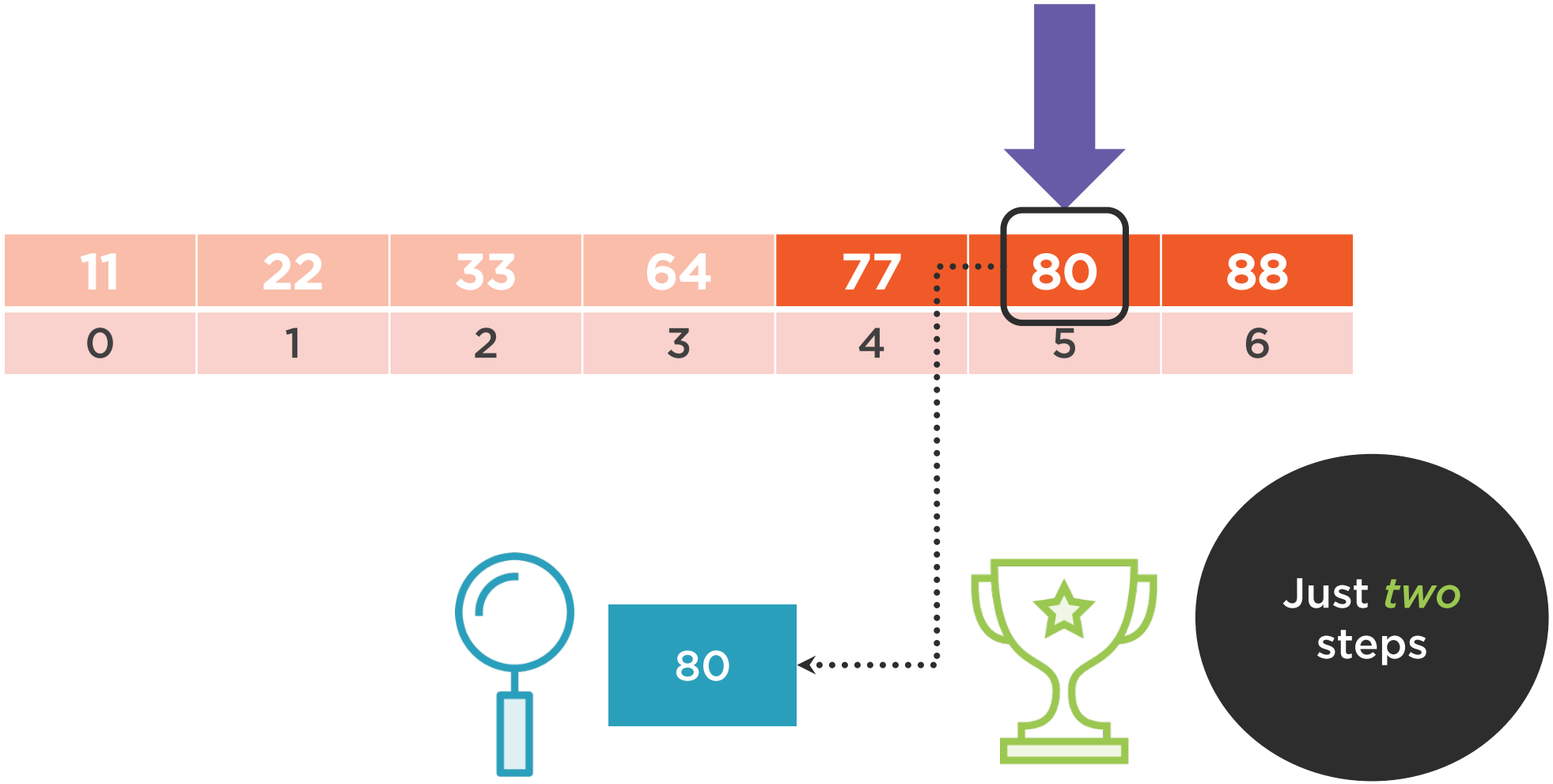
Search item (80)

==

Current item (80)



Binary Search



Linear Search

11	22	33	64	77	80	88
0	1	2	3	4	5	6

80

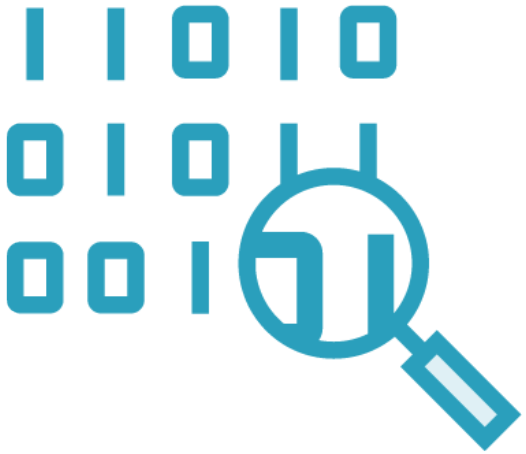
Linear search:

