

# **Data Structures and Algorithms 2**

## **Chapter 2**

## **Algorithm Analysis**

**Dr. Fouzia ANGUEL**  
**2<sup>nd</sup> year / S3**

**September 2024 – January 2025**

# Course Outline

- ❖ Algorithms efficiency
  - Machine-dependent vs Machine-independent
- ❖ Function ordering
  - Order of growth
  - Weak Order;
  - Landau symbols Big-Oh ; Big-omega ; big theta and Little-oh .
- ❖ Algorithm complexity analysis
  - Rules for complexity analysis
  - Analysis of various types of algorithms
  - Master Theorem

# Algorithm Efficiency

## Example : Shortest path problem

- A city has  $n$  view points
- Buses move from one view point to another
- A bus driver wishes to follow the shortest path (travel time wise).
- Every view point is connected to another by a road.
- However, some roads are less congested than others.
- Also, roads are one-way, i.e., the road from view point 1 to 2, is different from that from view point 2 to 1.

# Algorithm Efficiency

Example : Shortest path problem

How to find the shortest path between any two pairs?

## → Naïve approach

- ◆ List all the paths between a given pair of view points
- ◆ Compute the travel time for each.
- ◆ Choose the shortest one.

How many paths are there between any two view points  
(without revisits)?

$$n! \approx (n/e)^n$$

→ It will be impossible to run the algorithm for  $n = 30$

# Algorithm efficiency

- Run time in the computer is Machine dependent

**Example** : Need to multiply two positive integers a and b

**Subroutine 1**: Multiply a and b

**Subroutine 2**:  $V = a, \quad W = b$

While  $W > 1$

$V \rightarrow V + a; W \rightarrow W - 1$

Output V

# Algorithm efficiency

First subroutine has 1 multiplication.

Second has  $b$  additions and subtractions.

For some architectures, 1 multiplication is more expensive than  $b$  additions and subtractions.

Ideally, we would like to program all choices and run all of them in the machine we are going to use and find which is efficient!

# Machine Independent Analysis

We assume that every **basic operation** takes **constant** time

Example **Basic** Operations:

Addition, Subtraction, Multiplication, Memory Access

**Non-basic** Operations:

Sorting, Searching

**Efficiency** of an algorithm is the **number of basic operations** it performs

We do not distinguish between the basic operations.

Subroutine 1 uses **1** basic operation (\*)

Subroutine 2 uses **2b** basic operations (+, -)

Subroutine **1** is **more efficient**.

This measure is good for all large input sizes

In fact, we will not worry about the **exact values**, but will look at “**broad classes**” of values.

Let there be **n inputs**.

If an algorithm needs **n** basic operations and another needs **2n** basic operations, we will consider them to be in the **same efficiency category**.

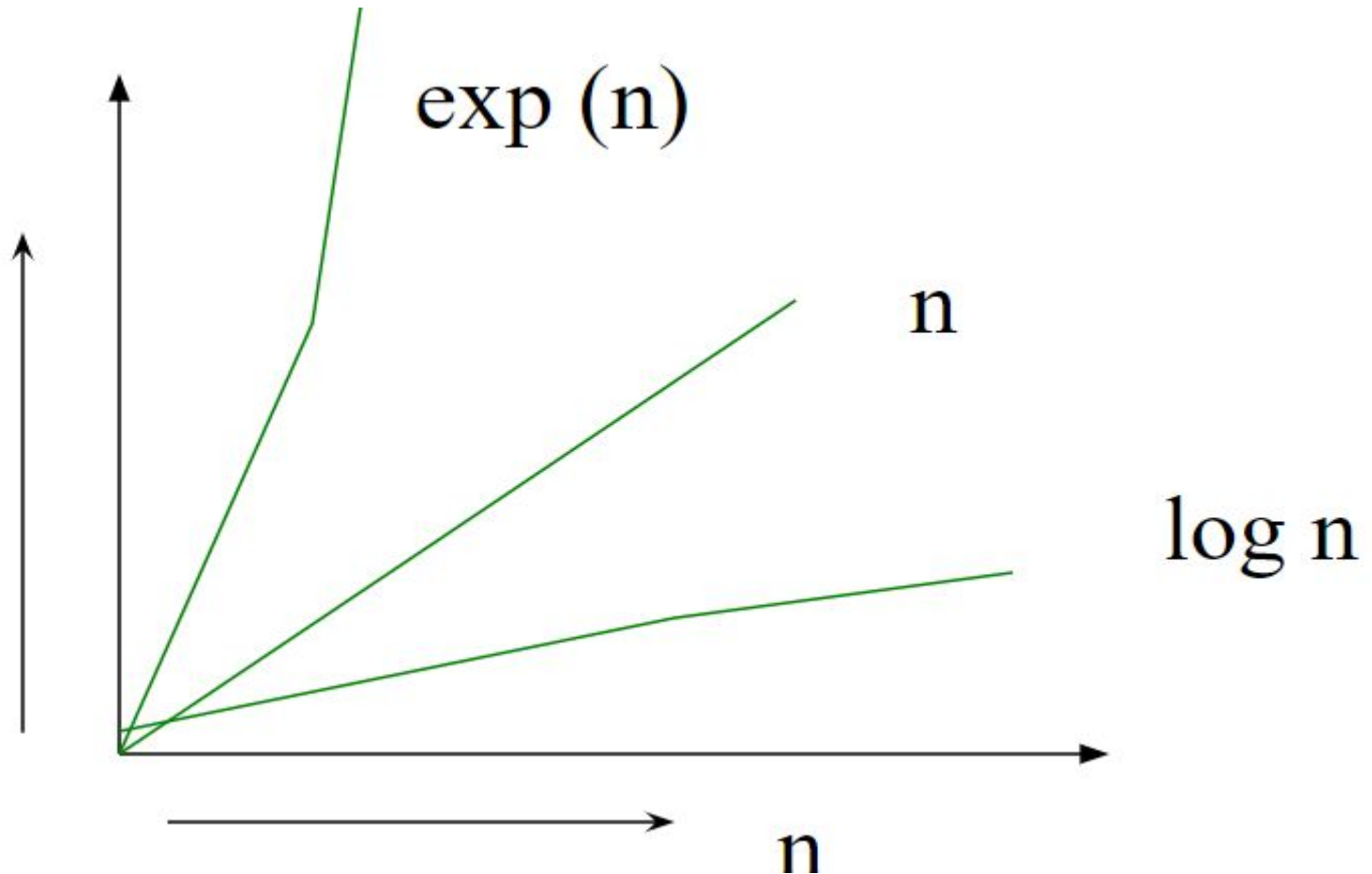
However, we distinguish between **exp(n), n, log(n)**



