Algorithms and Data Structures

Lec01

Introduction

Dr. Mohammad Ahmad

Algorithms and Data Structures

In this course, we will look at:

- *Algorithms* for solving problems efficiently
- *Data structures* for efficiently storing, accessing, and modifying data

We will see that all data structures have trade-offs

- There is no *ultimate* data structure...
- The choice depends on our requirements

What is an algorithm?

An ***algorithm*** is a sequence of unambiguous instructions for solving a problem, i.e., for obtaining a required output for any legitimate input in a finite amount of time.

What about data structures

How data is organized

- *A data structure is defined by*
  - *the logical arrangement of data elements, combined with:*
  - *the set of operations we need to access the elements.*

Algorithms + Data Structures = Programs

Algorithms $\leftrightarrow$ Data Structures

What is this Course About?

Clever ways to organize information in order to enable efficient computation

- What do we mean by clever?
- What do we mean by efficient?

Clever? Efficient?

Array, Lists, Stacks, Queues
Heaps
Binary Search Trees
AVL Trees
Hash Tables
Graphs
Disjoint Sets

Data Structures

Insert
Delete
Find
Merge
Shortest Paths
Union

Algorithms

Motivation Example

- Consider searching for an element in an array
  - In an array, we can access it using an index `array[k]`
- Consider searching for an entry in a sorted array
  - In a sorted array, we use a fast binary search
    - Very fast

Mathematical Background

- The ceiling and floor functions
- L'Hôpital's rule
- Logarithms
- Arithmetic and other polynomial series
  - Mathematical induction
- Geometric series
- Recurrence relations
- Combinations

Floor and ceiling functions

The *floor* function maps any real number x onto the greatest integer less than or equal to x:

$$\lfloor 3.2 \rfloor = \lfloor 3 \rfloor = 3$$

$$\lfloor -5.2 \rfloor = \lfloor -6 \rfloor = -6$$

The *ceiling* function maps x onto the least integer greater than or equal to x:

$$\lceil 3.2 \rceil = \lceil 4 \rceil = 4$$

$$\lceil -5.2 \rceil = \lceil -5 \rceil = -5$$

L'Hôpital's rule

If you are attempting to determine

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

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

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

Repeat as necessary…

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

Logarithms

We will begin with a review of logarithms:

If $n = e^m$, we define

$$m = \ln(n)$$

It is always true that $e^{\ln(n)} = n$; however, $\ln(e^n) = n$ requires that n is real

