The Power of Algorithms

SMASH program CS Electives

Hall of Fame

Intro to Algorithm

What is Algorithm?

What is Algorithm?

TL;DR: Just a bunch of steps.

What is Algorithm?

Algorithm is a set of rules to be followed in calculations or other problem-solving operations, or a procedure for solving a mathematical problem in a finite number of steps.

Therefore Algorithm refers to a sequence of finite steps to solve a particular problem.

Algorithms are everywhere

- Social media (Tiktok’s FYP)
- Music (Spotify’s playlist)
- Video (Netflix’s recommendations system)
- Navigation (Google map’s best route or least walking or lowest cost)
- Search (Google search)
- Cryptocurrency (Bitcoin’s SHA-256)

Game No.1: Guessing numbers!

Two students as a pair, guessing each other’s number.

Note: Remember to track and note down each guess on the paper!

Discussion

Observe the numbers you have guessed, can you find any pattern? Or did you use any trick to allow you to get the number faster?

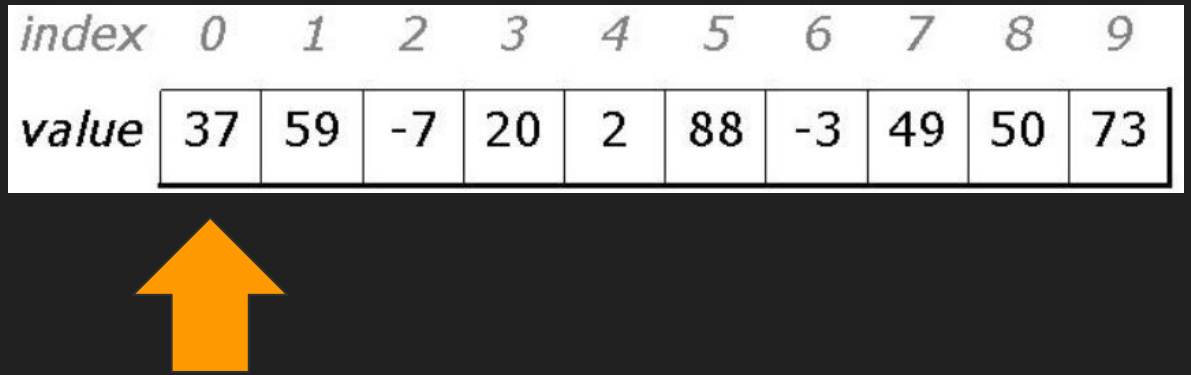
Further question: We’re like so smart, and good-looking, attractive, pretty, handsome, and beautiful, but what about computers? what strategy can they use to search a number?

Next a few slides are provided through the courtesy of my 5008 Professor Michael Shah.

It’s him

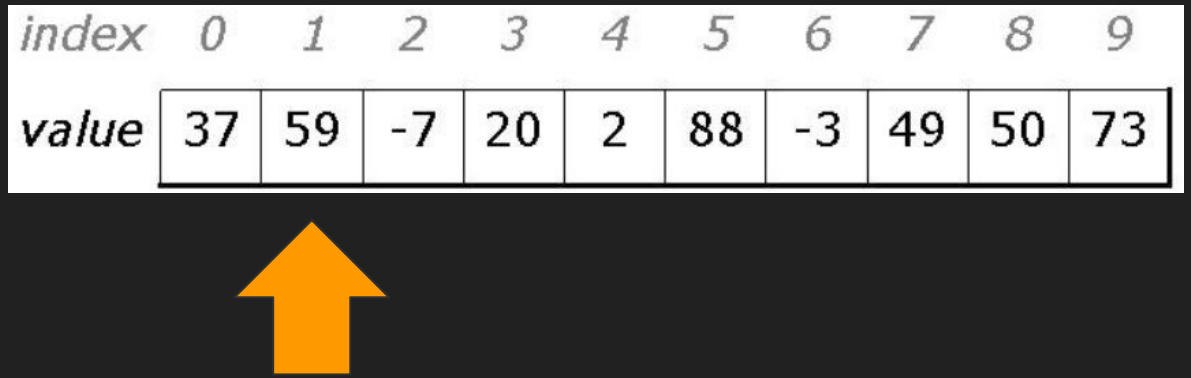
Linear Search - to check the number one by one

Let’s say we want to find if 88 is in the group?