# Function Ordering

## Order of Increase(order of growth)

We worry about the speed of our algorithms for large input sizes.

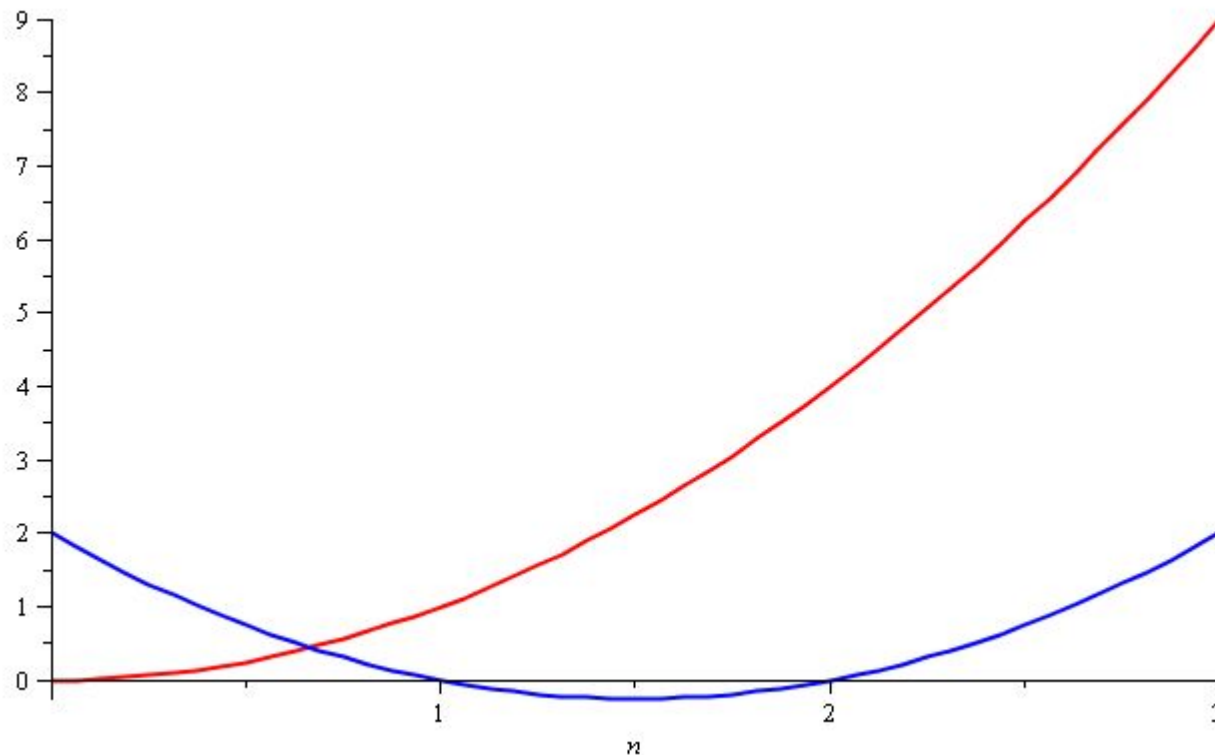


# Quadratic Growth

Consider the two functions

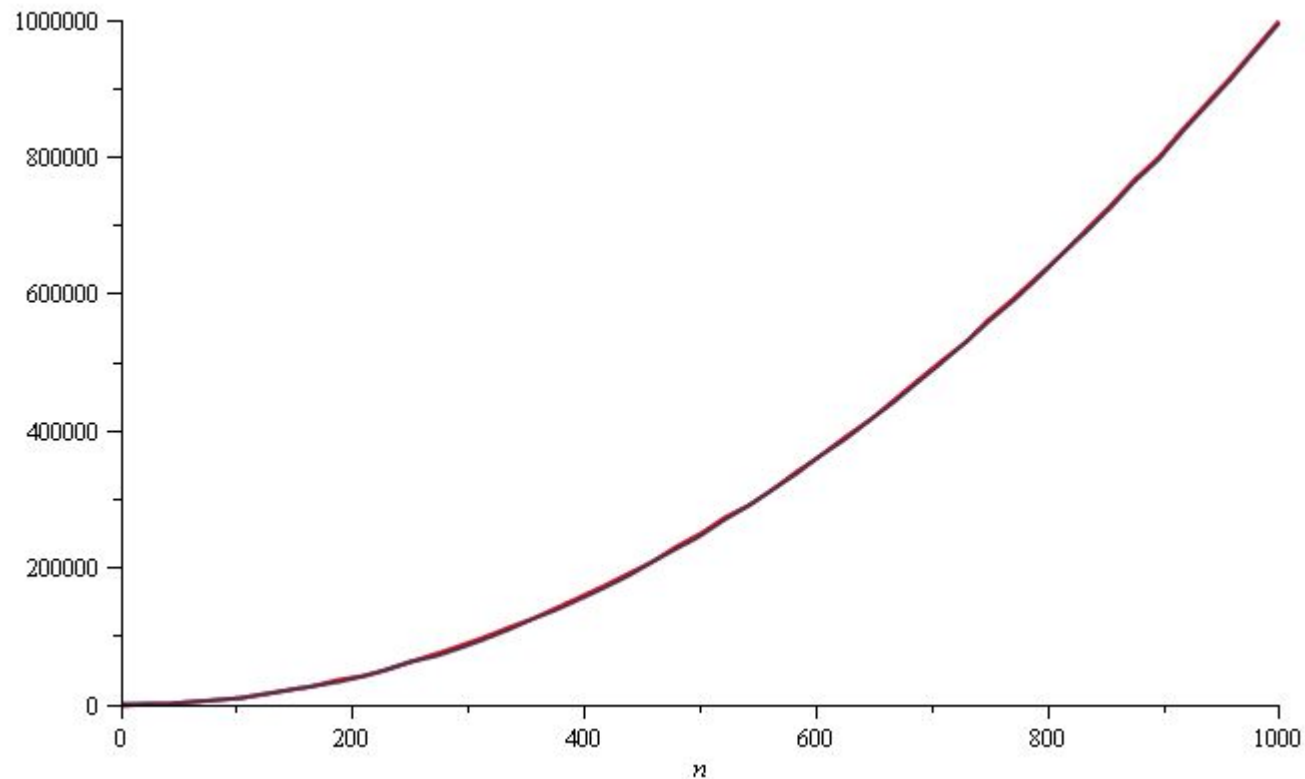
$$f(n) = n^2 \text{ and } g(n) = n^2 - 3n + 2$$

Around  $n = 0$ , they look very different