Logarithms

Exponentials grow faster than any non-constant polynomial

$$\lim_{n\to\infty}\frac{e^n}{n^d}=\infty$$

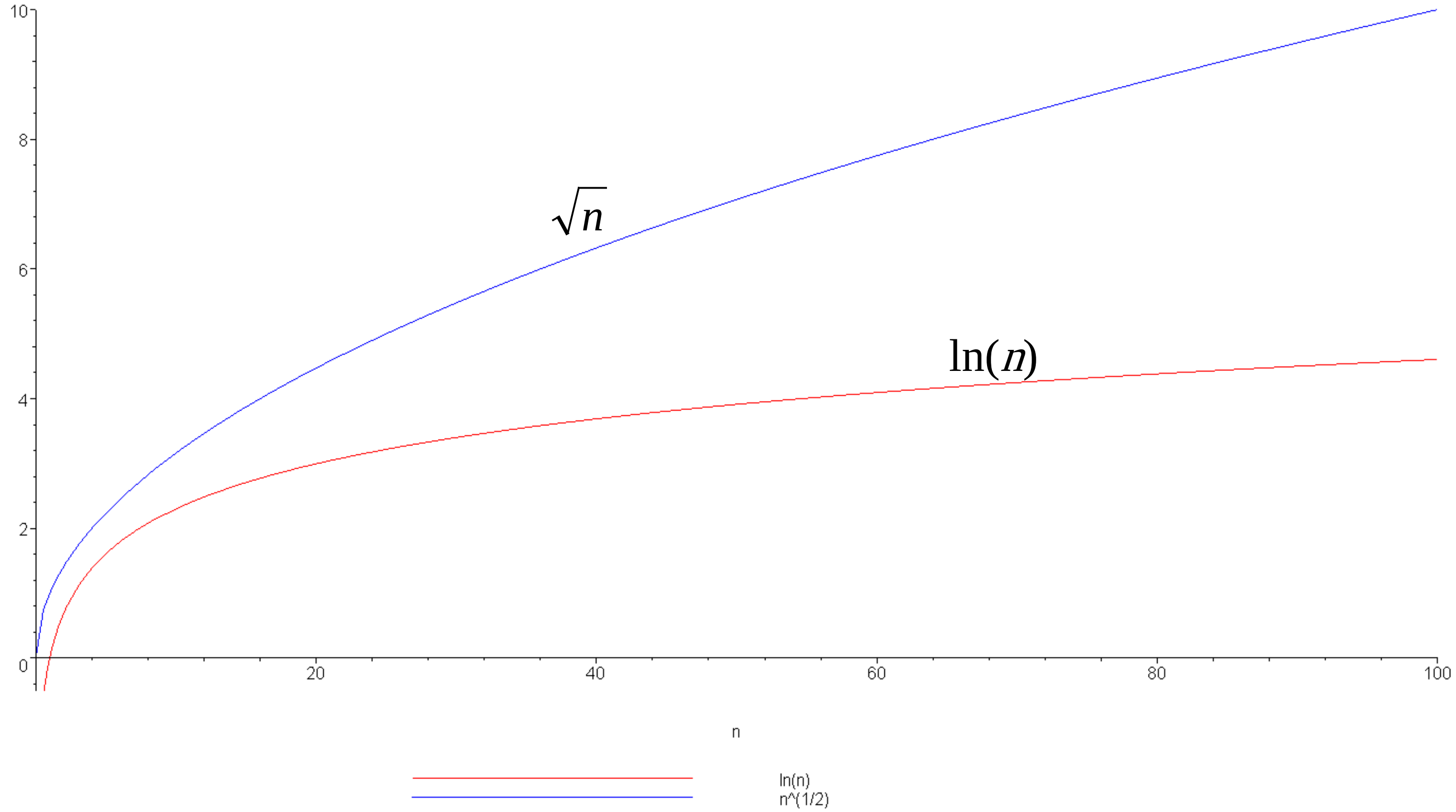
for any $d > 0$

$$\lim_{n\to\infty}\frac{\ln(n)}{n^d}=0$$

