### CSE 548: Analysis of Algorithms

Rezaul A. Chowdhury
Department of Computer Science
SUNY Stony Brook
Fall 2019

"Theory is when you know everything but nothing works. Practice is when everything works but no one knows why.

In our lab, theory and practice are combined:

nothing works and no one knows why."

— A practical theoretician

### Basic Logistics: Who/Where/When

Lecture Time: MW 7:00 pm - 8:20 pm

Location: Harriman Hall 137, West Campus

Instructor: Rezaul A. Chowdhury

- Office Hours: M 4:00 pm - 5:00 pm

F 6:00 pm - 8:00 pm

239 Computer Science

Email: rezaul@cs.stonybrook.edu

— **TA**: TBA

Class Webpage:

http://www3.cs.stonybrook.edu/~rezaul/CSE548-F19.html

### 2

### **Prerequisites**

- Required: Some background (undergrad level) in the design and analysis of algorithms and data structures
  - fundamental data structures (e.g., lists, stacks, queues and arrays)
  - discrete mathematical structures (e.g., graphs, trees, and their adjacency lists & adjacency matrix representations)
  - fundamental programming techniques (e.g., recursion, divide-and-conquer, and dynamic programming)
  - basic sorting and searching algorithms
  - fundamentals of asymptotic analysis (e.g., O(  $\cdot$  ),  $\Omega(\,\cdot\,$  ) and  $\Theta(\,\cdot\,)$  notations)
- Required: Some background in programming languages ( C / C++)



### Topics to be Covered

The following topics will be covered (hopefully)

- recurrence relations and divide-and-conquer algorithms
- dynamic programming
- graph algorithms (e.g., network flow)
- amortized analysis
- advanced data structures (e.g., Fibonacci heaps)
- cache-efficient and external-memory algorithms
- high probability bounds and randomized algorithms
- parallel algorithms and multithreaded computations
- NP-completeness and approximation algorithms
- the alpha technique (e.g., disjoint sets, partial sums)
- FFT ( Fast Fourier Transforms )



### **Grading Policy**

- Four Homework Problem Sets
   ( highest score 15%, lowest score 5%, and others 10% each ): 40%
- Two Exams ( higher one 30%, lower one 15% ): 45%
  - Midterm (in-class): Oct 16Final (in-class): Dec 2
- Class participation & attendance: 5%

Scribe note (one lecture): 10%

5

### <u>Textbooks</u>

### Required

Thomas Cormen, Charles Leiserson, Ronald Rivest, and Clifford Stein.
 Introduction to Algorithms (3rd Edition), MIT Press, 2009.

### Recommended

- Sanjoy Dasgupta, Christos Papadimitriou, and Umesh Vazirani.
   Algorithms (1st Edition), McGraw-Hill, 2006.
- Jon Kleinberg and Éva Tardos.
   Algorithm Design (1st Edition), Addison Wesley, 2005.
- Rajeev Motwani and Prabhakar Raghavan.
   Randomized Algorithms (1st Edition), Cambridge University Press, 1995.
- Vijay Vazirani.
   Approximation Algorithms, Springer, 2010.
- Joseph JáJá.
   An Introduction to Parallel Algorithms (1st Edition), Addison Wesley, 1992.

# What is an Algorithm?

An algorithm is a *well-defined computational procedure* that solves a well-specified computational problem.

It accepts a value or set of values as *input*, and produces a value or set of values as *output* 

**Example**: *mergesort* solves the *sorting problem* specified as a relationship between the input and the output as follows.

**Input:** A sequence of n numbers  $\langle a_1, a_2, ..., a_n \rangle$ .

**Output:** A permutation  $\langle a'_1, a'_2, ..., a'_n \rangle$  of the input sequence such that  $a'_1 \leq a'_2 \leq \cdots \leq a'_n$ .

 $\rightarrow$ 

# **Desirable Properties of an Algorithm**

- √ Correctness
  - Designing an incorrect algorithm is straightforward
- √ Efficiency
  - Efficiency is easily achievable if we give up on correctness

Surprisingly, sometimes incorrect algorithms can also be useful!

- If you can control the error rate
- Tradeoff between correctness and efficiency:
  - Randomized algorithms

( Monte Carlo: always efficient but sometimes incorrect,

Las Vegas: always correct but sometimes inefficient )

Approximation algorithms

( always incorrect! )

-

### **How Do You Measure Efficiency?**

We often want algorithms that can use the available resources efficiently.

Some measures of efficiency

- time complexity
- space complexity
- cache complexity
- I/O complexity
- energy usage
- number of processors/cores used
- network bandwidth

### **Goal of Algorithm Analysis**

Goal is to predict the behavior of an algorithm without implementing it on a real machine.

But predicting the exact behavior is not always possible as there are too many influencing factors.

Runtime on a serial machine is the most commonly used measure.

We need to model the machine first in order to analyze runtimes.

But an exact model will make the analysis too complicated!

So we use an approximate model (e.g., assume unit-cost Random Access Machine model or RAM model).

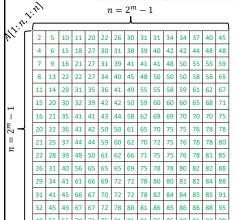
We may need to approximate even further: e.g., for a sorting algorithm we may count the comparison operations only.

So the predicted running time will only be an approximation!

### **Performance Bounds**

- worst-case complexity: maximum complexity over all inputs of a given size
- average complexity: average complexity over all inputs of a given size
- amortized complexity: worst-case bound on a sequence of operations
- expected complexity: for algorithms that make random choices during execution ( randomized algorithms )
- **high-probability bound:** when the probability that the complexity holds is  $\geq 1-\frac{c}{n^\alpha}$  for input size n, positive constant c and some constant  $\alpha\geq 1$

# Searching in a Sorted Grid



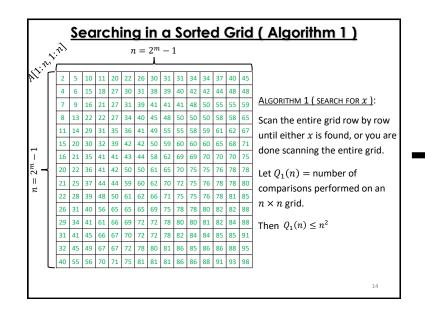
You are given an  $n \times n$  grid A[1:n,1:n], where  $n=2^m-1$  for some integer m>0.