# Quadratic Growth

Yet on the range  $n = [0, 1000]$ , they are (relatively) indistinguishable:



## Quadratic Growth

The absolute difference is large, for example,

$$f(1000) = 1\,000\,000$$

$$g(1000) = 997\,002$$

but the relative difference is very small

$$\left| \frac{f(1000) - g(1000)}{f(1000)} \right| = 0.002998 < 0.3\%$$

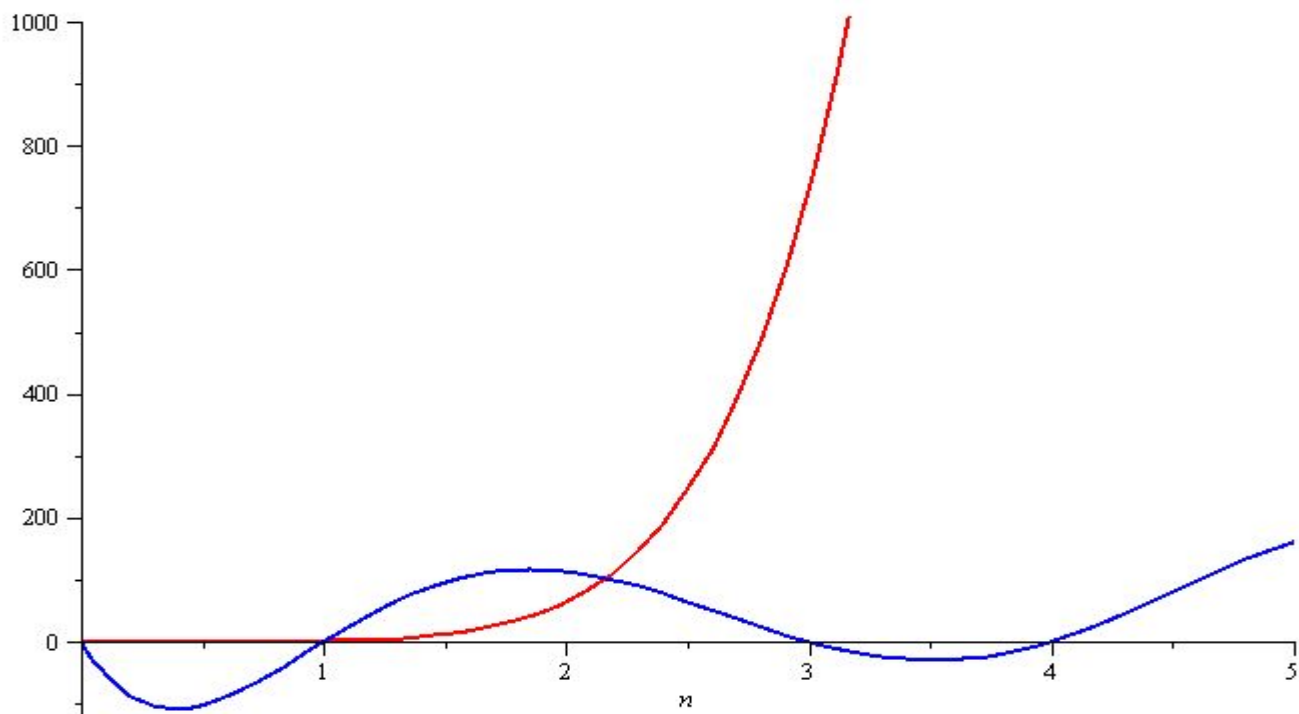
and this difference goes to zero as  $n \rightarrow \infty$