6 steps

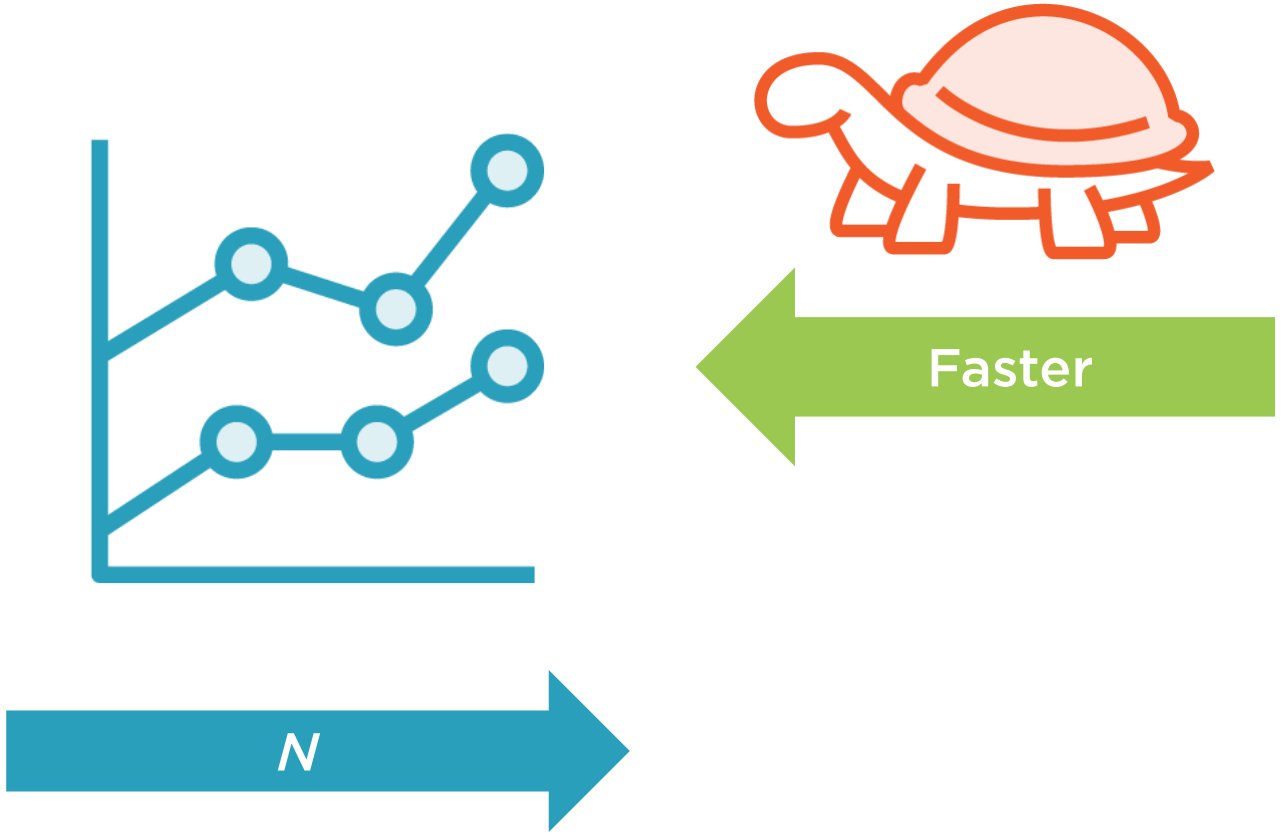
Introducing the «Big O» Notation

Introducing the «Big O» Notation

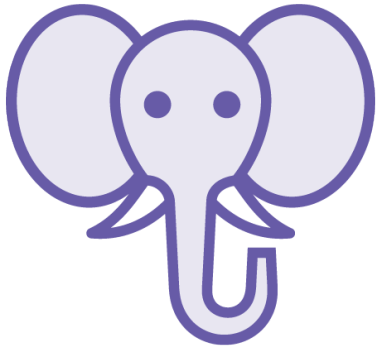
Trend



Estimate the Runtime Performance



«Big O» for Data Transfer



1-hour drive

Internet

Transferring Big (~100 GB) Files

Internet Speed

5 Mbps

0.4 Mbps

Transferring Big (~100 GB) Files

1-hour drive

Download: 5 Mbps
Size: 100 GB

+50 Hours

Data Transfer Algorithms

«Big-O» of Data Transfer Algorithms

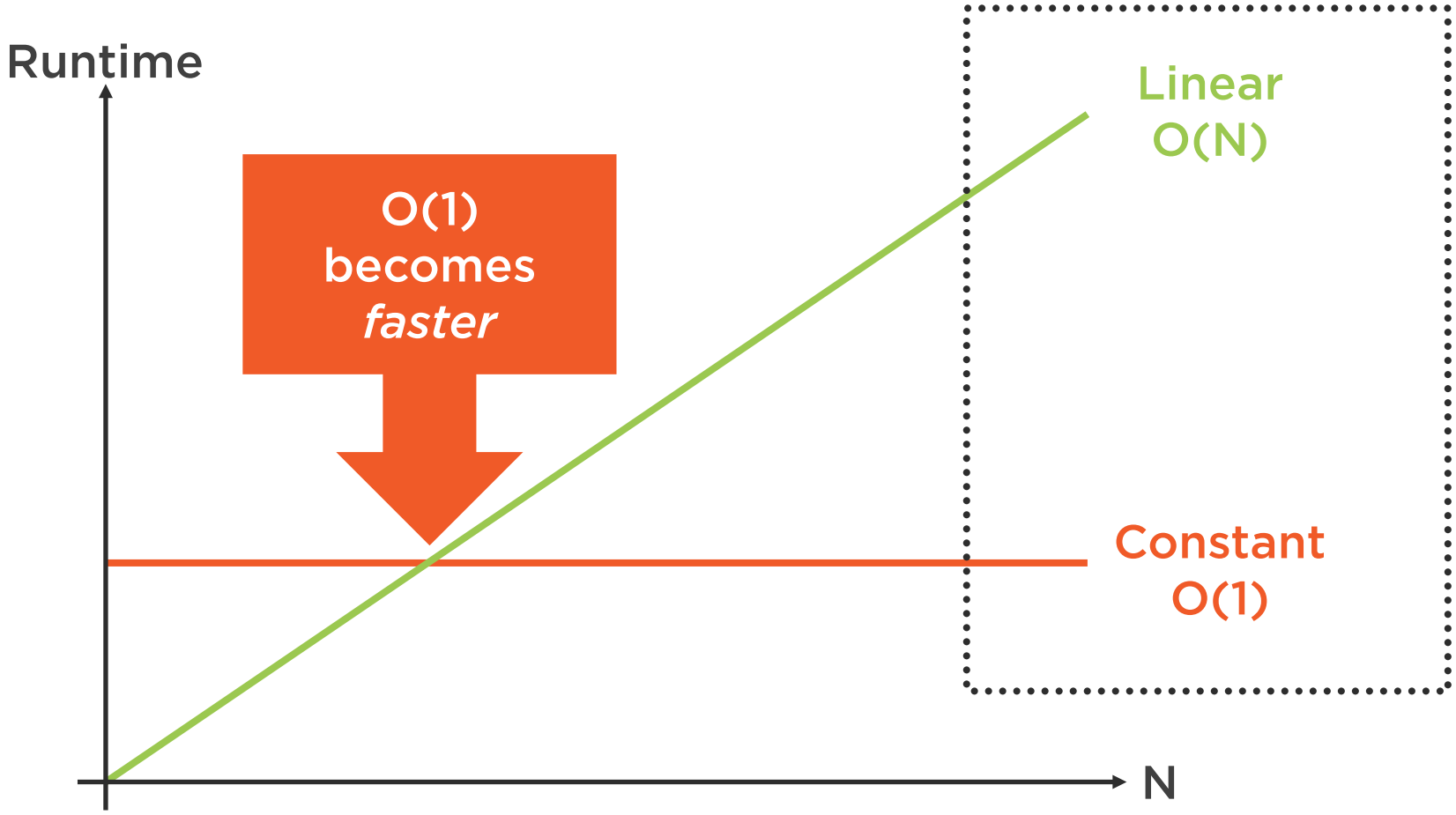
O(1)