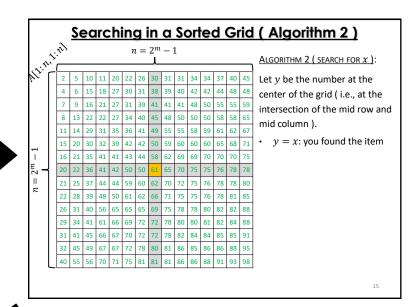
Each grid cell contains a number.

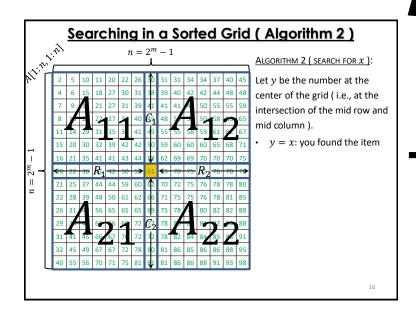
The numbers in each row are sorted in non-decreasing order from left to right.

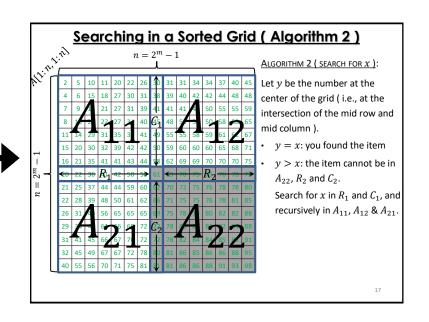
The numbers in each column are sorted in non-decreasing order from top to bottom.