# Polynomial Growth

To demonstrate with another example,

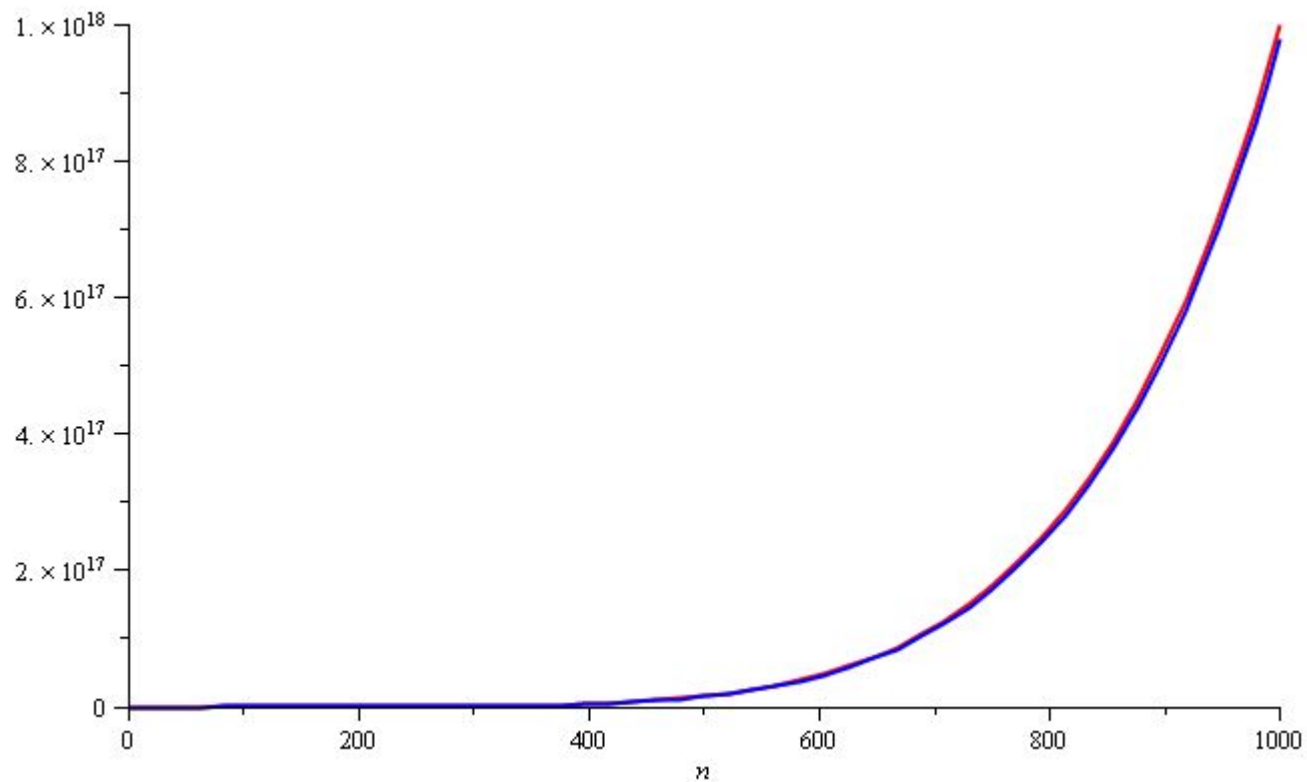
$$f(n) = n^6 \quad \text{and} \quad g(n) = n^6 - 23n^5 + 193n^4 - 729n^3 + 1206n^2 - 648n$$

Around  $n = 0$ , they are very different



# Polynomial Growth

Still, around  $n = 1000$ , the relative difference is less than 3%



# Polynomial Growth

The justification for both pairs of polynomials being similar is that, in both cases, they each had the **same leading term**:

$n^2$  in the first case,  $n^6$  in the second

Suppose however, that the coefficients of the leading terms were different

- In this case, both functions would exhibit the same rate of growth, however, one would always be proportionally larger

# Weak ordering

Consider the following definitions:

- We will consider two functions to be equivalent,  $f \sim g$ , if

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c \quad \text{where} \quad 0 < c < \infty$$

- We will state that  $f < g$  if  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$

For functions we are interested in, these define a *weak ordering*



# Weak ordering

17

$f$  and  $g$  are functions from the set of natural numbers to itself.

Let  $f(n)$  and  $g(n)$  describe the run-time of two algorithms

- If  $f(n) \sim g(n)$ , then it is always possible to improve the performance of one function over the other by purchasing a **faster** computer
- If  $f(n) < g(n)$ , then you can **never** purchase a computer **fast enough** so that the second function always runs in less time than the first

Note that for **small values** of  $n$ , it may be reasonable to use an algorithm that is asymptotically more expensive, but we will consider these on a **one-by-one** basis

# Function orders “Landau Symbols”

we will make some assumptions:

- Our functions will describe the time or memory required to solve a problem of size  $n$
- We are restricting to certain functions :
  - They are defined for  $n \geq 0$
  - They are strictly positive for all  $n$ 
    - In fact,  $f(n) > c$  for some value  $c > 0$
    - That is, any problem requires at least one instruction and byte
  - They are increasing (monotonic increasing)

# Function orders “Landau Symbols”

## Big Oh Notation

A function  **$f(n)$**  is  **$O(g(n))$**  if the **rate of growth** of  $f(n)$  is not greater (not faster) than that of  $g(n)$ .

### Definition 1

**$f(n) = O(g(n))$**  if there are a number  **$n_0$**  and a nonnegative  **$c$**  such that

for all  $n \geq n_0$ ,  $0 \leq f(n) \leq cg(n)$ .