Constant time
Independent of data size

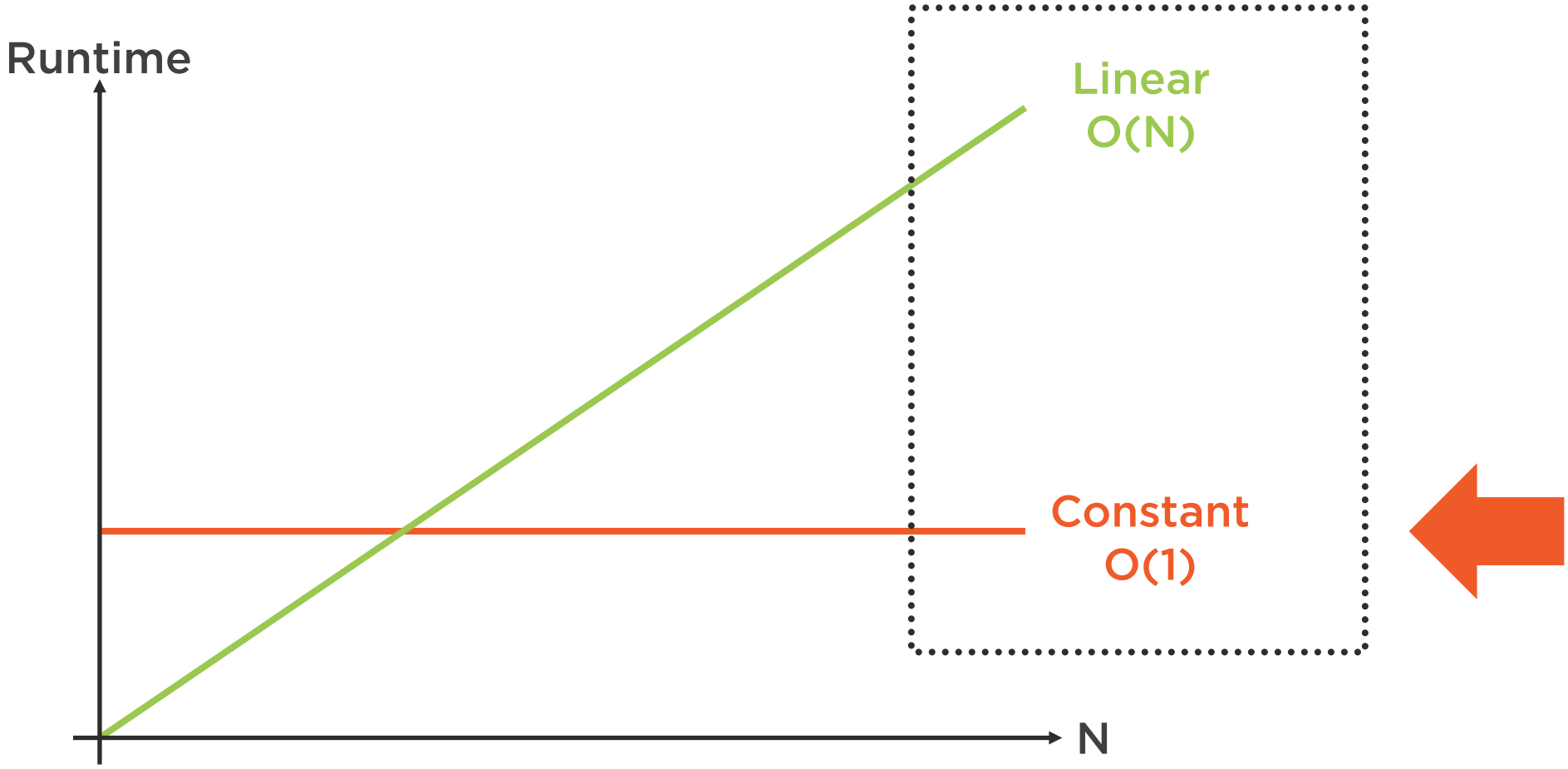
Linear
Proportional to data size

O(N)