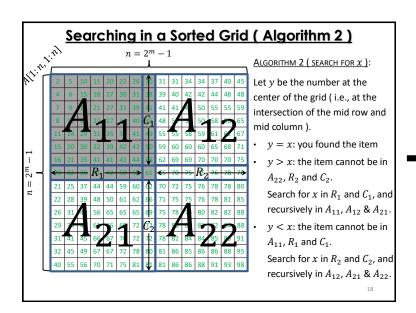
13

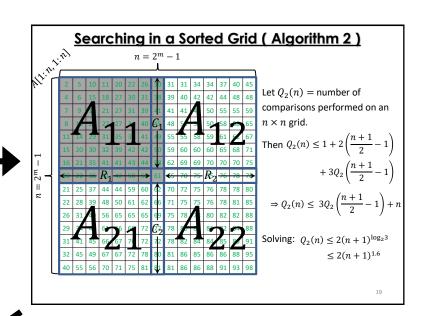


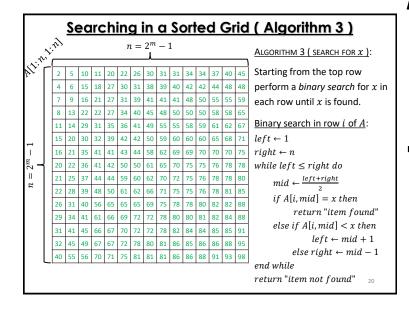


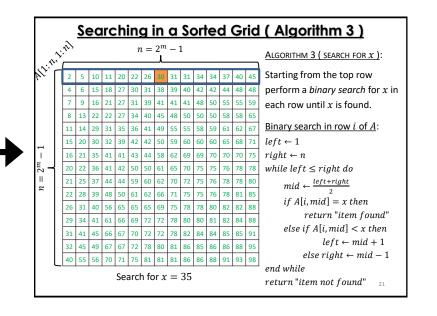


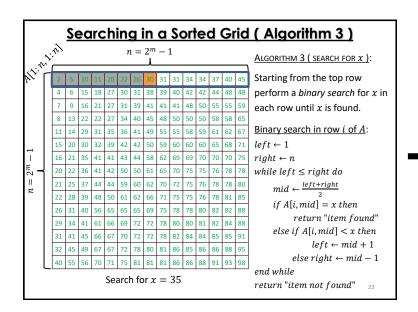


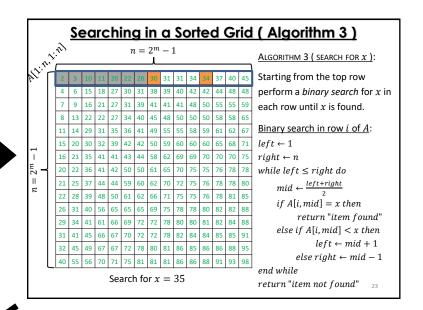


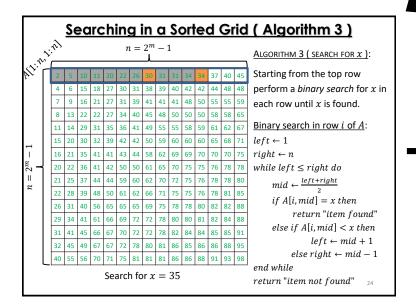


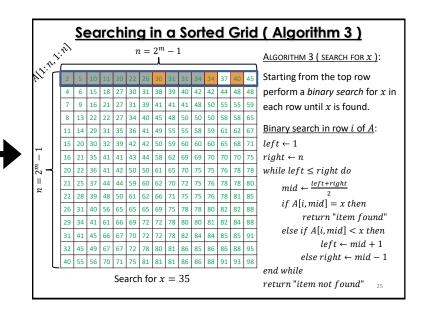


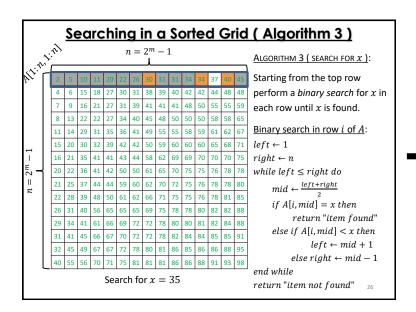


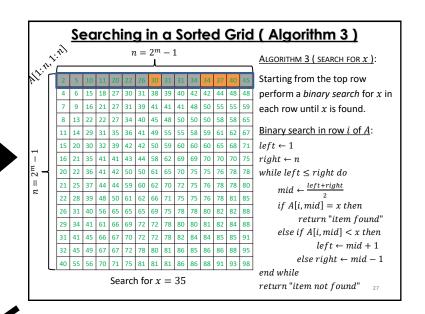


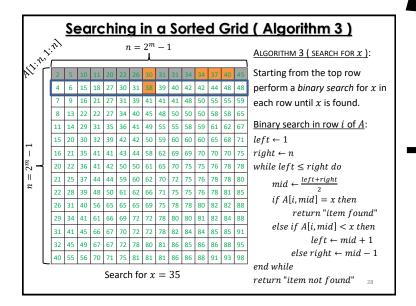


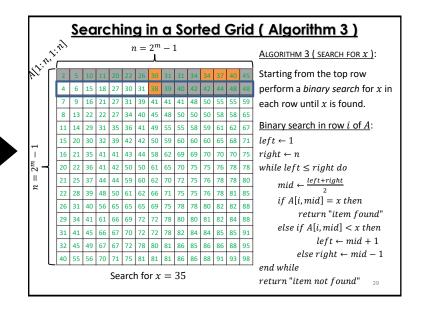


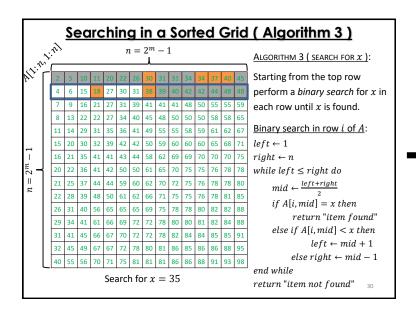


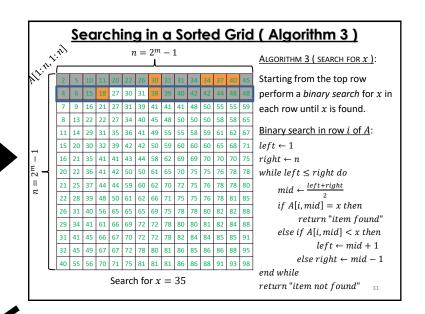


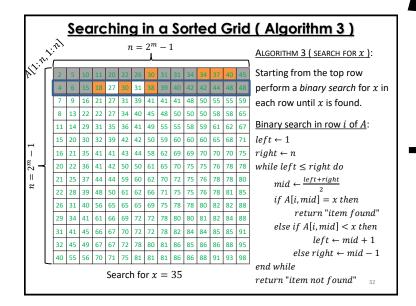


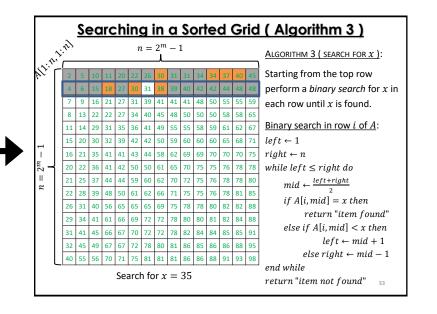


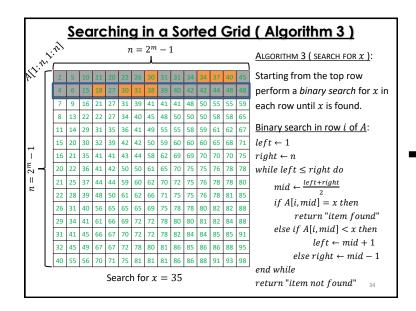


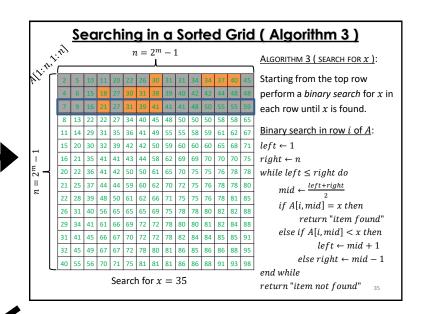


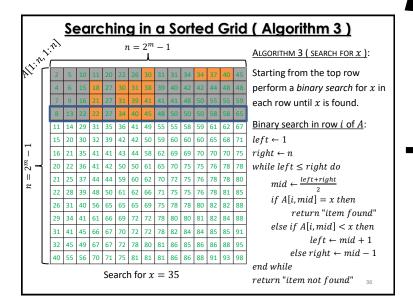


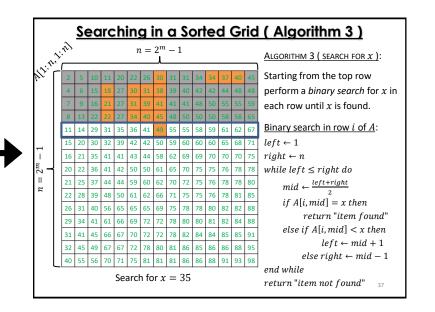


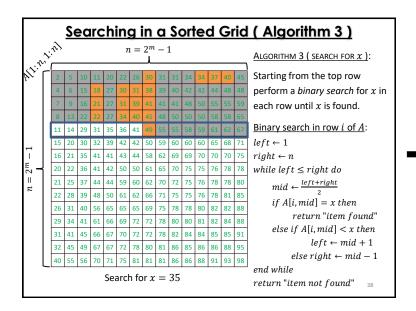


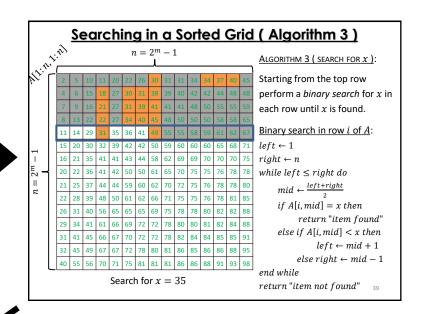


