Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

COMP2521 25T3 Analysis of Algorithms

Sim Mautner cs2521@cse.unsw.edu.au

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

- Program efficiency is critical for many applications:
 - Finance, robotics, games, database systems, ...
- We may want to compare programs to decide which one to use
- We may want to determine whether a program will be "fast enough"

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Mariatoria

Multiple Variables

Appendix

What determines how fast a program runs?

- The operating system?
- Compilers?
- Hardware?
 - E.g., CPU, GPU, cache
- Load on the machine?
- Most important: the data structures and algorithms used

Motivation Efficiency

Time

Searching

Empirical

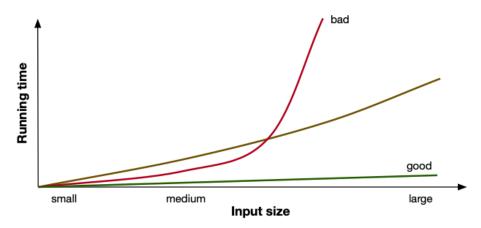
Analysis

Theoretical Analysis

Binary Search

Multiple Variables

- The running time of an algorithm tends to be a function of input size
- Typically: larger input ⇒ longer running time
 - Small inputs: fast running time, regardless of algorithm
 - Larger inputs: slower, but how much slower?



Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis Binary Search

Multiple Variables

- Best-case performance
 - Not very useful
 - Usually only occurs for specific types of input
- Average-case performance
 - Difficult; need to know how the program is used
- Worst-case performance
 - Most important; determines how long the program could possibly run

Time Complexity

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

Time complexity is the amount of time it takes to run an algorithm, as a function of the input size

Time Complexity

Motivation

Efficiency

Time Complexity

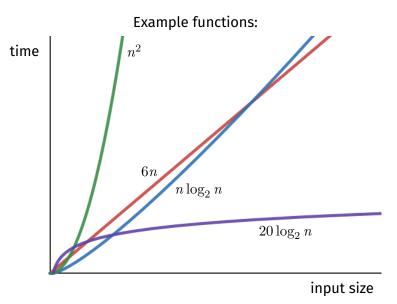
Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables



Analysing Time Complexity

Motivation

Efficiency

Time Complexity

Searching Empirical

Analysis
Theoretical
Analysis

Binary Search

Multiple Variables

Appendix

The time complexity of an algorithm can be analysed in two ways:

- Empirically: Measuring the time that a program implementing the algorithm takes to run
- Theoretically: Counting the number of operations or "steps" performed by the algorithm as a function of input size

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

The search problem:

Given an array of size n and a value, return the index containing the value if it exists, otherwise return -1.

[0]	[1]	[2]	[3]	[4]	[5]	[6]
2	16	11	1	9	4	15

Efficiency

Time Complexity

Searching

Empirical

Analysis Measuring ru

Demonstration

Theoretical Analysis

Binary Search

Diliary Searci

Multiple Variables

- 1 Write a program that implements the algorithm
- 2 Run the program with inputs of varying size and composition
- 3 Measure the running time of the algorithm
- Plot the results

Motivation Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring running time

Demonstrati Limitations

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

We can measure the running time of an algorithm using clock(3).

• The *clock()* function determines the amount of processor time used since the start of the process.

```
#include <time.h>
clock_t start = clock();
// algorithm code here...
clock_t end = clock();

double seconds = (double)(end - start) / CLOCKS_PER_SEC;
```

Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring running

Demonstrati

Theoretical Analysis

Binary Search

Multiple

Variables
Appendix

between machines, between languages ...so we're not interested in absolute time.

We are interested in the *relative* change as the input size increases

Absolute times will differ

Motivation Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring running

Demonstration

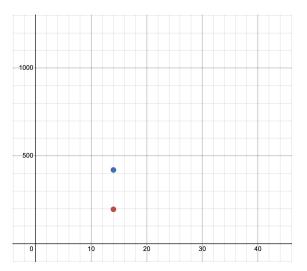
Theoretical Analysis

Binary Search

Multiple Variables

Appendix

Which algorithm is more efficient?