If  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)}$  **exists** and is **finite**, then  $f(n)$  is  **$O(g(n))$** .

Intuitively, (not exactly)  $f(n)$  is  $O(g(n))$  means  $f(n) \leq g(n)$  for all  $n$  beyond some value  $n_0$ ; i.e.  $g(n)$  is an **upper bound** for  $f(n)$ .

# Function orders “Landau Symbols”

## Example Functions

$\text{sqrt}(n)$  ,  $n$ ,  $2n$ ,  $\ln n$ ,  $\exp(n)$ ,  $n + \text{sqrt}(n)$  ,  $n + n^2$

$$\lim_{n \rightarrow \infty} \text{sqrt}(n) / n = 0, \quad \text{sqrt}(n) \text{ is } O(n)$$

$$\lim_{n \rightarrow \infty} n / \text{sqrt}(n) = \text{infinity}, \quad n \text{ is not } O(\text{sqrt}(n))$$

$$\lim_{n \rightarrow \infty} n / 2n = 1/2, \quad n \text{ is } O(2n)$$

$$\lim_{n \rightarrow \infty} 2n / n = 2, \quad 2n \text{ is } O(n)$$

$$\lim_{n \rightarrow \infty} \ln(n) / n = 0,$$

$\ln(n)$  is  $O(n)$

$$\lim_{n \rightarrow \infty} n / \ln(n) = \text{infinity},$$

$n$  is not  $O(\ln(n))$

$$\lim_{n \rightarrow \infty} \exp(n) / n = \text{infinity},$$

$\exp(n)$  is not  $O(n)$

$$\lim_{n \rightarrow \infty} n / \exp(n) = 0,$$

$n$  is  $O(\exp(n))$

$$\lim_{n \rightarrow \infty} (n + \sqrt{n}) / n = 1,$$

$n + \sqrt{n}$  is  $O(n)$

$$\lim_{n \rightarrow \infty} n / (\sqrt{n} + n) = 1,$$

$n$  is  $O(n + \sqrt{n})$

$$\lim_{n \rightarrow \infty} (n + n^2) / n = \text{infinity},$$

$n + n^2$  is not  $O(n)$

$$\lim_{n \rightarrow \infty} n / (n + n^2) = 0,$$

$n$  is  $O(n + n^2)$

# Implication of big-Oh notation

Suppose we know that our algorithm uses at most  $O(f(n))$  basic steps for any  $n$  inputs, and  $n$  is sufficiently large,

- then we know that our algorithm will terminate after executing at most  $f(n)$  basic steps.
- We know that a basic step takes a constant time in a machine.

Hence, our algorithm will terminate in a constant time  $f(n)$  units of time, for all large  $n$ .

# Function orders “Landau Symbols”

23

## $\Omega$ “Omega” Notation

Now a lower bound notation,  $\Omega$

### Definition 2

$f(n) = \Omega(g(n))$  if there are a number  $n_0$  and a nonnegative  $c$  such that

for all  $n \geq n_0$ ,  $f(n) \geq cg(n)$ .

If  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} > 0$  if  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)}$  exists

We say  $g(n)$  is a **lower bound** on  $f(n)$ ,  
i.e. no matter what specific inputs we have, the algorithm  
will not run faster than this lower bound.

Suppose, an algorithm has complexity  $\Omega(f(n))$ . This means that there exists a **positive constant  $c$**  such that for **all sufficiently large  $n$** , there exists **at least one input** for which the algorithm consumes **at least  $c \cdot f(n)$**  steps.

# Function orders “Landau Symbols”

## $\Theta$ “theta ” Notation

### Definition 3

**$f(n) = \Theta(g(n))$**  if and only if  $f(n)$  is  $O(g(n))$  and  $\Omega(g(n))$

**$f(n) = \Theta(g(n))$**  if there exist positive  $n_o$ ,  $c_1$ , and  $c_2$  such that

$$c_1 g(n) \leq f(n) \leq c_2 g(n) \quad \text{whenever } n \geq n_o$$

- $\Theta(g(n))$  is “asymptotic equality”

- $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)}$  is a **finite, positive constant**, if it exists

A function  $f(n)$  is  $\Theta(g(n))$  if The function  $f(n)$  has a **rate of growth equal** to that of  $g(n)$  .  $\Theta$  represents a **tight bound** in asymptotic analysis, which means it captures both the **upper** and **lower** bounds of a function's growth.



# Function orders “Landau Symbols”

## Little-oh Notation

### Definition 4

**$f(n) = o(g(n))$**  if for all positive constant  $c$ , there exists an  $n_0$  such that :

$$f(n) < cg(n) \text{ when } n > n_0$$

Less formally,  **$f(n) = o(g(n))$**  if  $f(n) = O(g(n))$  and  $f(n) \neq \theta(g(n))$ .

“asymptotic strict inequality”

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0 \quad \text{if} \quad \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} \text{ exists}$$

# Function orders “Landau Symbols”

Suppose that  $f(n)$  and  $g(n)$  satisfy  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c$

If  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c$  where  $0 < c < \infty$ , it follows that  $f(n) = \Theta(g(n))$

If  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c$  where  $0 \leq c < \infty$ , it follows that  $f(n) = \mathbf{O}(g(n))$

If  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$ , it follows that  $f(n) = o(g(n))$

# Function orders “Landau Symbols”

$$f(n) = \mathbf{o}(g(n))$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$$

$$f(n) = \mathbf{O}(g(n))$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty$$

$$f(n) = \mathbf{\Theta}(g(n))$$

$$0 < \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty$$

$$f(n) = \mathbf{\Omega}(g(n))$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} > 0$$

# Function orders “Landau Symbols”

## Terminology

Asymptotically less than or equal to  $\mathbf{O}$

Asymptotically greater than or equal to  $\mathbf{\Omega}$

Asymptotically equal to  $\mathbf{\theta}$

Asymptotically strictly less  $\mathbf{o}$

# Little-o as a Weak Ordering

We can show that, for example

$$\ln(n) = o(n^p) \quad \text{for any } p > 0$$

Proof: Using l'Hôpital's rule.