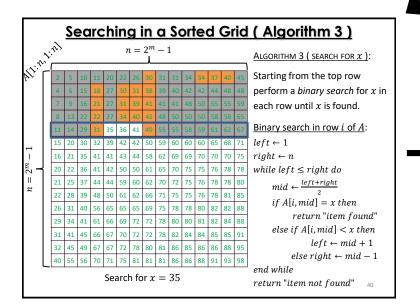


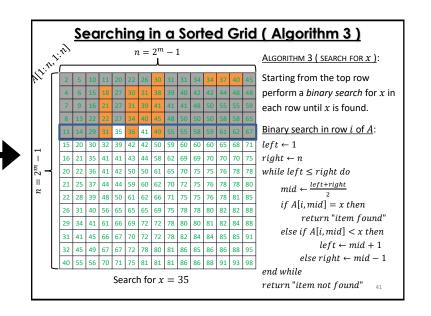


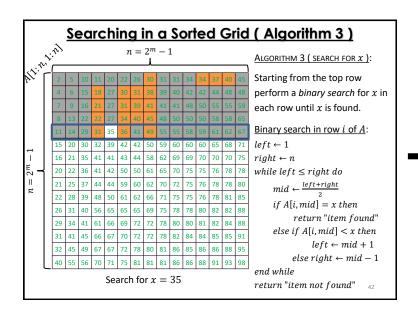


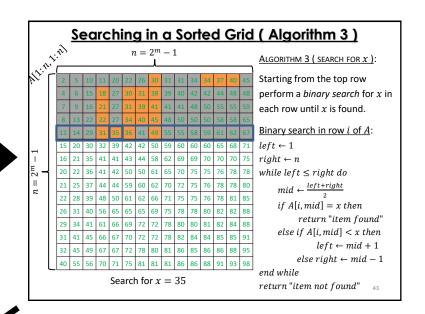


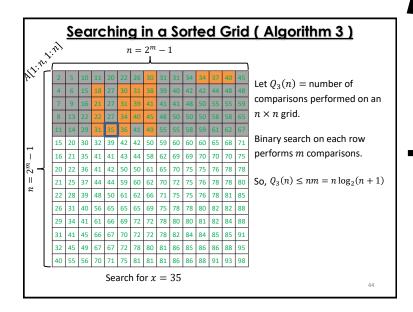


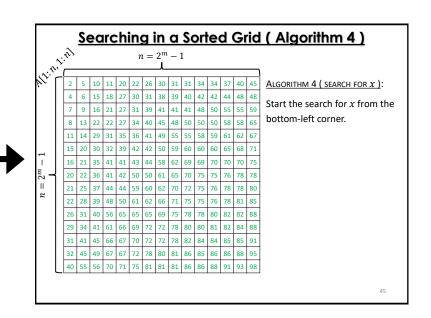


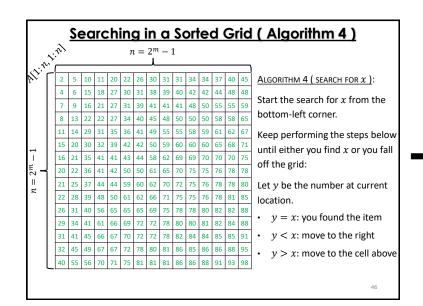


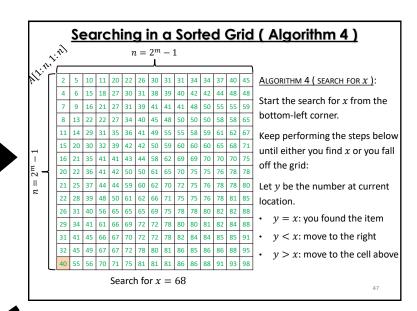


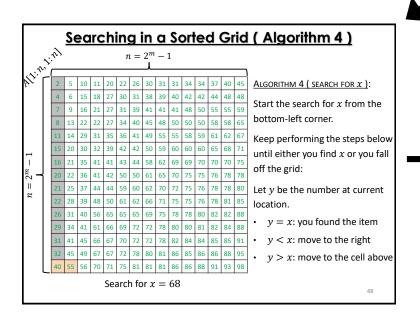


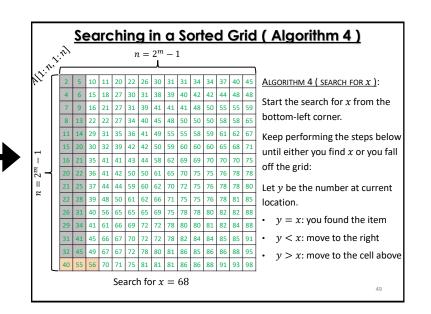


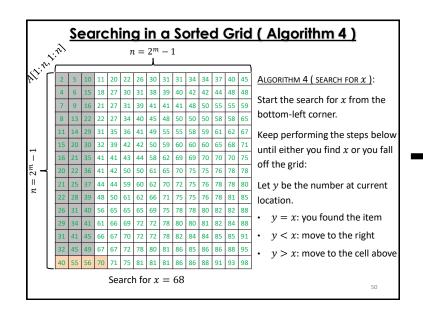


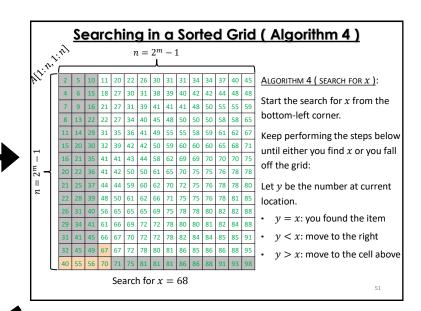


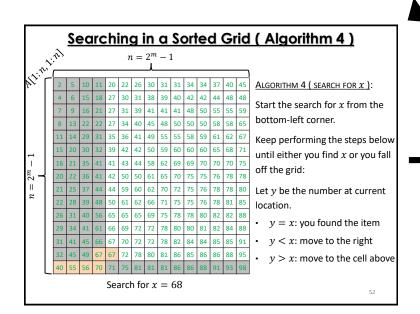


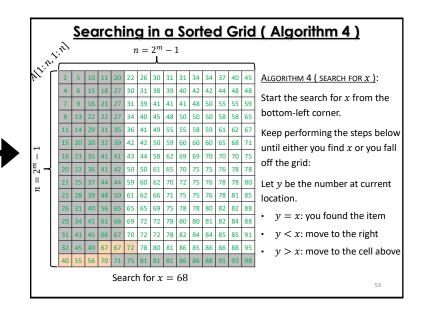


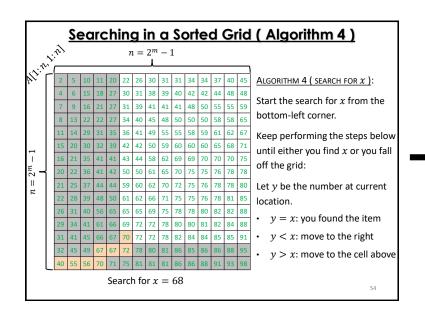


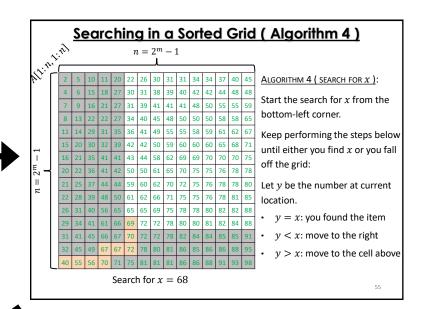


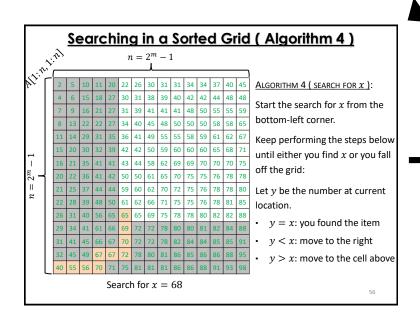


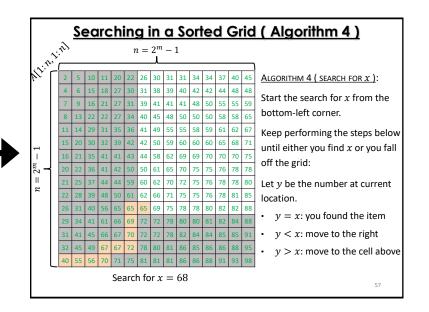


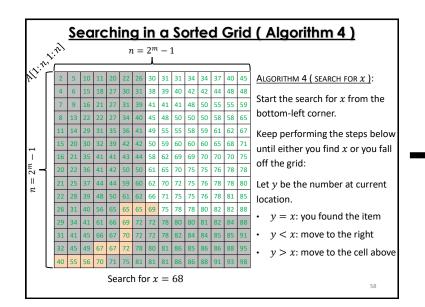


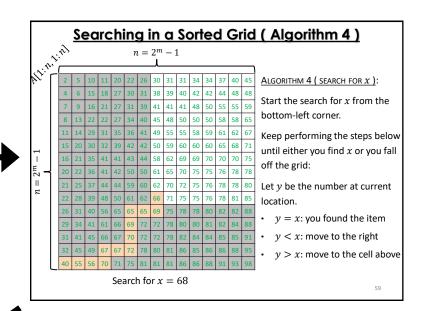


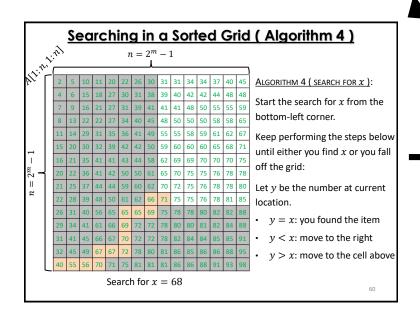