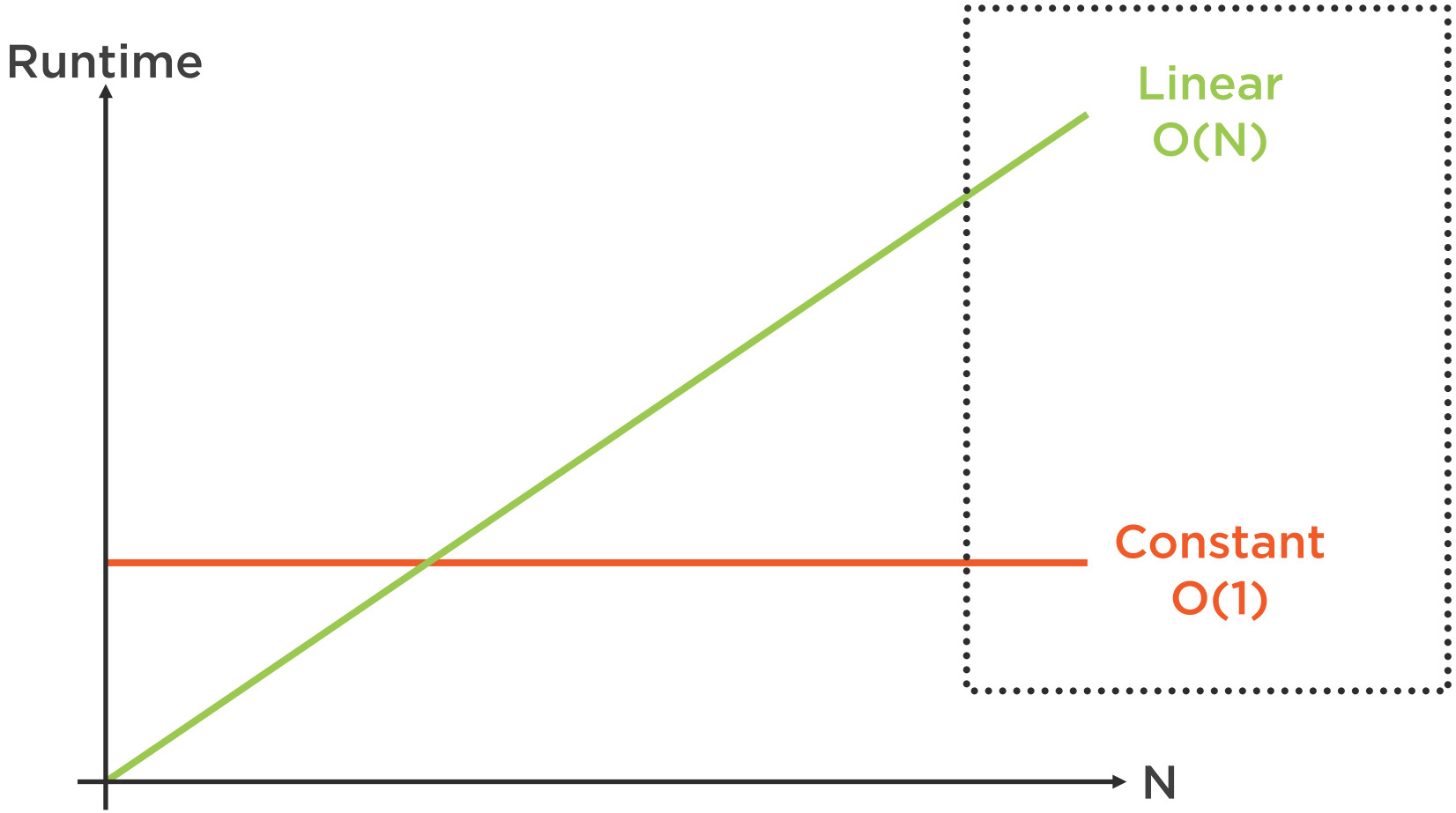
Linear vs. Constant Complexity



Linear vs. Constant Complexity



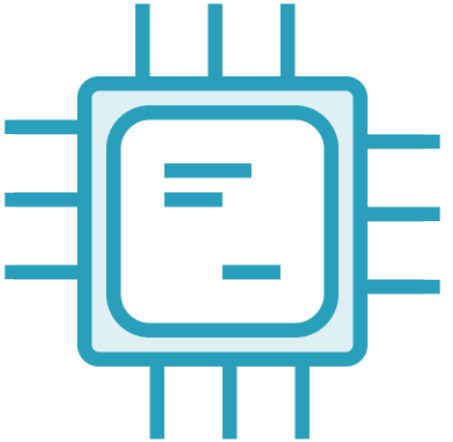