If you are attempting to determine  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)}$

but both  $\lim_{n \rightarrow \infty} f(n) = \lim_{n \rightarrow \infty} g(n) = \infty$ , it follows

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \lim_{n \rightarrow \infty} \frac{f^{(1)}(n)}{g^{(1)}(n)}$$

Repeat as necessary...

Note: the  $k^{\text{th}}$  derivative will always be shown as  $f^{(k)}(n)$

$$\lim_{n \rightarrow \infty} \frac{\ln(n)}{n^p} = \lim_{n \rightarrow \infty} \frac{1/n}{pn^{p-1}} = \lim_{n \rightarrow \infty} \frac{1}{pn^p} = \frac{1}{p} \lim_{n \rightarrow \infty} n^{-p} = 0$$

# Big- $\Theta$ as an Equivalence Relation

If we look at the first relationship, we notice that  $f(n) = \Theta(g(n))$  seems to describe an equivalence relation:

1.  $f(n) = \Theta(g(n))$  if and only if  $g(n) = \Theta(f(n))$
2.  $f(n) = \Theta(f(n))$
3. If  $f(n) = \Theta(g(n))$  and  $g(n) = \Theta(h(n))$ , it follows that  $f(n) = \Theta(h(n))$

Consequently, we can group all functions into **equivalence classes**, where all functions within one class are big-theta  $\Theta$  of each other

# Big- $\Theta$ as an Equivalence Relation

For example, all of

$$\begin{array}{ccc} n^2 & 100000 n^2 - 4 n + 19 & n^2 + 1000000 \\ 323 n^2 - 4 n \ln(n) + 43 n + 10 & & 42n^2 + 32 \\ & n^2 + 61 n \ln^2(n) + 7n + 14 \ln^3(n) + \ln(n) & \end{array}$$

are big- $\Theta$  of each other

*E.g.*,  $42n^2 + 32 = \Theta( 323 n^2 - 4 n \ln(n) + 43 n + 10 )$

We will select just one element to represent the entire class of these functions:  **$n^2$**

- We could choose any function, but this is the simplest

# Function orders “Landau Symbols”

## Terminology

The most common classes are given names:

$\Theta(1)$	constant
$\Theta(\ln(n))$ or $\Theta(\log(n))$	logarithmic
$\Theta(n)$	linear
$\Theta(n \ln(n))$	“ $n \log n$ ”
$\Theta(n^2)$	quadratic
$\Theta(n^3)$	cubic
$2^n, e^n, 4^n, \dots$	exponential

Recall that all logarithms are scalar multiples of each other

Therefore  $\log_b(n) = \Theta(\ln(n))$  for any base  $b$



# Function orders “Landau Symbols”

## Example Functions

$\text{sqrt}(n)$  ,  $n$ ,  $2n$ ,  $\ln n$ ,  $\exp(n)$ ,  $n + \text{sqrt}(n)$  ,  $n + n^2$

$$\lim_{n \rightarrow \infty} \text{sqrt}(n) / n = 0,$$

$\text{sqrt}(n)$  is  $o(n)$  and  $O(n)$

$$\lim_{n \rightarrow \infty} n / \text{sqrt}(n) = \text{infinity},$$

$n$  is  $\Omega(\text{sqrt}(n))$

$$\lim_{n \rightarrow \infty} n / 2n = 1/2,$$

$n$  is  $\theta(2n)$

$$\lim_{n \rightarrow \infty} 2n / n = 2,$$

$2n$  is  $\theta(n)$

$$\lim_{n \rightarrow \infty} \ln(n) / n = 0,$$

$\ln(n)$  is  $o(n)$

$$\lim_{n \rightarrow \infty} n / \ln(n) = \text{infinity},$$

$n$  is  $\Omega(\ln(n))$

$$\lim_{n \rightarrow \infty} \exp(n) / n = \text{infinity},$$

$\exp(n)$  is  $\Omega(n)$

$$\lim_{n \rightarrow \infty} n / \exp(n) = 0,$$

$n$  is  $o(\exp(n))$

$$\lim_{n \rightarrow \infty} (n + \sqrt{n}) / n = 1,$$

$n + \sqrt{n}$  is  $\theta(n)$

$$\lim_{n \rightarrow \infty} n / (\sqrt{n} + n) = 1,$$

$n$  is  $\theta(n + \sqrt{n})$ ,

$$\lim_{n \rightarrow \infty} (n + n^2) / n = \text{infinity},$$

$n + n^2$  is  $\Omega(n)$

$$\lim_{n \rightarrow \infty} n / (n + n^2) = 0,$$

$n$  is  $o(n + n^2)$

# Algorithms Analysis

An algorithm is said to have **polynomial** *time complexity* if its run-time may be described by  $O(n^d)$  for some fixed  $d \geq 0$

- We will consider such algorithms to be **efficient**

**Problems** that have no known polynomial-time algorithms are said to be **intractable**

- *Traveling salesman problem*: find the shortest path that visits  $n$  cities
- Best run time:  $\Theta(n^2 2^n)$

## Complexity of a Problem Vs Algorithm

A **problem** is  $O(f(n))$  means there is some  $O(f(n))$  algorithm to solve the problem.