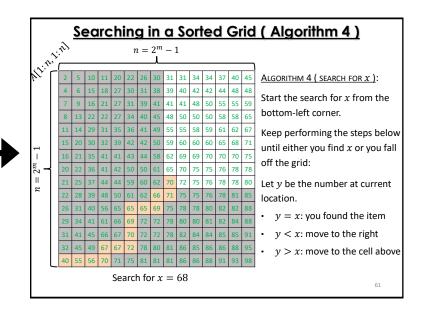


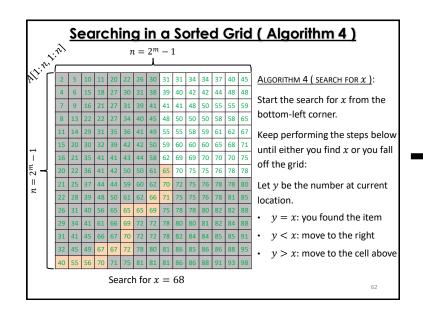


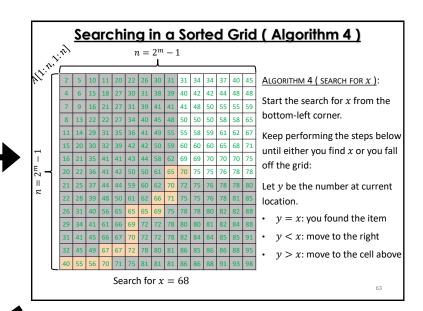


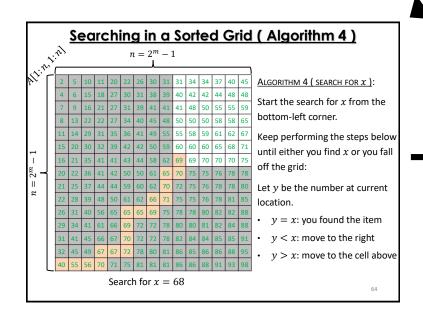


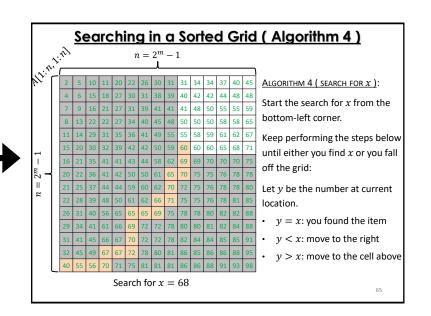


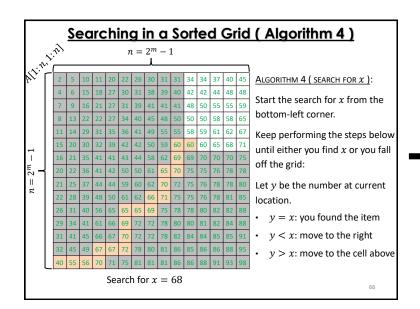


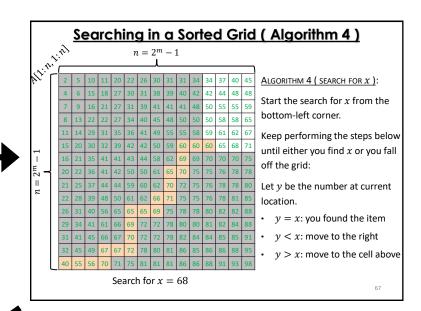


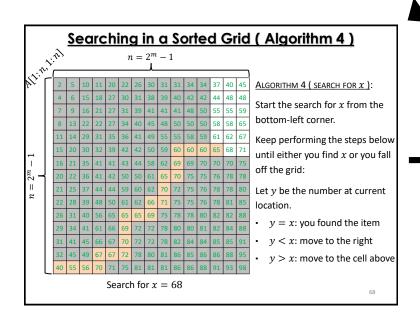


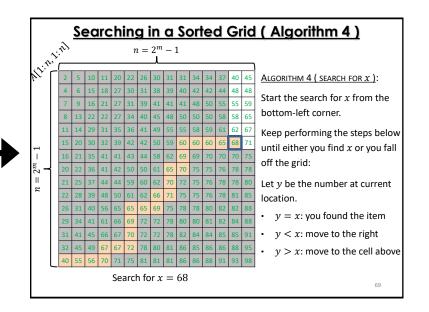


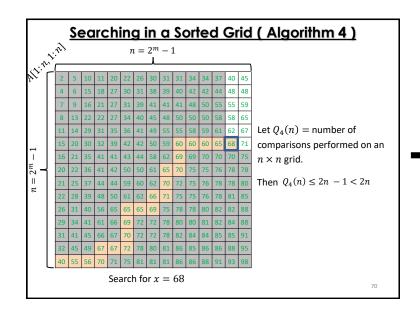


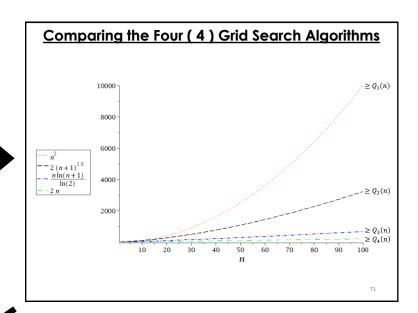




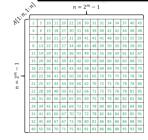








# Searching in a Sorted Grid ( Algorithm 1 )



### ALGORITHM 1 ( SEARCH FOR x ):

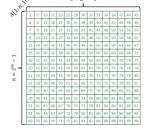
- 1. for i = 1 to n do
- 2. for j = 1 to n do
- 3. if A[i,j] = x then return "item found"
- 4. end for
- 5. end for
- 6. return "item not found"

Let  $T_1(n) = \text{running time of algorithm 1 on an } n \times n \text{ grid.}$ 