Linear vs. Constant Complexity



Asymptotic

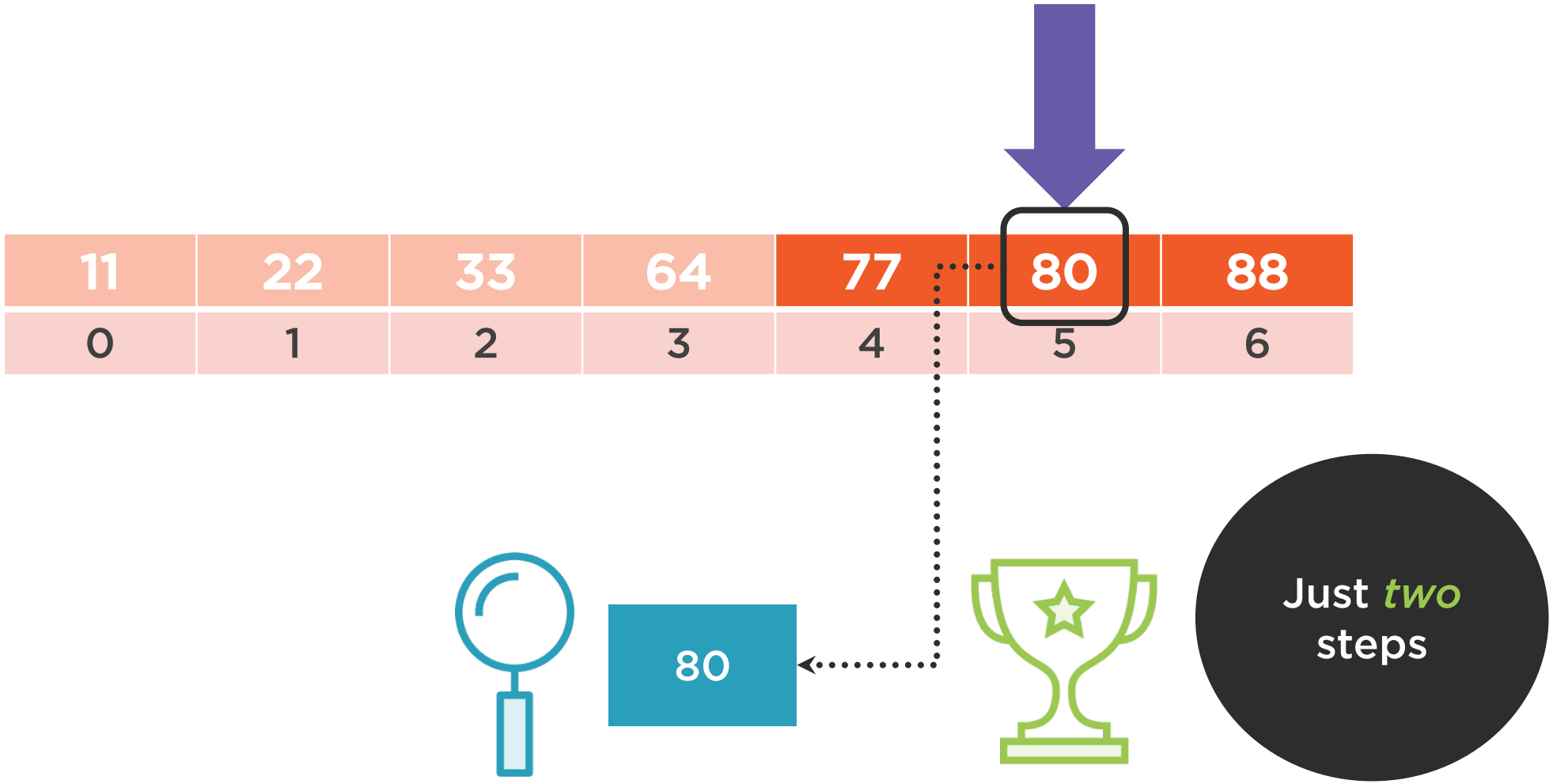
N >= 1,000,000

«Big-O» → Trends

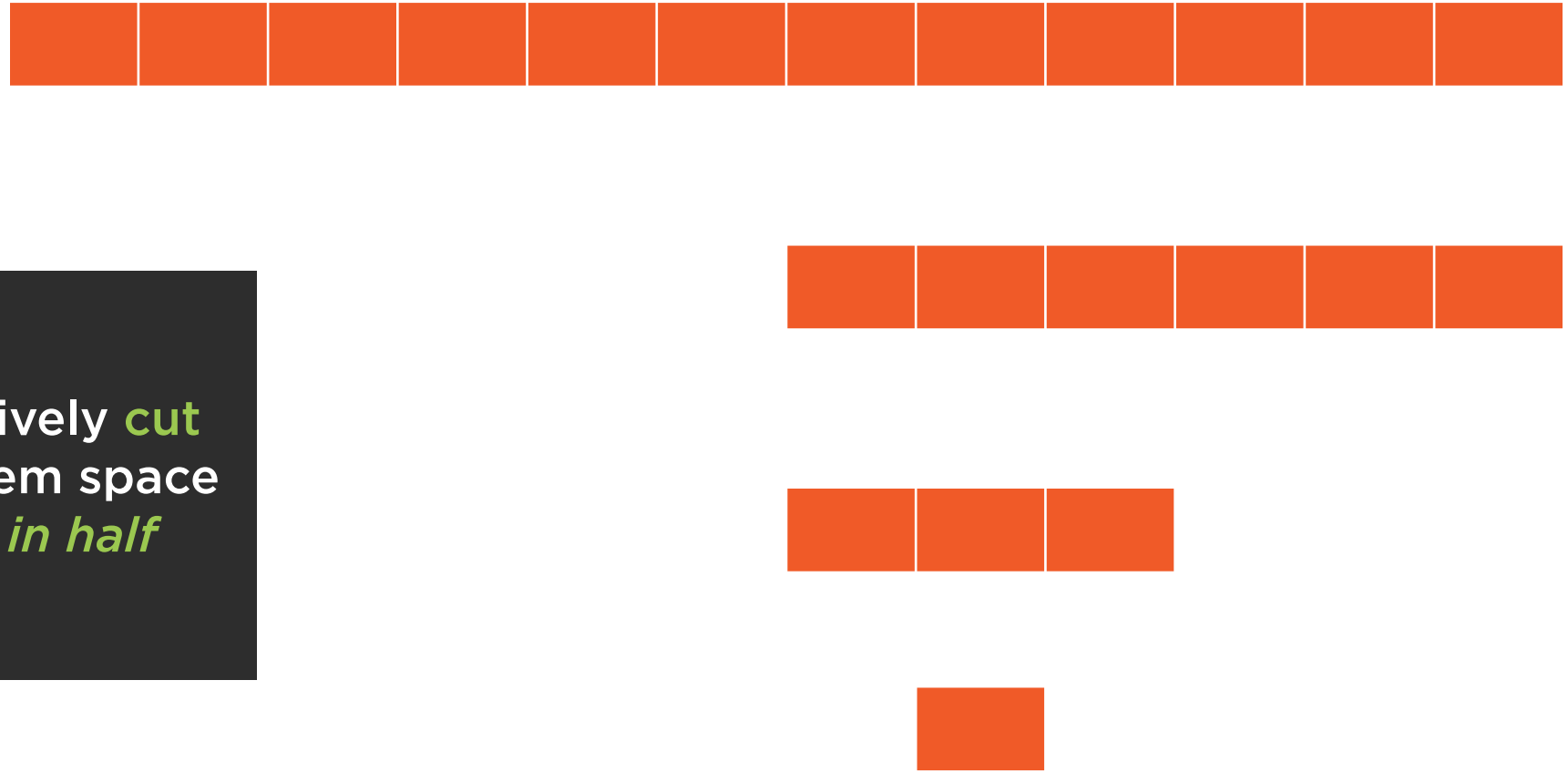


Cache-friendly?