A **problem** is  $\Omega(f(n))$  means every algorithm that can solve the problem is  $\Omega(f(n))$

# Rules for arithmetic with big-O symbols

## Rule 1

If  $T_1(n) = O(f(n))$  and  $T_2(n) = O(g(n))$ , then

**(a)**  $T_1(n) + T_2(n) = O(f(n) + g(n))$  (intuitively and less formally it is  $O(\max(f(n), g(n)))$ ),

**(b)**  $T_1(n) * T_2(n) = O(f(n) * g(n))$ .

## Rule 2

If  $T(n)$  is a polynomial of degree  $k$ , then  $T(n) = \theta(n^k)$ .

## Rule 3

- $\log^k n = O(n)$  for any constant  $k$ .

This tells us that logarithms grow very slowly.

# Rules for arithmetic with big-O symbols

38

## Rule 4

If  $f(n) = \mathbf{O}(g(n))$ , then

$$c * f(n) = \mathbf{O}(g(n)) \text{ for any constant } c.$$

## Rule 5

If  $f_1(n) = \mathbf{O}(g(n))$  but  $f_2(n) = \mathbf{o}(g(n))$ , then

$$f_1(n) + f_2(n) = \mathbf{O}(g(n)).$$

## Rule 6

If  $f(n) = \mathbf{O}(g(n))$ , and  $g(n) = \mathbf{o}(h(n))$ , then

$$f(n) = \mathbf{o}(h(n)). \quad (\text{complexity of } f \circ g)$$

These are not all of the rules, but they're enough for most purposes.

# Algorithm Complexity Analysis

39

- Three cases for which the efficiency of algorithms has to be determined :
  - *worst case* : is when an algorithm requires a maximum number of steps,
  - the *best case* : is when the number of steps is the smallest, and
  - the *average case* falls between these extremes.
- We define  $T_{avg}(N)$  and  $T_{worst}(N)$ , as the average and worst-case running time, resp., used by an algorithm on input of size  $N$ . Clearly,  $T_{avg}(N) \leq T_{worst}(N)$ .
- Average-case performance often reflects *typical behavior*
- Worst-case performance represents a *guarantee for performance* on *any possible* input.
- The best-case performance of an algorithm is of little interest: does not represent the typical behavior. It is occasionally analyzed.

# Algorithm Complexity Analysis

## Example

1. <code>diff = sum = 0;</code>	• Line 1 takes 2 basic steps
2. <code>for (k=0: k &lt; N; k++)</code>	• in every iteration of first loop
3. <code>sum → sum + 1;</code>	Line 3 takes 2 basic steps.
4. <code>diff → diff - 1;</code>	Line 4 takes 2 basic steps
	First loop runs N times
5. <code>for (k=0: k &lt; 3N; k++)</code>	• in every iteration of second
6. <code>sum → sum - 1;</code>	Line 6 loop takes 2 basic step
	• Second loop runs for 3N times

**Overall,  $2 + 4N + 6N$  steps** ( without counting the test and increment operations for each iteration in the the two loops)

**This is  $O(N)$**



# Algorithm Complexity Analysis

## General Rules

### Rule 1- Consecutive Statements:

This just add , which means that the maximum is that counts .

### Rule 2- Complexity of a loop:

The running time of a loop is at most the running time of the statements inside the loop (including tests) times the number of iterations.

**$O(\text{Number of iterations in a loop} * \text{maximum complexity of each iteration})$**

### Rule 3- Nested Loops:

The running time of a group of nested loops is the running time inside a group of nested loops multiplied by the product of the sizes of all the loops .

**$\text{Complexity of an outer loop} = \text{number of iterations in this loop} * \text{complexity of inner loop, etc.}$**

# Algorithm Complexity Analysis

42

## Example

<pre>1. sum = 0; 2. for (i=0; i &lt; N; i++) 3.     for (j=0; j &lt; N; j++) 4.         sum → sum + 1;</pre>	<p>Outer loop: <math>N</math> iterations</p> <p>Inner loop: <math>O(N)</math></p> <p><b>Overall: <math>O(N^2)</math></b></p>
<pre>1. for (i=0; i &lt; N; i++) 2.     a[i] = 0; 3. for (i=0; i &lt; N; i++) 4.     for (j=0; j &lt; N; j++) 5.         a[i] = a[j] + i+j ;</pre>	<p>First loop <math>O(N)</math></p> <p>Inner loop: <math>O(N)</math></p> <p>Outer loop: <math>N</math> iterations</p> <p><b>Overall: <math>O(N) + O(N^2)</math> So <math>O(N^2)</math></b></p>

# Algorithm Complexity Analysis

## General Rules

### Rule 3- If else

For the fragment

If (Condition)

S1

Else

Maximum of the two complexities

S2

The running time of an if/else statement is never more than the running time of **the test plus the larger** of the running times of S1 and S2

If (yes)

print(1,2,...**1000N**)

else print(1,2,...**N<sup>2</sup>**)

overall **O(N<sup>2</sup>)**

the basic strategy is analyzing from the inside (or deepest part) out. If there are function calls, these must be analyzed first.

# Algorithm Complexity Analysis

## Analysis of recursion

44

- If the recursion is really just a for loop , the analysis is usually trivial .

```
Long factorial (int n) {  
    if (n <= 1) O(N)  
        return 1;  
    else  
        return n*factorial(n - 1);  
}
```

- However, if the recursion is properly used . The analysis will involve a recurrence relation.