Motivation Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring running time

Demonstrati Limitations

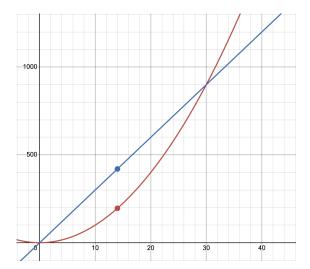
Theoretical Analysis

Binary Search

Multiple Variables

Appendix

Compare growth rates, not absolute times



Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring running

Demonstration

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

Let's empirically analyse the following search algorithm:

```
// Returns the index of the given value in the array
// if it exists, or -1 otherwise
int linearSearch(int arr[], int size, int val) {
    for (int i = 0; i < size; i++) {
        if (arr[i] == val) {
            return i;
        }
    }
}</pre>
```

Empirical Analysis

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring running time

Demonstration Limitations

Theoretical Analysis

Binary Search

Multiple

Variables

Appendix

```
// Returns the index of the given value in the array if it exists,
// or -1 otherwise
int linearSearch(int arr[], int size, int val) {
    for (int i = 0; i < size; i++) {
        if (arr[i] == val) {
            return i;
        }
    }
    return -1;
}</pre>
```

Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring runnir time

Demonstration

Theoretical Analysis

Binary Search

Multiple

Variables

Appendix

Sample results:

Input Size	Running Time
1,000,000	0.002
10,000,000	0.023
100,000,000	0.240
200,000,000	0.471
300,000,000	0.702
400,000,000	0.942
500,000,000	1.196
1,000,000,000	2.384
200,000,000 300,000,000 400,000,000 500,000,000	0.471 0.702 0.942 1.196

The worst-case running time of linear search grows linearly as the input size increases.

Empirical Analysis

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Measuring running time

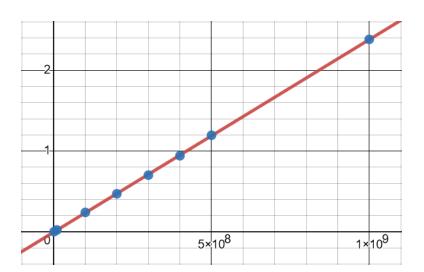
Demonstration

Theoretical Analysis

Binary Search

Multiple

Variables



Limitations of Empirical Analysis

Motivation

Efficiency

Time Complexit

Searching

Empirical Analysis

time

Limitations

Theoretical Analysis

Binary Search

Multiple Variables

- Requires implementation of algorithm
- Different choice of input data ⇒ different results
 - · Choosing good inputs is extremely important
- Timing results affected by runtime environment
 - E.g., load on the machine
- In order to compare two algorithms...
 - Need "comparable" implementation of each algorithm
 - Must use same inputs, same hardware, same O/S, same load

Theoretical Analysis

Motivation

Efficiency

Time Complexi

Searching Empirical

Analysis
Theoretical

Theoretic Analysis

Primitive operations
Asymptotic analysis
Big-Oh notation
Analysing complexity

Binary Search

Multiple Variables

- Uses high-level description of algorithm (pseudocode)
 - Can use the code if it is implemented already
- Characterises running time as a function of input size
- Allows us to evaluate the efficiency of the algorithm
 - Independent of the hardware/software environment

Efficiency

Time Complexit

Searching Empirical

Analysis Theoretic

Theoretical Analysis

Primitive operation
Asymptotic analysis

Binary Search

Multiple Variables

Annendix

- Pseudocode is a plain language description of the steps in an algorithm
- Uses structural conventions of a regular programming language
 - if statements, loops
- Omits language-specific details
 - variable declarations
 - allocating/freeing memory

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Pseudocode

Asymptotic analys

Analysing complex

Binary Search

Multiple Variables

Appendix

Pseudocode for linear search:

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical

Pseudocode

Asymptotic analysis

Big-Oh notation Analysing complexit

Binary Search

Multiple Variables

Annond

Every algorithm uses a core set of basic operations.