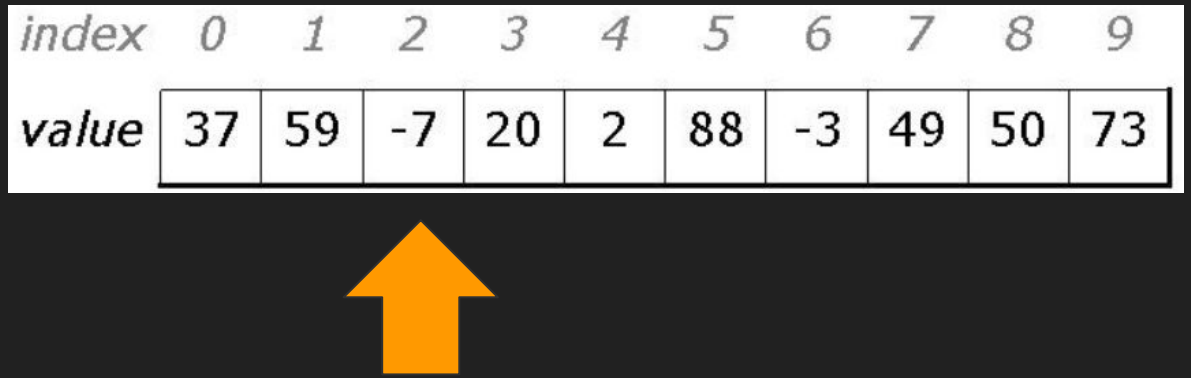


index	0	1	2	3	4	5	6	7	8	9
value	37	59	-7	20	2	88	-3	49	50	73

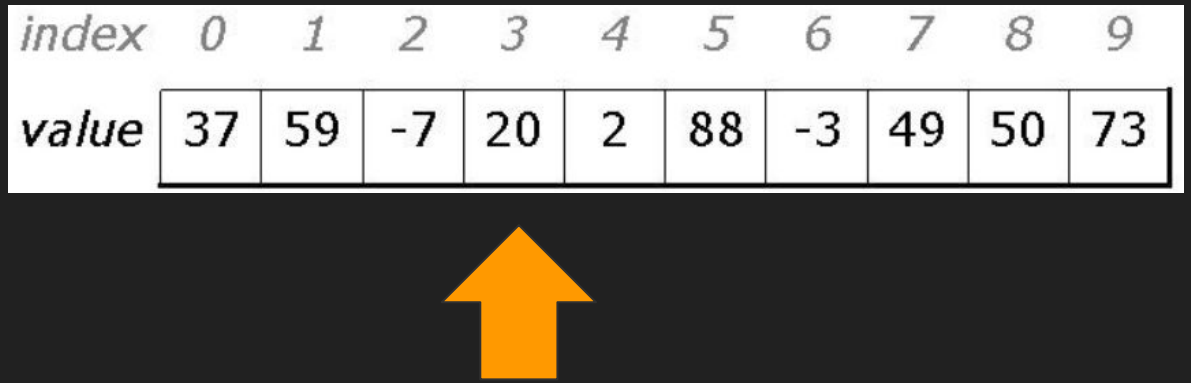
Linear Search - to check the number one by one