Though we were able to compute an exact worst-case bound for  $Q_1(n)$ , the same cannot be done for  $T_1(n)$  because it depends on many other external factors such as CPU speed, programming style, compiler and optimization level used, etc.

But for large values of n,  $T_1(n)$ 's worst-case value will be within a constant factor of that of  $Q_1(n)$ . That constant is generally unknown, and depends on the specific hardware and compiler used, expertise of the programmer, etc.

# Searching in a Sorted Grid ( Algorithm 1 )



### ALGORITHM 1 ( SEARCH FOR x ):

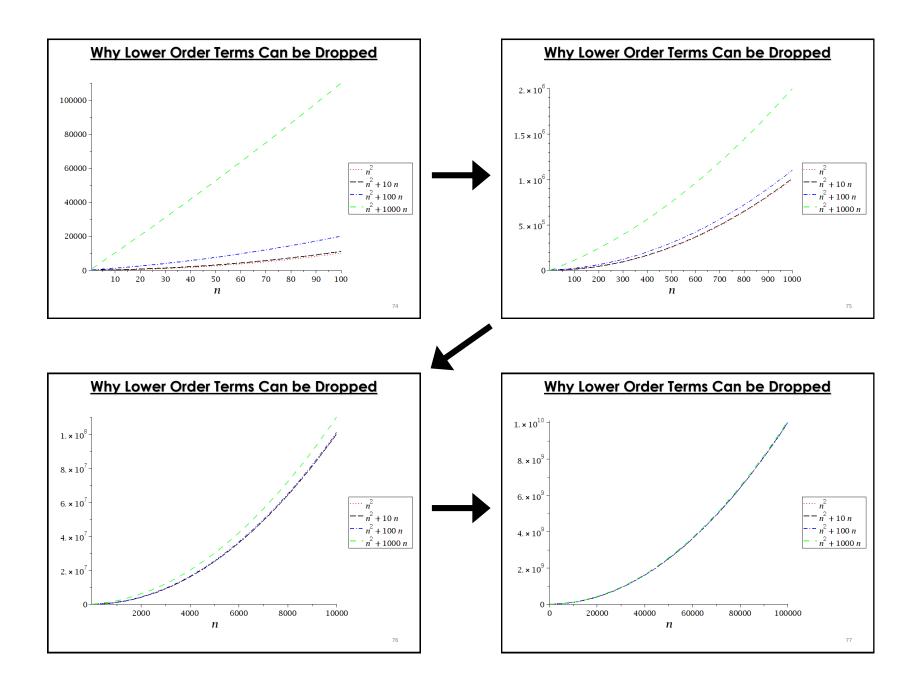
- 1. for i = 1 to n do
- $2. \quad for j = 1 to n do$
- 3. if A[i,j] = x then return "item found"
- i. end for
- 5. end for
- 6. return "item not found"

### In the worst case,

- line 3 will be executed  $n^2$  times,
- variable j in line 2 will be updated  $n^2$  times,
- variable i in line 1 will be updated n times, and
- line 6 will be executed will be executed 1 time.