Examples:

- Assignment
- Indexing into an array
- Calling/returning from a function
- Evaluating an expression
- Increment/decrement

We call these operations **primitive** operations.

Assume that primitive operations take the same constant amount of time.

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Pseudocode

Primitive operations
Asymptotic analysis

Big-Oh notation
Analysing complexi

Binary Search

Multiple Variables

Appendix

How many primitive operations are performed by this line of code?

```
for (int i = 0; i < n; i++)
```

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical

Allatysis

Pseudocod

Primitive operations
Asymptotic analysis

Big-Oh notation

Analysing comple

Binary Search

Multiple Variables

Annendix

How many primitive operations are performed by this line of code?

The assignment i = 0 occurs 1 time The comparison i < n occurs n + 1 times The increment i++ occurs n times

Total: 1 + (n+1) + n primitive operations

Counting Primitive Operations

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical

Analysis

Primitive operations

Asymptotic analysis Big-Oh notation

Dinami Caaral

Binary Search

Multiple Variables

Appendix

By inspecting the pseudocode, we can determine the maximum number of primitive operations executed by an algorithm as a function of the input size.

Counting Primitive Operations

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Decudosos

Primitive operations
Asymptotic analysis
Big-Oh notation

Analysing complex

Binary Search

Multiple Variables

Appendix

Linear search requires 4n+3 primitive operations in the worst case.

If the time taken by a primitive operation is c, then the time taken by linear search in the worst case is c(4n+3).

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Pseudocode

Primitive operation

Asymptotic analysis

Big-Oh notation

Binary Search

Multiple Variables

Appendix

We are mainly interested in how the running time of an algorithm changes as the input size increases.

This is called the **asymptotic behaviour** of the running time.

Lower-Order Terms

Motivation

Efficiency

Time Complexit

Searching

Empirical Analysis

Theoretical

Pseudocode
Primitive operations
Asymptotic analysis
Big-Oh notation

Binary Search

Multiple Variables

Appendix

Asymptotic behaviour is not affected by lower-order terms.

- For example, suppose the running time of an algorithm is 4n + 100.
- As n increases, the lower-order term (i.e., 100) becomes less significant (i.e., becomes a smaller proportion of the running time)

Constant Factors

Motivation

Efficiency

Time Complexit

Searching

Empirical Analysis

Theoretical

Pseudocod

Primitive operations
Asymptotic analysis
Big-Oh notation

Binary Search

Multiple Variables

Appendix

Asymptotic behaviour is not affected by constant factors.

Example: Suppose the running time T(n) of an algorithm is n^2 .

• What happens when we double the input size?

$$T(2n) = (2n)^2$$
$$= 4n^2$$
$$= 4T(n)$$

When we double the input size, the time taken quadruples.

Constant Factors

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Pseudococ

Primitive opera

Asymptotic analysis

Big-Oh notation
Analysing comple

Binary Search

Multiple Variables

Appendix

Example: Now suppose the running time T(n) of an algorithm is $10n^2$.

Now what happens when we double the input size?

$$T(2n) = 10 \times (2n)^{2}$$
$$= 10 \times 4n^{2}$$
$$= 4 \times 10n^{2}$$
$$= 4T(n)$$

When we double the input size, the time taken also quadruples!

Efficiency

Searching

Empirical Analysis Theoretical

Asymptotic analysis

Binary Search

Multiple Variables

To summarise:

- Asymptotic behaviour is unaffected by lower-order terms
- Asymptotic behaviour is unaffected by constant factors

This means we can ignore lower-order terms and constant factors when characterising the asymptotic behaviour of an algorithm.

Examples:

- If T(n) = 100n + 500, ignoring lower-order terms and constant factors gives n
- If $T(n) = 5n^2 + 2n + 3$, ignoring lower-order terms and constant factors gives n^2

Motivation Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Pseudocode

Primitive operations

Asymptotic analysis

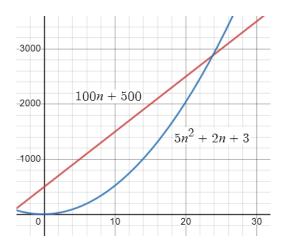
Big-Oh notation

Binary Search

Multiple Variables

Appendix

This also means that for sufficiently large inputs, the algorithm that has the running time with the highest-order term will always take longer.