Thus, their inverses—logarithms—grow slower than any polynomial

Logarithms

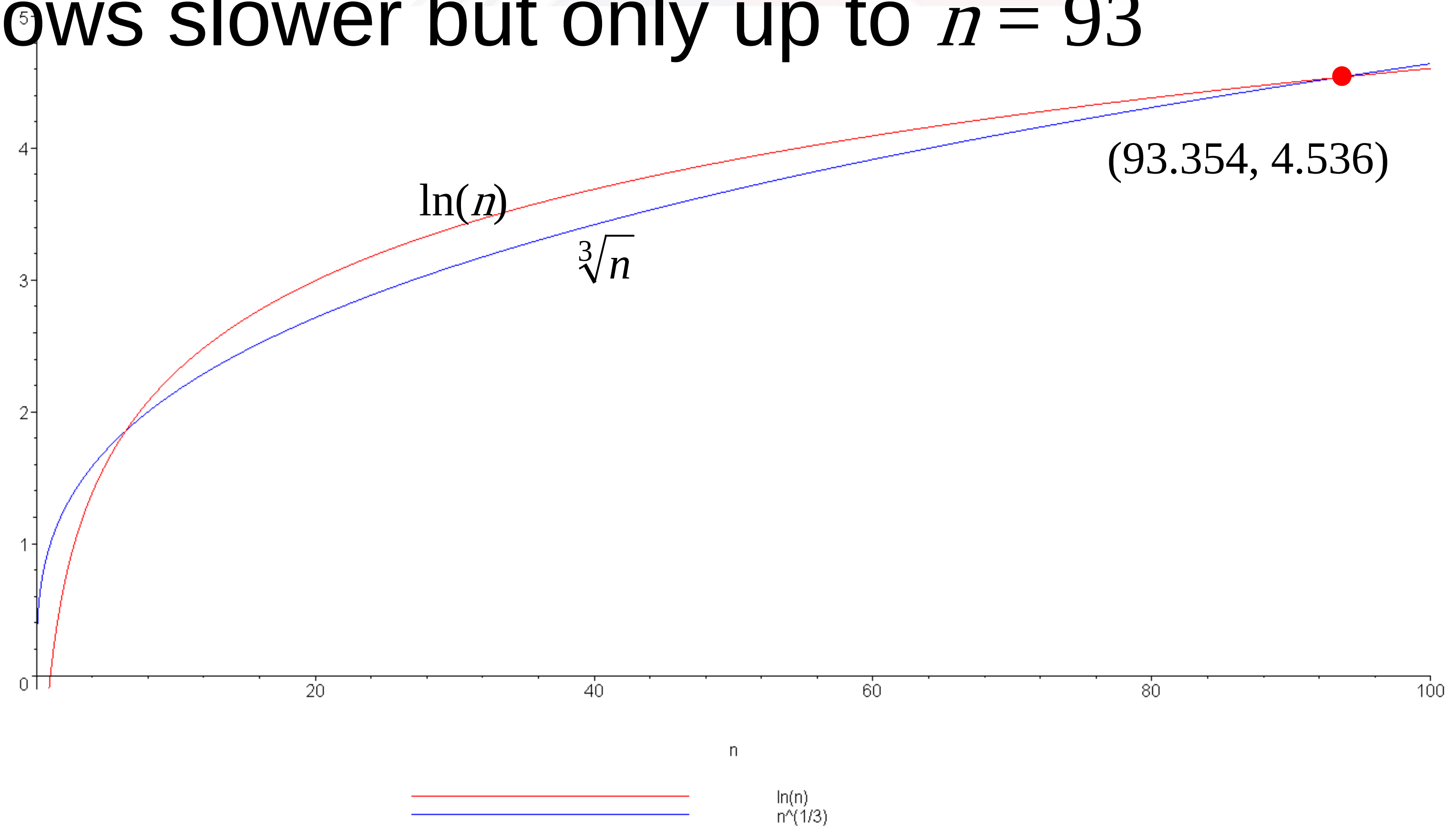
Example: $f(n) = n^{1/2} = \sqrt{n}$ is strictly greater than $\ln(n)$