Hence,  $T_1(n) \le a_1 n^2 + a_2 n + a_3$ , where  $a_1$ ,  $a_2$  and  $a_3$  are constants.

Clearly,  $T_1(n) \le (a_1 + 1)n^2 = (a_1 + 1)Q_1(n)$ , when  $n \ge a_2 + a_3$ .



### Running Times of the Four (4) Algorithms for Large n

GRID SEARCHING ALGORITHM	WORST-CASE BOUND ON #COMPARISONS	WORST-CASE BOUND ON RUNNING TIMES
ALGORITHM 1	$Q_1(n) \le n^2$	$T_1(n) \le c_1 n^2$
ALGORITHM 2	$Q_2(n) \le 2(n+1)^{1.6}$	$T_2(n) \le c_2(n+1)^{1.6}$
ALGORITHM 3	$Q_3(n) \le n \log_2(n+1)$	$T_3(n) \le c_3 n \log_2(n+1)$
ALGORITHM 4	$Q_4(n) \le 2n$	$T_4(n) \le c_4 n$

 $c_1, c_2, c_3$  and  $c_4$  are constants

# Why Faster Algorithms?

As the input gets large a faster algorithm run on a slow computer will eventually beat a slower algorithm run on a fast computer!

Suppose we run Algorithm 4 on computer A that can execute only 1 million instructions per second. The algorithm was implemented by an inexperienced programmer, and so  $c_4=10$ .

Suppose we run Algorithm 1 on computer B that is 1000 times faster than A, and the algorithm was implemented by an expert programmer, and so  $c_1=1$ .

Let's run both algorithm on a large grid with n = 100,000.

Then Algorithm 1 will require up to  $\frac{1\times(100000)^2}{1000000000}=10$  seconds, while Algorithm 4 will terminate in only  $\frac{10\times100000}{1000000}=1$  second!

78

### **Asymptotic Bounds**

We compute performance bounds as functions of input size n.

Asymptotic bounds are obtained when  $n \to \infty$ .

Several types of asymptotic bounds

- upper bound ( O-notation )
- strict upper bound ( o-notation )
- lower bound (  $\Omega$ -notation )
- strict lower bound (  $\omega$ -notation )

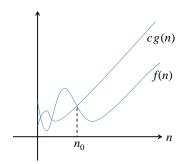
- tight bound (  $\Theta$ -notation )



Asymptotic Stickman ( by Aleksandra Patrzalek )

80

### Asymptotic Upper Bound (O-notation)



 $O\big(g(n)\big) = \left\{ \begin{array}{l} f(n): \text{there exist positive constants c and } n_0 \text{ such that} \\ 0 \leq f(n) \leq cg(n) \text{ for all } n \geq n_0 \end{array} \right\}$ 

$$O(g(n)) = \begin{cases} f(n): \text{there exists a positive constant } c \text{ such that} \\ \lim_{n \to \infty} \left( \frac{f(n)}{g(n)} \right) \le c \end{cases}$$

### Asymptotic Upper Bound (O-notation)

Recall that for Algorithm 1 we had,  $T_1(n) \le f(n)$ , where,

$$f(n) = a_1 n^2 + a_2 n + a_3$$
, for constants  $a_1$ ,  $a_2$  and  $a_3$ .

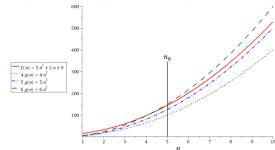
Suppose,  $a_1 = 5$ ,  $a_2 = 2$  and  $a_3 = 9$ .

Then  $f(n) = 5n^2 + 2n + 9$ .

We will now derive asymptotic bounds for f(n).







Let  $g(n) = n^2$ .

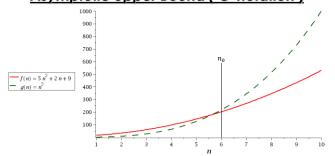
Then  $f(n) = O(n^2)$  because:

$$0 \le f(n) \le cg(n)$$
 for  $c = 6$  and  $n \ge 5$ .

83

82

# Asymptotic Upper Bound ( O-notation )



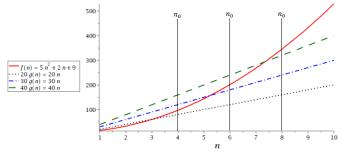
Let  $g(n) = n^3$ .

Then  $f(n) = O(n^3)$  because:

 $0 \le f(n) \le cg(n)$  for c = 1 and  $n \ge 6$ .

 $\rightarrow$ 

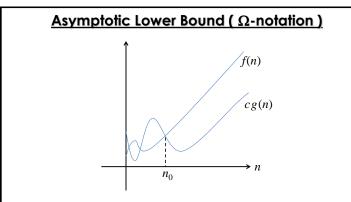
# <u>Asymptotic Upper Bound (O-notation)</u>

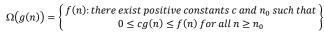


Let g(n) = n.

Then  $f(n) \neq O(n)$  because:

f(n) > cg(n) for any c and  $n \ge \frac{c}{5}$ .





$$\Omega \big( g(n) \big) = \left\{ \begin{array}{l} f(n) \text{: there exists a positive constant } c \text{ such that} \\ \lim\limits_{n \to \infty} \left( \frac{f(n)}{g(n)} \right) \geq c \end{array} \right.$$





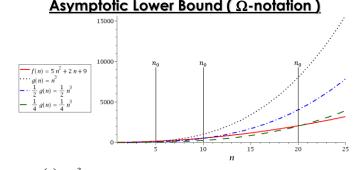
Asymptotic Lower Bound ( $\Omega$ -notation)

Let  $g(n) = n^2$ .