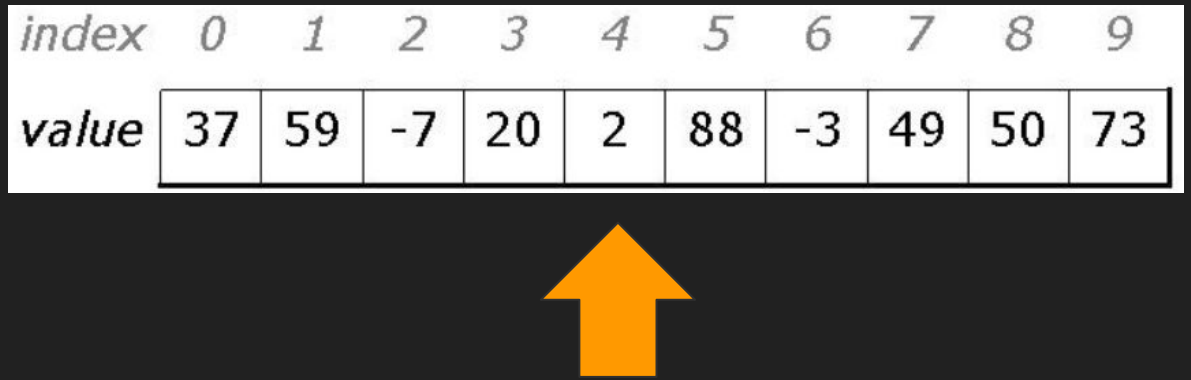
Linear Search - to check the number one by one



Linear Search - to check the number one by one



Linear Search - to check the number one by one



Linear Search - to check the number one by one

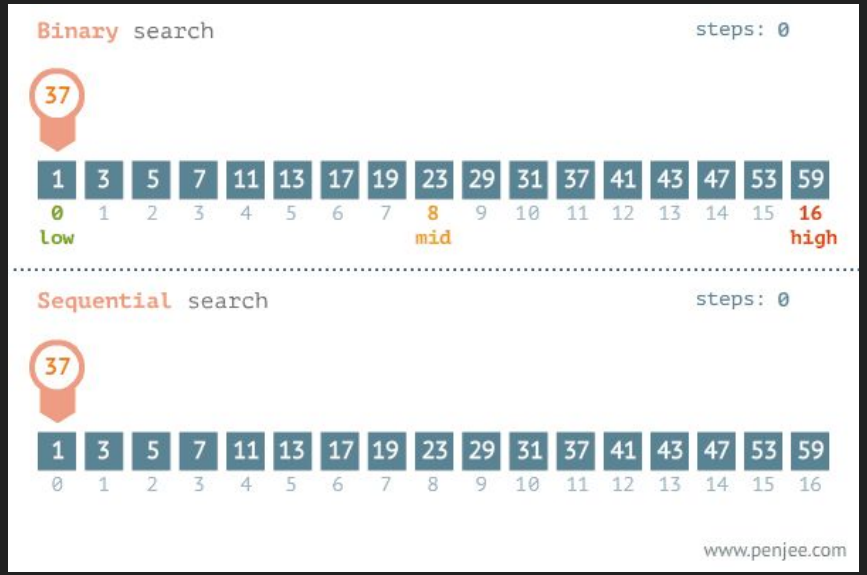
index	0	1	2	3	4	5	6	7	8	9
value	37	59	-7	20	2	88	-3	49	50	73

Ah 88! Found it!

What if we know the group of number is sorted?

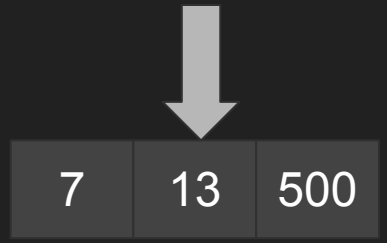
Binary Search - a better (than linear) search

If we know our data is sorted, if we know our data is sorted, if we know our data is sorted, if we know our data is sorted, if we know our data is sorted, if we know our data is sorted, if we know our data is sorted, if we know our data is sorted, if we know our data is sorted, we can totally do better than linear search!