Efficiency

Time Complexit

Searching

Empirical Analysis

Theoretical Analysis

Pseudocode

Assumptatic anal

Big-Oh notation

Analysing com

Binary Search

Multiple Variables

Appendix

Big-Oh notation is used to classify the asymptotic behaviour of an algorithm, and this is how we usually express time complexity in this course.

For example, linear search is O(n) in the worst case.

Efficiency

Time Complexit

Searching

Empirical Analysis

Theoretical

Primitive operation
Asymptotic analyse
Big-Oh notation

Binary Search

Multiple Variables

Appendix

Big-Oh notation allows us to easily compare the efficiency of algorithms

• For example, if algorithm A has a time complexity of O(n) and algorithm B has a time complexity of $O(n^2)$, then we can say that for sufficiently large inputs, algorithm A will perform better.

Efficiency

Time Complexit

Searching

Empirical Analysis

Theoretical

Pseudocode

Asymptotic analysis Big-Oh notation

Analysing complexi

Binary Search

Multiple Variables

Append

Formally, big-Oh is actually a notation used to describe the asymptotic relationship between functions.

Formally:

Given functions f(n) and g(n), we say that f(n) is O(g(n)) if:

- There are positive constants c and n_0 such that:
 - $f(n) \le c \cdot g(n)$ for all $n \ge n_0$

Informally:

Given functions f(n) and g(n), we say that f(n) is O(g(n)) if for sufficiently large n, f(n) is bounded above by some multiple of g(n).

Relatives of Big-Oh

All The Mathematics!

Motivation

Efficiency

Time Complexit

Searching

Empirical Analysis

Theoretical

Decudered

Primitive operatio

Asymptotic analys

Big-Oh notation

Analysing complex

Binary Search

Multiple Variables

Appendix

 $f(n) \mbox{ is } O(g(n)) \\ \mbox{if } f(n) \mbox{ is asymptotically less than or equal to } g(n) \\$

 $f(n) \text{ is } \Omega(g(n)) \\$ if f(n) is asymptotically greater than or equal to g(n)

$$f(n) \text{ is } \Theta(g(n)) \\ \text{if } f(n) \text{ is asymptotically equal to } g(n) \\$$

Motivation Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical

Pseudocode Primitive operation

Asymptotic analysis
Big-Oh notation

Analysing complexity

Binary Search

Multiple Variables

Appendix

Since time complexity is not affected by constant factors, instead of counting primitive operations, we can simply count line executions.

```
linearSearch(A, value):
    Input: array A of size n, value
   Output: index of value in A if it exists
            -1 otherwise
    for i from 0 up to n-1:
        if A[i] = value:
            return i
   return -1
                                 2n + 1 (total)
```

Worst-case time complexity: O(n)

Efficiency

Time Complexit

Searching Empirical

Analysis
Theoretical

Analysis

Primitive operations
Asymptotic analysis
Big-Oh notation

Analysing complexity

Binary Search

Multiple Variables

Appendi

To determine the worst-case time complexity of an algorithm:

- Determine the number of line executions performed in the worst case in terms of the input size
- Discard lower-order terms and constant factors
- The worst-case time complexity is then the big-Oh of the term that remains

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical

Analysis

Primitive operati

Asymptotic analysis

Big-Oh notation

Analysing complexity

Binary Search

billary Searci

Multiple Variables

Appendix

Commonly encountered functions in algorithm analysis:

• Constant: 1

• Logarithmic: $\log n$

• Linear: *n*

• N-Log-N: $n \log n$

• Quadratic: n^2

• Cubic: n^3

• Exponential: 2^n

• Factorial: n!