Logarithms

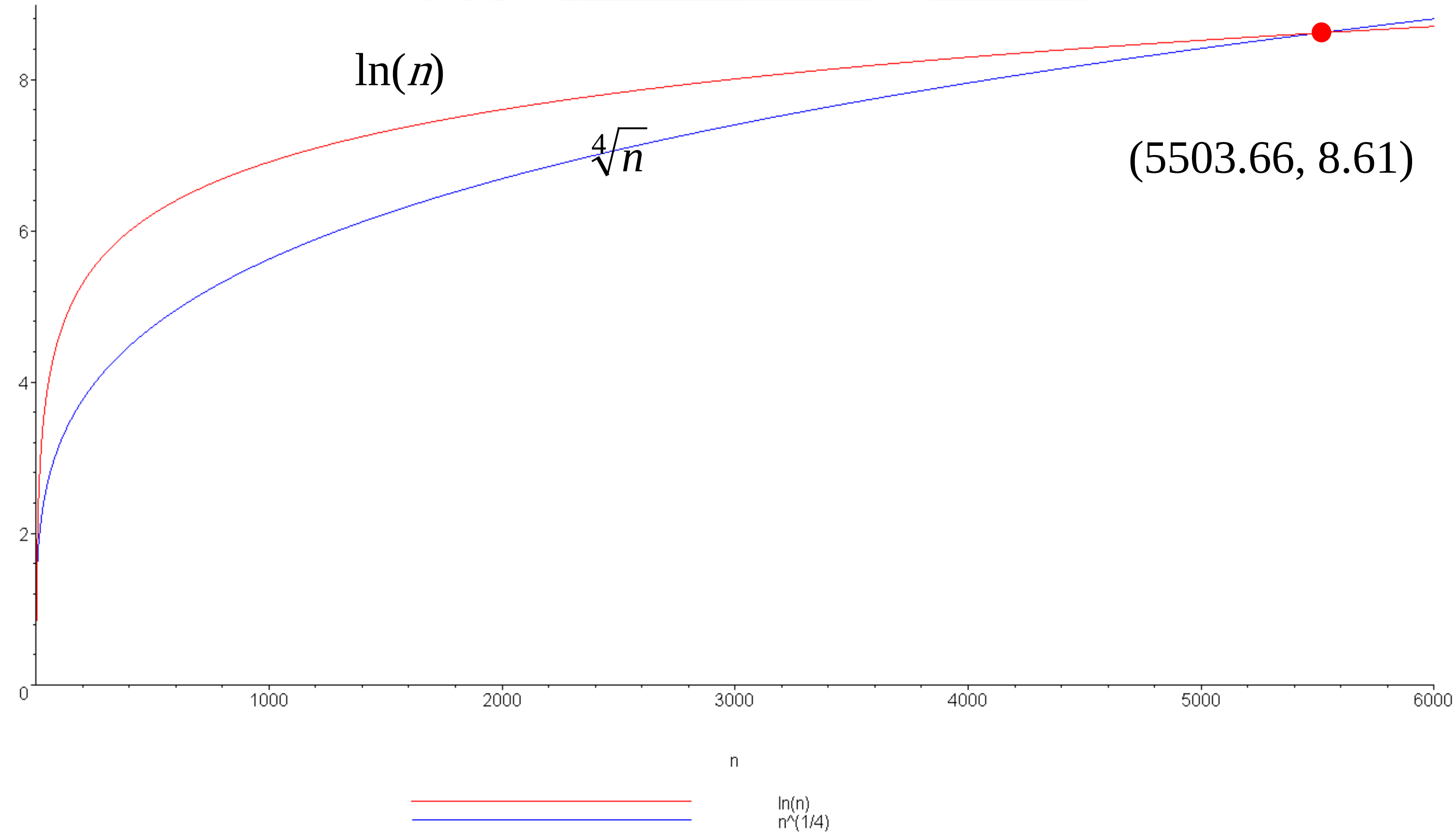
$$f(n) = n^{1/3} = \sqrt[3]{n}$$

grows slower but only up to $n = 93$



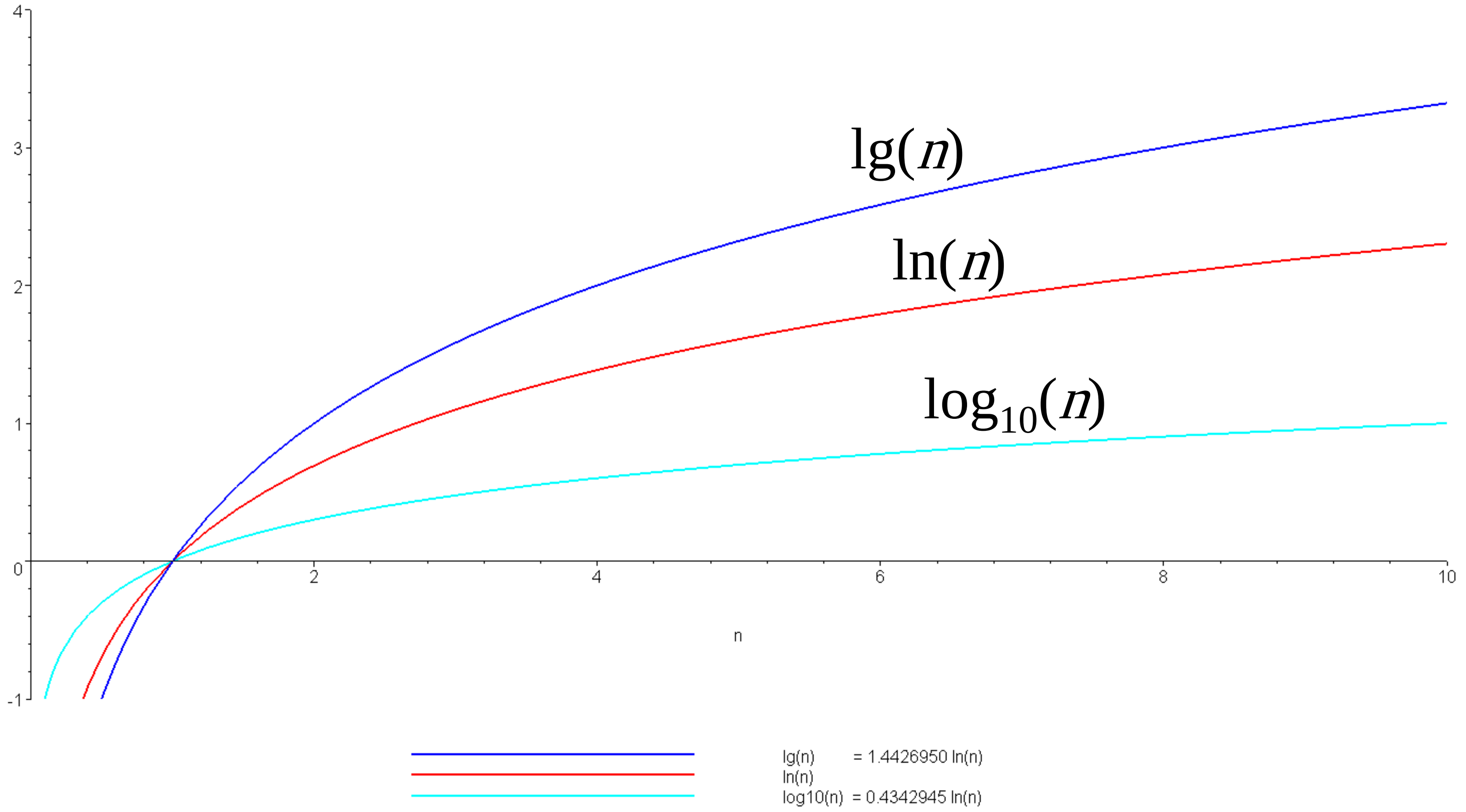
Logarithms

You can view this with any polynomial

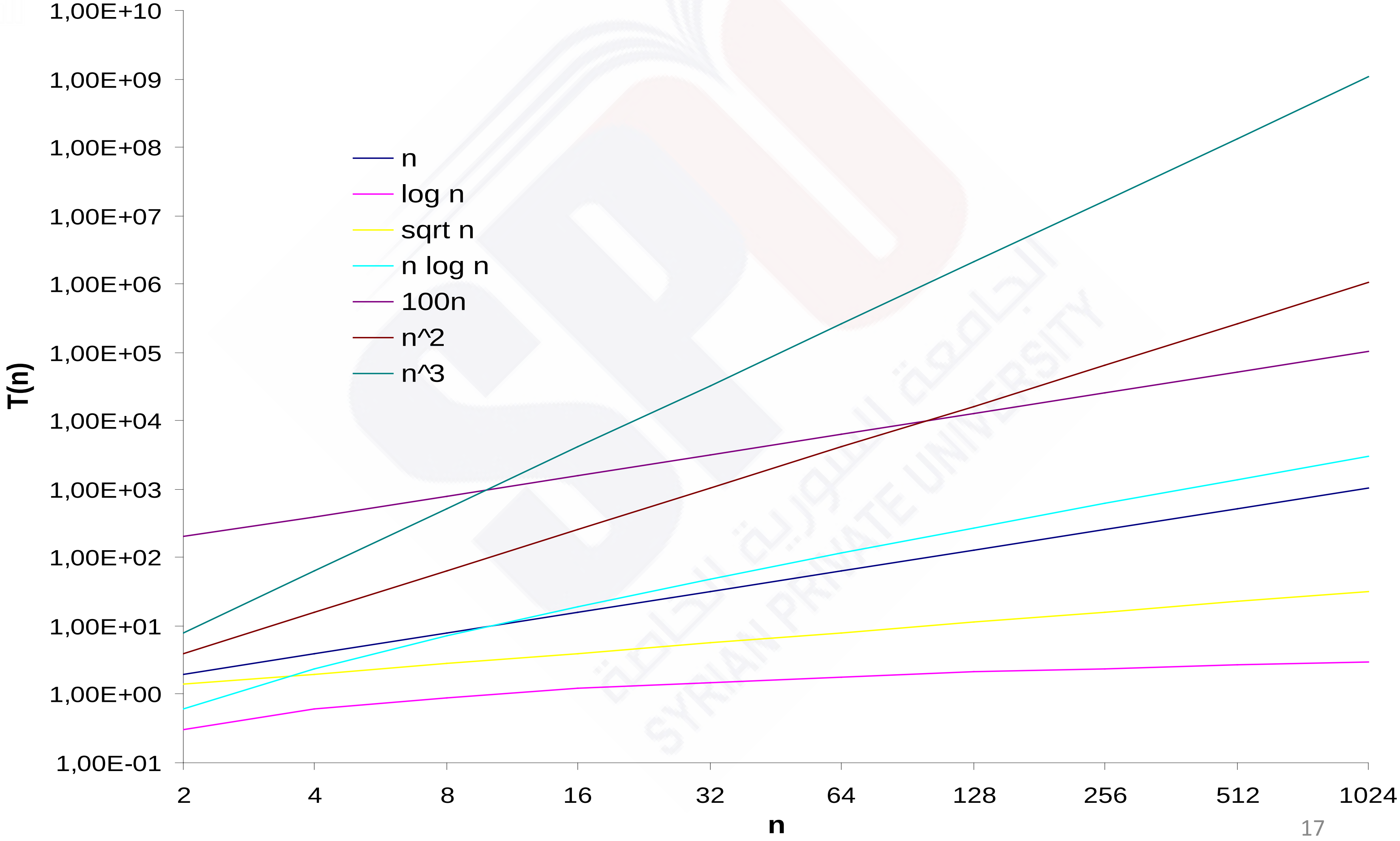


Logarithms

A plot of $\log_2(n) = \lg(n)$, $\ln(n)$, and $\log_{10}(n)$



Growth Functions



Arithmetic series

Next we will look various series

Each term in an arithmetic series is increased by a constant value (usually 1) :

$$0+1+2+3+\cdots+n=\sum_{k=0}^{n} k=\frac{n(n+1)}{2}$$

Other polynomial series

We could repeat this process, after all:

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \qquad \sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4}$$

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} \approx \frac{n^2}{2} \qquad \sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \approx \frac{n^3}{3}$$

however, it is easier to see the pattern:

$$\sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4} \approx \frac{n^4}{4}$$

Geometric series

The next series we will look at is the geometric series with common ratio r:

$$\sum_{k=0}^{n} r^k = \frac{1-r^{n+1}}{1-r}$$

and if $|r| < 1$ then it is also true that

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$$

Geometric series

A common geometric series will involve the ratios r = ½ and r = 2:

$$\sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^i = 2$$

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

$$\sum_{k=0}^{n}2^k = \frac{1-2^{n+1}}{1-2} = 2^{n+1}-1$$

Recurrence relations

Finally, we will review recurrence relations:

– Sequences may be defined explicitly: $x_n = 1/n$

$$1, {}^{1}/_{2}, {}^{1}/_{3}, {}^{1}/_{4}, \ldots$$

– A recurrence relationship is a means of defining a sequence based on previous values in the sequence

– Such definitions of sequences are said to be *recursive*

Recurrence relations

Define an initial value: *e.g.*, $x_1 = 1$

Defining x_n in terms of previous values:

– For example,

$$x_n = x_{n-1} + 2$$

$$x_n = 2x_{n-1} + n$$

$$x_n = x_{n-1} + x_{n-2}$$

Recurrence relations

Given the two recurrence relations

$$x_n = x_{n-1} + 2 \qquad\qquad x_n = 2x_{n-1} + n$$

and the initial condition $x_1 = 1$ we would like to find explicit formulae for the sequences

In this case, we have

$$x_n = 2n - 1 \qquad\qquad x_n = 2^{n+1} - n - 2$$

respectively

Recurrence relations

The previous examples using the functional notation are:

$$f(n) = f(n-1) + 2 \qquad g(n) = 2\,g(n-1) + n$$

With the initial conditions $f(1) = g(1) = 1$, the solutions are:

$$f(n) = 2n - 1 \qquad g(n) = 2^{n+1} - n - 2$$

Recurrence relations

In some cases, given the recurrence relation, we can find the explicit formula:

– Consider the Fibonacci sequence:

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

$$f(0) = f(1) = 1$$

Combinations

Given n distinct items, in how many ways can you choose k of these?

- I.e., "In how many ways can you combine k items from n?"
- For example, given the set $\{1, 2, 3, 4, 5\}$, I can choose three of these in any of the following ways:

$$\{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 3, 4\}, \{1, 3, 5\},$$
$$\{1, 4, 5\}, \{2, 3, 4\}, \{2, 3, 5\}, \{2, 4, 5\}, \{3, 4, 5\}$$

The number of ways such items can be chosen is written

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

where $\binom{n}{k}$ is read as "n choose k"s

There is a recursive definition: $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

Combinations

The most common question we will ask in this vein:

- Given n items, in how many ways can we choose two of them?
- In this case, the formula simplifies to: $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}$

For example, given $\{0, 1, 2, 3, 4, 5, 6\}$, we have the following 21 pairs:

$$\begin{aligned}
&\{0, 1\}, \{0, 2\}, \{0, 3\}, \{0, 4\}, \{0, 5\}, \{0, 6\},\\
&\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{1, 6\},\\
&\{2, 3\}, \{2, 4\}, \{2, 5\}, \{2, 6\},\\
&\{3, 4\}, \{3, 5\}, \{3, 6\},\\
&\{4, 5\}, \{4, 6\},\\
&\{5, 6\}
\end{aligned}$$

Formulation

Often $F(n)$ is an equation:

– For example, $F(n)$ may be an equation such as:

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} \quad \text{for } n \geq 0$$

$$\sum_{k=1}^{n} 2k - 1 = n^2 \quad \text{for } n \geq 1$$

$$\sum_{k=0}^{n} 2^k = 2^{n+1} - 1 \quad \text{for } n \geq 0$$