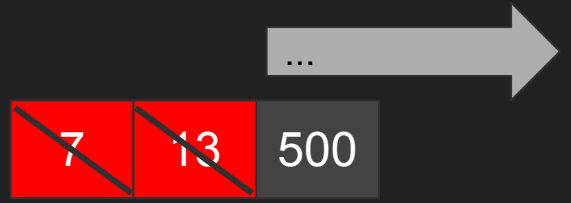


Binary Search - a better (than linear) search

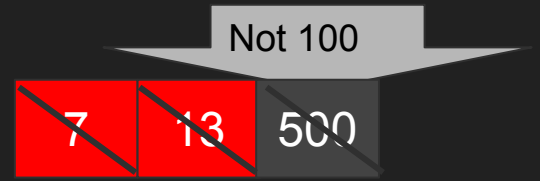
Step 1: Start from the middle of the list



Step 2: If the number we are looking for “100” is bigger than the middle, than keep everything to the right



Step 3: Repeat process until our ‘mid’ is either the number we are looking for, or we are down to 1 item. In this case, 100 is not found



Video: Linear Search and Binary Search

https://www.youtube.com/watch?v=FOwCCvHEfY0

Workshop time!

Starter Code Address:

https://bit.ly/3OST5G2

Workshop time!

Answer Code Address:

https://bit.ly/3P761rG

Modern Search Algorithms

Source code of Google Search Engine

nope, we can’t :)

Crawling? To find and get information

Google downloads text, images, and videos from pages it found on the internet with automated programs called crawlers.

Crawling - Google Spider

Google spider, also known as crawler, or Googlebot, is an algorithm that crawls websites and stores information for the search engine to index.

1. Work as a spider with “many legs”
2. Crawl into webpages
3. Download text, images, and videos from pages by “its legs”
4. Store the information for indexing

Indexing? To organize and analyze information

After a page is crawled, Google tries to understand what the page is about. This stage is called indexing and it includes processing and analyzing the content of webpages.

Serving search results? To serve th… c’mon it’s kinda obvious

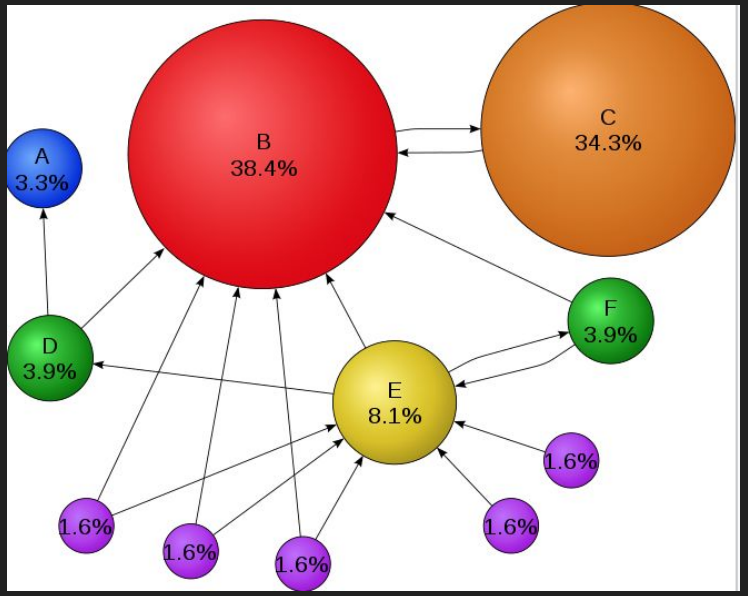
When a user enters a query, our machines search the index for matching pages and return the results we believe are the highest quality and most relevant to the user.

Relevancy is determined by hundreds of factors, which could include information such as the user's location, language, and device (desktop or phone).

Serving search results - PageRank

PageRank is an algorithm used by Google Search to rank websites in their search engine results. PageRank was named after Larry Page one of the founders of Google.

PageRank is a way of measuring the importance of website pages. PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is.



Video: How Google Search Works

https://www.youtube.com/watch?v=0eKVizvYSUQ

Three takeaways

1. Algorithms are just a bunch of steps to solve a certain problem.
2. Algorithms can save time, and time is XXX!
3. Algorithms can also be abused. (We'll talk more on Thursday!)

Bonus game: Fine the Golden Coin

Imagine it is your lucky day, and you are given 16 golden coins. Unfortunately 15 of the gold coins are fake. The fake gold coins all weigh 1 oz. but the 1 real gold weighs 1.0000000001 oz. You are also given one balancing scale that can precisely weigh each of the two sides. If one side is heavier than the other side, you will see the scale tip. How many weighing must you do to find the real coin given your algorithm?