Common Functions

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Proudocodo

Primitive operat

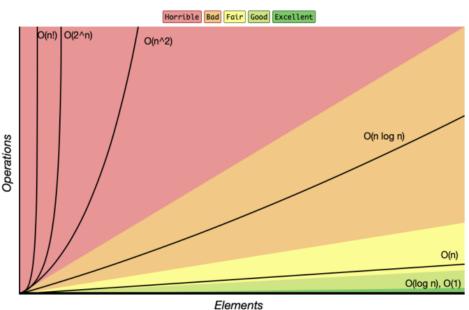
Asymptotic analys Big-Oh notation

Analysing complexity

Binary Search

Multiple Variables

Appendix



Back to Linear Search

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

Linear search requires 4n+3 primitive operations in the worst case.

Therefore, linear search is O(n) in the worst case.

Searching in a Sorted Array

Motivation Efficiency

Time

Searching

Empirical Analysis Theoretical

Binary Search

Dillary ocur

Multiple Variables

Appendix

Is there a faster algorithm for searching an array?

Yes... if the array is sorted.

[0]	[1]	[2]	[3]	[4]	[5]	[6]
1	2	4	9	11	15	16

Let's start in the middle.

- If a[N/2] = val, we found val; we're done!
- Otherwise, we split the array:

... if
$$val < a[N/2]$$
, we search the left half ($a[0]$ to $a[(N/2) - 1)$)

... if val > a[N/2], we search the right half (a[(N/2) + 1)] to a[N-1])

Efficiency

Time Complexity

Searching

Empirical

Analysis
Theoretical
Analysis

Binary Search

Multiple Variables

Appendix

Binary search is a more efficient search algorithm for **sorted arrays**:

```
int binarySearch(int arr[], int size, int val) {
    int lo = 0;
    int hi = size - 1;
    while (lo <= hi) {</pre>
        int mid = (lo + hi) / 2;
        if (val < arr[mid]) {</pre>
            hi = mid - 1;
        } else if (val > arr[mid]) {
            lo = mid + 1;
        } else {
            return mid;
    return -1;
```

Motivation Efficiency

Time Complexity

Searching

Empirical Analysis

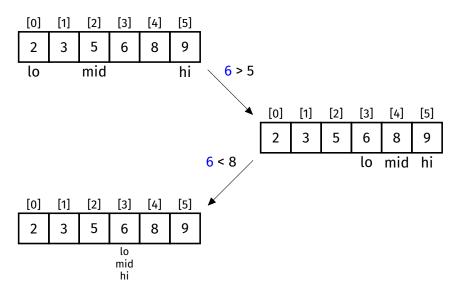
Theoretical Analysis Binary Search

Multiple

Variables

Appendix

Successful search for 6:





Binary Search

Example

Motivation Efficiency

Time Complexity

Searching

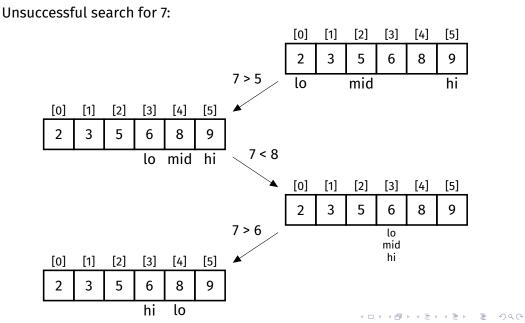
Empirical Analysis Theoretical

Analysis
Binary Search

Multiple

Variables

Appendix



Time Complexity

Searching Empirical

Analysis
Theoretical

Analysis

Binary Search

Multiple Variables

Appendix

How many iterations of the loop?

- Best case: 1 iteration
 - Item is found right away
- Worst case: $\log_2 n$ iterations
 - Item does not exist
 - Every iteration, the size of the subarray being searched is halved

Thus, binary search is $O(\log_2 n)$ or simply $O(\log n)$

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

$$O(\log_2 n) = O(\log n)$$

Why drop the base?