Binary Search

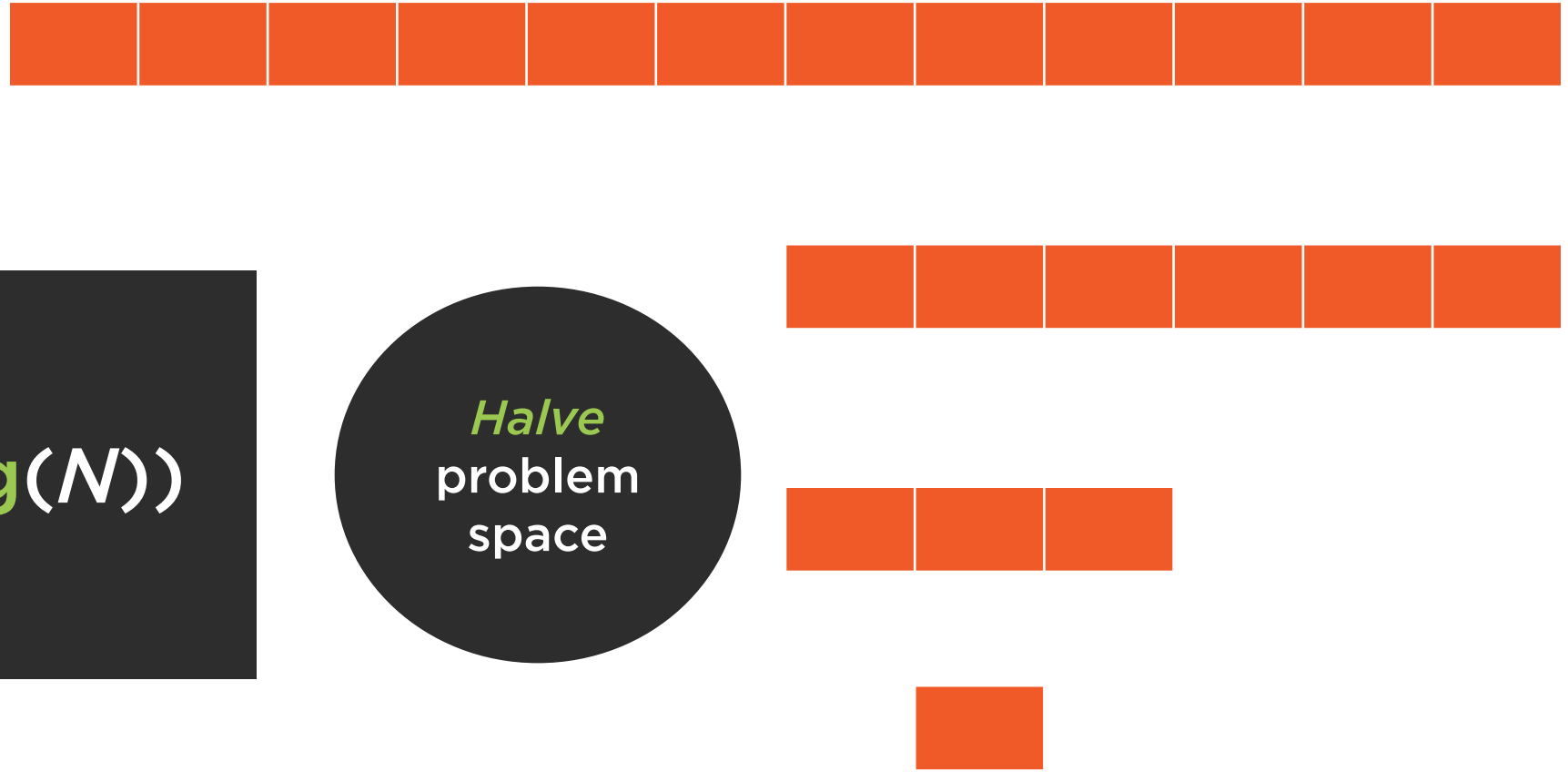


Binary Search

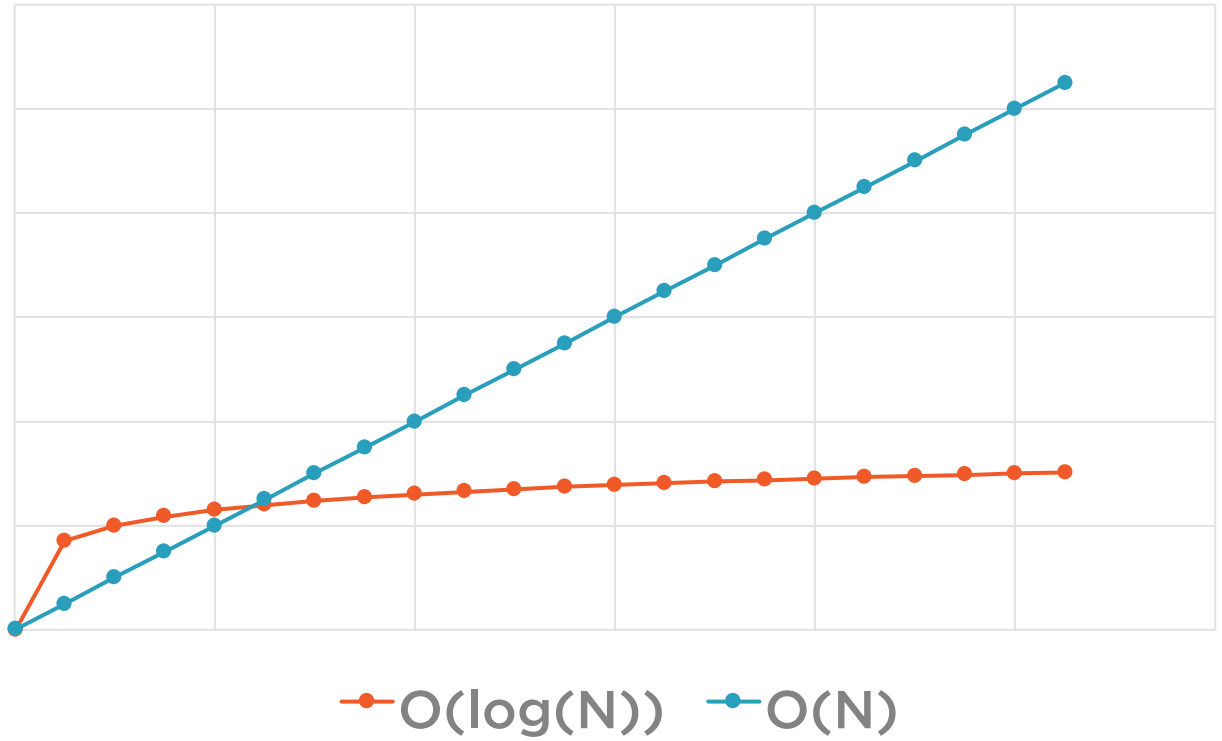


Progressively **cut** the problem space (array) ***in half***

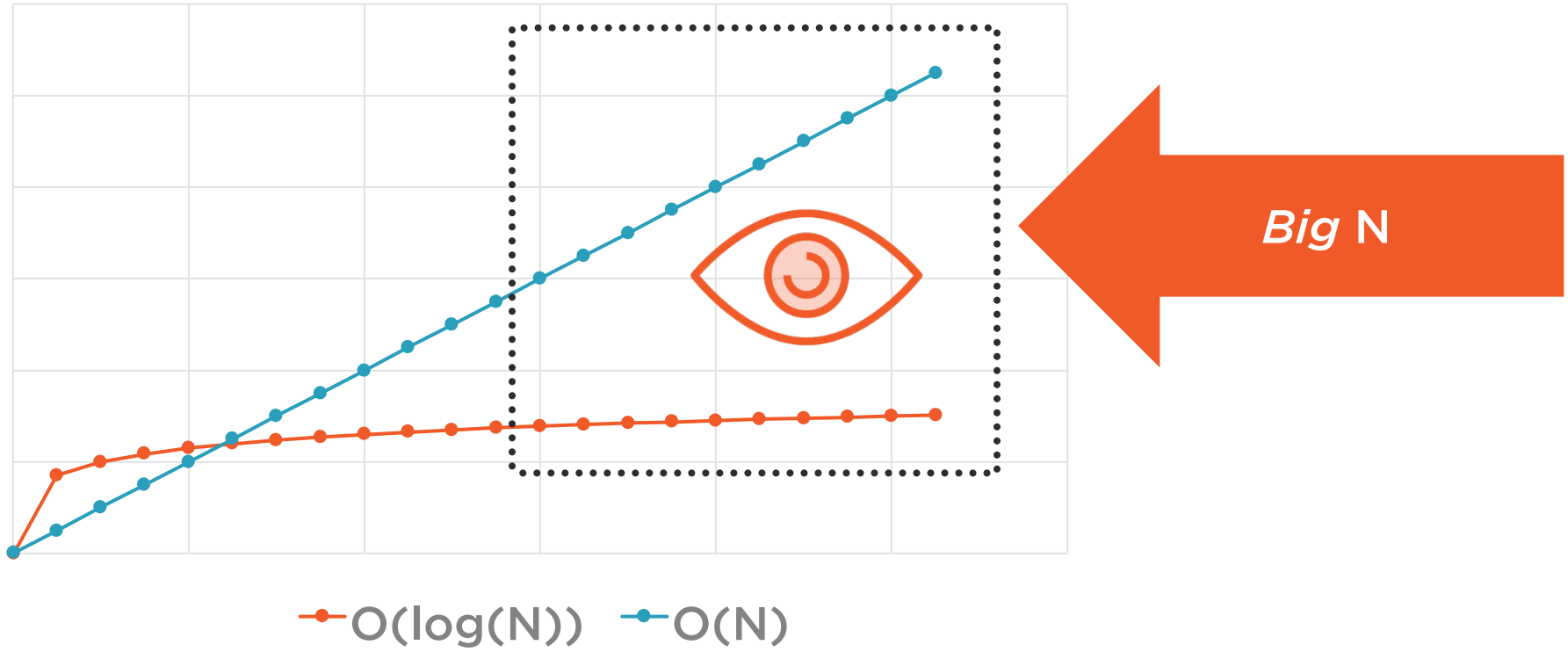
Binary Search