# Algorithm Complexity Analysis

## Analysis of recursion

45

- Suppose we have the following code :

```
Long fib (int n) {
```

```
1.  if (n <= 1)
```

```
2.      return 1;
```

```
    else
```

```
3.      return fib(n - 1) + fib(n - 2);
```

```
}
```

Let  $T(N)$  be the running time for the function call  $\text{fib}(n)$

if  $N=0$  or  $N=1$      $T(0) = T(1) = O(1)$

if  $n \geq 2$

$T(n) = \text{cost of constant op at line 1} + \text{cost of line 3 work}$

$T(n) = 1 \text{ op} + (\text{addition} + 2 \text{ function calls})$

# Algorithm Complexity Analysis

46

## Analysis of recursion

$$T(n) = 1 \text{ op} + (\text{addition} + \text{cost of fib}(n-1) + \text{cost fib}(n-2))$$

Thus ,

$$T(n) = T(n-1) + T(n-2) + 2$$

$$\text{Since fib}(n) = \text{fib}(n-1) + \text{fib}(n-2)$$

it is easy to show by induction that :

$$T(n) \geq \text{fib}(n)$$

- we have showed (in chapter 1) that  $\text{fib}(n) < (5/3)^n$
- a similar proof shows for  $n > 4$ ,  $\text{fib}(n) \geq (3/2)^n$

thus  $T(n) \geq (3/2)^n$  and so

the running time of the programme grows **exponentially**.  
This program is slow because there is a huge amount of redundant work being performed.

By using an array and a for loop,  
the programme running time can be reduced substantially.

# Algorithm Complexity Analysis

## Maximum Subsequence Problem

47

Given an array of N elements  $A_1, A_2, A_3, \dots, A_N$ , (possibly negative)

find the maximum value of  $\sum_{k=i}^j A_k$

Need to find  $i, j$  such that the sum of all elements between the  $i^{\text{th}}$  and  $j^{\text{th}}$  positions is maximum for all such sums

(for convenience, the maximum subsequence sums is 0 if all integers are negative)

### Example

for the input -2, 11, -4, 13, -5, -2 the answer is 20

We will discuss four algorithms to solve it, their performance varies :  
 $O(N)$ ,  $O(N \log N)$ ,  $O(N^2)$ ,  $O(N^3)$

# Running time of 4 algorithms for max subsequence sum

48

Input Size	Algorithm Time ( seconds)			
	1	2	3	4
	$O(N^3)$	$O(N^2)$	$O(N \log N)$	$O(N)$
$N = 100$	0.000159	0.000006	0.000005	0.000002
$N = 1,000$	0.095857	0.000371	0.000060	0.000022
$N = 10,000$	86.67	0.033322	0.000619	0.000222
$N = 100,000$	NA	3.33	0.006700	0.002205
$N = 1,000,000$	NA	NA	0.074870	0.022711

Figure [textbook Weiss, Figure 2.2]



# Maximum Subsequence Problem

## Algorithm 1

49

```
/**
 * Cubic maximum contiguous subsequence sum algorithm. */
int maxSubSum1( const vector<int> & a )
{
    int maxSum = 0;

    for( int i = 0; i < a.size( ); ++i )

        for( int j = i; j < a.size( ); ++j )
        {
            int thisSum = 0;
            for( int k = i; k <= j; ++k )
                thisSum += a[ k ];
            if( thisSum > maxSum )
                maxSum = thisSum;
        }
    return maxSum;
}
```

# Complexity of Algorithm 1

Because constants do not matter, the runtime is obtained from the sum :

We have 
$$\sum_{k=0}^{N-1} \sum_{k=j}^{N-1} \sum_{k=i}^j 1$$

inner loop 
$$\sum_{k=i}^j 1 = j - i + 1$$

Outer Loop 
$$\sum_{j=i}^{N-1} (j - i + 1) = \frac{(N-i+1)(N-i)}{2}$$

$$\sum_{i=0}^{N-1} \frac{(N-i+1)(N-i)}{2} = \frac{N^3 + 3N^2 + 2N}{6}$$

# Analysis of Algorithm 1

in Algorithm 1 can be made more efficient leading to  $O(N^2)$ .  
Thus, the cubic running time can be avoided by removing the innermost for loop, because :

$$\sum_{K=i}^j A_k = A_j + \sum_{K=i}^{j-1} A_k$$

# Maximum Subsequence Problem

## Algorithm 2

```
/**
 * Quadratic maximum contiguous subsequence sum algorithm.
 */
int maxSubSum2( const vector<int> & a )
{
    int maxSum = 0;
    for( int i = 0; i < a.size( ); ++i )
    {
        int thisSum = 0;
        for( int j = i; j < a.size( ); ++j )
        {
            thisSum += a[ j ];

            if( thisSum > maxSum )
                maxSum = thisSum;
        }
    }
    return maxSum;}
```

## Complexity of Algorithm 2

the runtime of algorithm2 is obtained from the two for loops :

$$\sum_{i=0}^{N-1} \sum_{j=i}^{N-1} 1 = \sum_{i=0}^{N-1} (N - i)$$

$$\sum_{i=0}^{N-1} (N - i) = N^2 - \frac{(N - 1)N}{2} = \frac{N(N + 1)}{2} = \frac{N^2 + N}{2}$$

$$O(N^2)$$

# Maximum Subsequence Problem

## Algorithm 3

### Divide and Conquer

#### **Divide-and- conquer strategy :**

- ❖ Split the big problem into “two” small sub-problems,
- ❖ Solve each of them efficiently,
- ❖ Combine the “two” solutions.