According to the change of base formula:

$$\log_a \frac{n}{n} = \frac{\log_b \frac{n}{n}}{\log_b a}$$

If a and b are constants, $\log_a n$ and $\log_b n$ differ by a constant factor

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

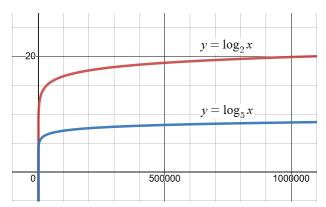
Binary Search

Multiple

Variables Appendix

For example:

$$\begin{aligned} \log_2 n &= \frac{\log_5 n}{\log_5 2} \\ &\approx 2.32193 \log_5 n \end{aligned}$$



Multiple Variables

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

What if an algorithm takes multiple arrays as input?

If there is no constraint on the relative sizes of the arrays, their sizes would be given as two variables, usually n and m

Multiple Variables

Motivation

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

Example time complexities with two variables:

$$O(n+m)$$

$$O(\max(n, m))$$

$$O(\min(n, m))$$

$$O(n \log m)$$

$$O(n\log m + m\log n)$$

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

Problem:

Given two arrays, where each array contains no repeats, find the number of elements in common

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix

```
numCommonElements(A, B):
          array A of size n
    Input:
            array B of size m
    Output: number of elements in common
    numCommon = 0
    for i from 0 up to n-1:
        for j from 0 up to m-1:
            if A[i] = B[j]:
                numCommon = numCommon + 1
    return numCommon
```

Time complexity: O(nm)

Efficiency

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

Appendix

Efficiency

Time Complexity

Searching Empirical

Analysis Theoretical

Analysis

Binary Search

Multiple Variables

Appendix Exercise Analysis Examp If I know my algorithm is quadratic (i.e., $O(n^2)$), and I know that for a dataset of 1000 items, it takes 1.2 seconds to run ...

- how long for 2000?
- how long for 10,000?
- how long for 100,000?
- how long for 1,000,000?

(answers on the next slide)

Efficiency

Time Complexity

Searching Empirical

Analysis
Theoretical

Analysis

Binary Search

Multiple Variables

Appendix Exercise

If I know my algorithm is quadratic (i.e., $O(n^2)$), and I know that for a dataset of 1000 items, it takes 1.2 seconds to run ...

- how long for 2000?4.8 seconds
- how long for 10,000?
 120 seconds (2 mins)
- how long for 100,000?
 12000 seconds (3.3 hours)
- how long for 1,000,000?
 1200000 seconds (13.9 days)

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
for (int i = 1; i <= n; i++) {
    for (int j = 1; j <= n; j++) {
        for (int k = 1; k <= n; k++) {
            // constant-time statement
        }
    }
}</pre>
```

Time complexity?

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
for (int i = 1; i <= n; i++) {
    for (int j = 1; j <= n; j++) {
        for (int k = 1; k <= n; k++) {
            // constant-time statement
        }
    }
}</pre>
```

Time complexity? $O(n^3)$

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
for (int i = 1; i <= n; i++) {
    for (int j = 1; j <= i; j++) {
        // constant-time statement
    }
}</pre>
```

Time complexity?

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
for (int i = 1; i <= n; i++) {
    for (int j = 1; j <= i; j++) {
        // constant-time statement
    }
}</pre>
```

Time complexity? $O(n^2)$

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
for (int i = 1; i <= n; i = i * 2) {
    // constant-time statement
}</pre>
```

Time complexity?

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
for (int i = 1; i <= n; i = i * 2) {
    // constant-time statement
}</pre>
```

Time complexity? $O(\log n)$

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
int p = 0;
for (int i = 1; p <= n; i++) {
    p = p + i;
}</pre>
```

Time complexity?

```
Motivation
```

Time Complexity

Searching

Empirical Analysis

Theoretical Analysis

Binary Search

Multiple Variables

Appendix Exercise

Analysis Examples

```
int p = 0;
for (int i = 1; p <= n; i++) {
    p = p + i;
}</pre>
```

Time complexity? $O(\sqrt{n})$