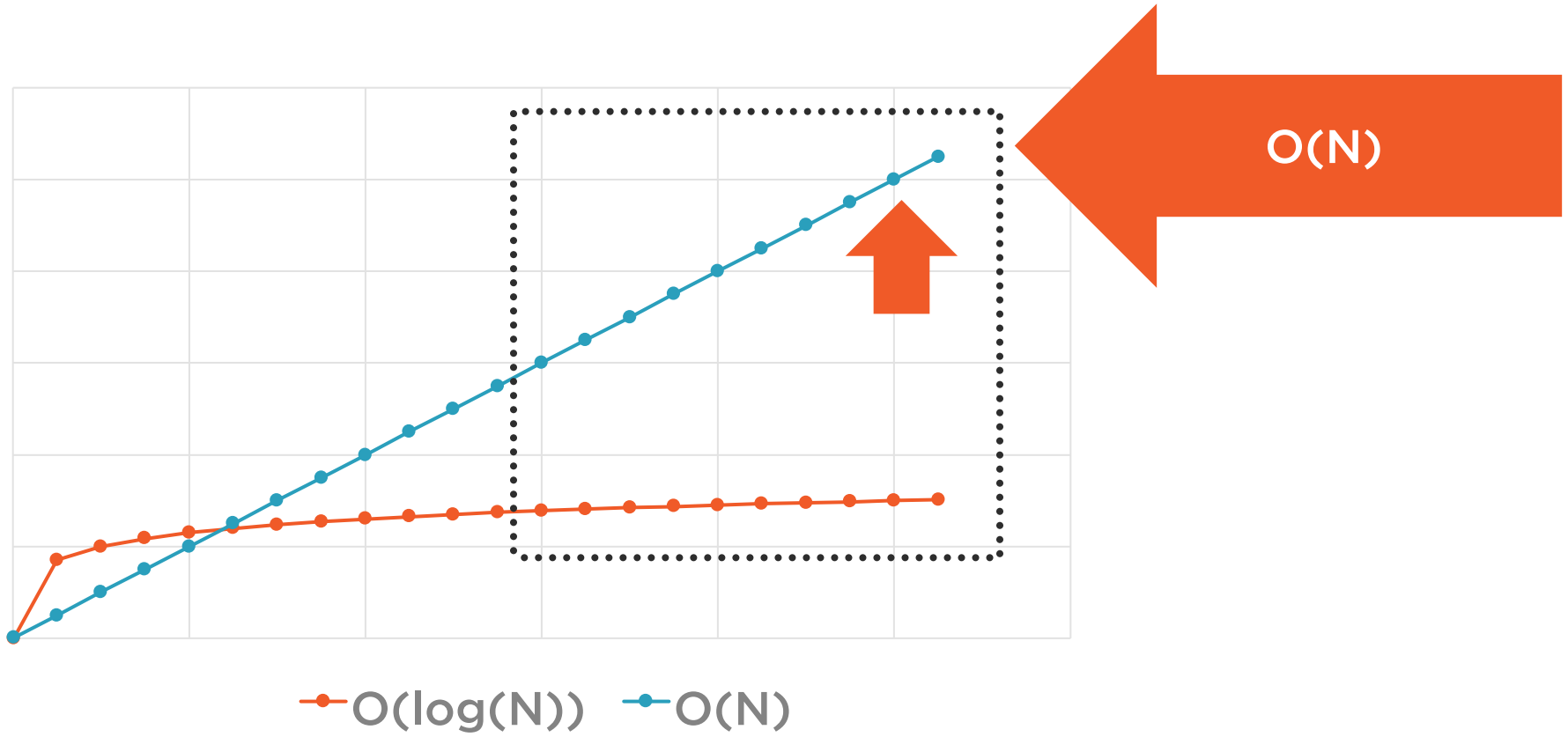
Asymptotic Analysis



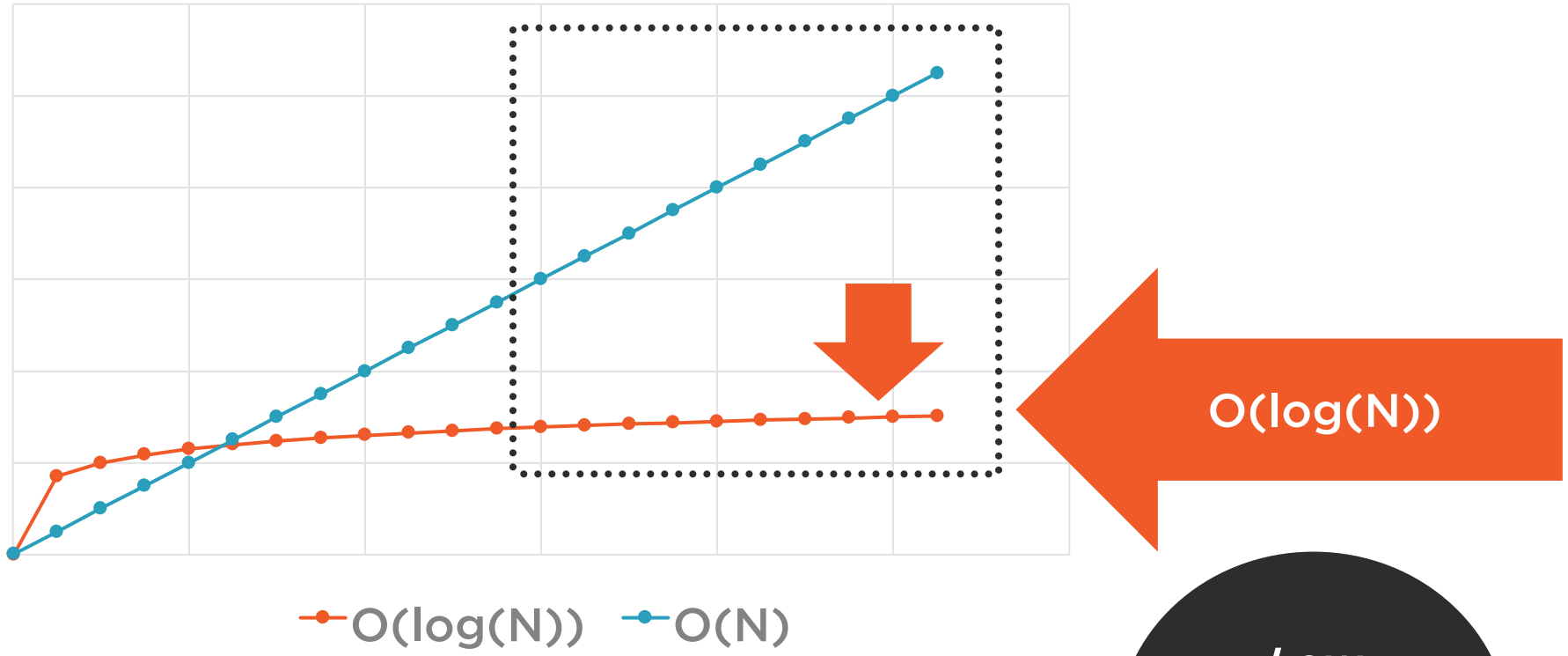
Asymptotic Analysis



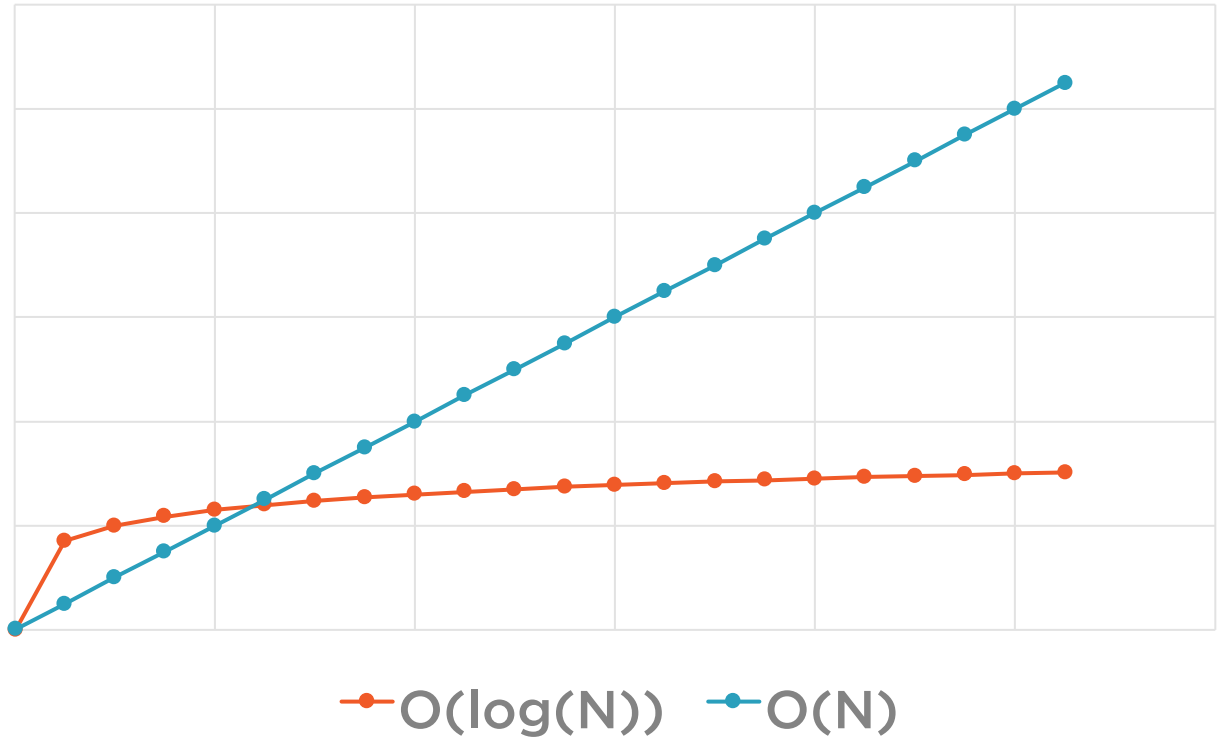
Asymptotic Analysis

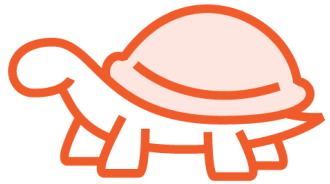


Asymptotic Analysis

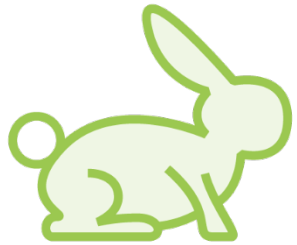


Asymptotic Analysis





Linear Search



Binary Search

For *sorted* arrays,
prefer *binary* search
to linear search

Summary

Linear search

Binary search

Runtime asymptotic analysis

O(*N*) vs. *O*(*log*(*N*))