Then  $f(n) = \Omega(n^2)$  because:

 $0 \le cg(n) \le f(n)$  for c = 5 and  $n \ge 1$ .



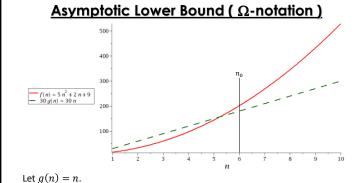


Let  $g(n) = n^3$ .

Then  $f(n) \neq \Omega(n^3)$  because:

cg(n) > f(n) for any c and  $n \ge \frac{5}{c}$ .

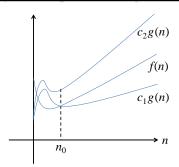
88



Then  $f(n) = \Omega(n)$  because:

 $0 \le cg(n) \le f(n)$  for c = 30 and  $n \ge 6$ .

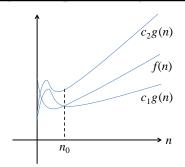
### Asymptotic Tight Bound ( Θ-notation )



$$\Theta(g(n)) = \begin{cases} f(n): \text{there exist positive constants } c_1, c_2 \text{ and } n_0 \text{ such that } \\ 0 \le c_1 g(n) \le f(n) \le c_2 g(n) \text{ for all } n \ge n_0 \end{cases}$$

$$\Theta(g(n)) = \begin{cases} f(n): \text{ there exist positive constants } c_1 \text{ and } c_2 \text{ such that } \\ c_1 \le \lim_{n \to \infty} \left( \frac{f(n)}{g(n)} \right) \le c_2 \end{cases}$$

### Asymptotic Tight Bound ( ⊗-notation )



$$(f(n) = O(g(n))) \land (f(n) = \Omega(g(n))) \Leftrightarrow (f(n) = \Theta(g(n)))$$

91

# Asymptotic Tight Bound ( Θ-notation )

92

$$f(n) = 5n^2 + 2n + 9$$

 $f(n) = \Theta(n^2)$  because both  $f(n) = O(n^2)$  and  $f(n) = \Omega(n^2)$  hold.

 $f(n) \neq \Theta \! \left( n^3 \right) \text{ because though } f(n) = O \! \left( n^3 \right) \text{ holds, } f(n) \neq \Omega \! \left( n^3 \right) .$ 

 $f(n) \neq \Theta(n)$  because though  $f(n) = \Omega(n)$  holds,  $f(n) \neq O(n)$ .

# <u>Asymptotic Strict Upper Bound (o-notation)</u>

$$O\big(g(n)\big) = \left\{ \begin{array}{c} f(n) \text{: there exist positive constants $c$ and $n_0$ such that} \\ 0 \leq f(n) \leq cg(n) \text{ for all } n \geq n_0 \end{array} \right\}$$

$$O\big(g(n)\big) = \left\{ \begin{array}{c} f(n) \text{: there exists a positive constant c such that} \\ \lim\limits_{n \to \infty} \left(\frac{f(n)}{g(n)}\right) \leq c \end{array} \right\}$$

$$o(g(n)) = \left\{ f(n) : \lim_{n \to \infty} \left( \frac{f(n)}{g(n)} \right) = 0 \right\}$$

### Asymptotic Strict Lower Bound (ω-notation)

$$\Omega \big(g(n)\big) = \left\{ \begin{array}{c} f(n) \text{: there exist positive constants $c$ and $n_0$ such that} \\ 0 \leq cg(n) \leq f(n) \text{ for all } n \geq n_0 \end{array} \right\}$$

$$\Omega \big(g(n)\big) = \left\{ \begin{array}{l} f(n) \text{: there exists a positive constant } c \text{ such that} \\ \lim\limits_{n \to \infty} \left(\frac{f(n)}{g(n)}\right) \geq c \end{array} \right.$$

$$\omega(g(n)) = \left\{ f(n) : \lim_{n \to \infty} \left( \frac{g(n)}{f(n)} \right) = 0 \right\}$$

### **Comparing Functions: Transitivity**

$$f(n) = O(g(n))$$
 and  $g(n) = O(h(n))$   $\Rightarrow$   $f(n) = O(h(n))$ 

$$f(n) = \Omega(g(n))$$
 and  $g(n) = \Omega(h(n))$   $\Rightarrow$   $f(n) = \Omega(h(n))$ 

$$f(n) = \Theta(g(n))$$
 and  $g(n) = \Theta(h(n))$   $\Rightarrow$   $f(n) = \Theta(h(n))$ 

$$f(n) = o(g(n))$$
 and  $g(n) = o(h(n))$   $\Rightarrow$   $f(n) = o(h(n))$ 

$$f(n) = \omega(g(n))$$
 and  $g(n) = \omega(h(n))$   $\Rightarrow$   $f(n) = \omega(h(n))$ 

4

94

### **Comparing Functions: Reflexivity**

$$f(n) = O(f(n))$$

$$f(n) = \Omega(f(n))$$

$$f(n) = \Theta(f(n))$$

 $f(n) = \Theta(g(n))$  if and only if  $g(n) = \Theta(f(n))$ 

**Comparing Functions: Symmetry** 

### **Comparing Functions: Transpose Symmetry**

$$f(n) = O\big(g(n)\big) \quad \text{if and only if} \quad g(n) = O\big(f(n)\big)$$
 
$$f(n) = O\big(g(n)\big) \quad \text{if and only if} \quad g(n) = O\big(f(n)\big)$$

98

100

### **Adding Functions**

$$O(f(n)) + O(g(n)) = O(f(n) + g(n))$$
  

$$\Omega(f(n)) + \Omega(g(n)) = \Omega(f(n) + g(n))$$

$$\Theta(f(n)) + \Theta(g(n)) = \Theta(f(n) + g(n))$$

99

### **Multiplying Functions by Constants**

$$0(c f(n)) = 0(f(n))$$

$$\Omega\bigl(cf(n)\bigr)=\Omega\bigl(f(n)\bigr)$$

$$\Theta(cf(n)) = \Theta(f(n))$$



# **Multiplying Two Functions**

$$O(f(n)) \times O(g(n)) = O(f(n) \times g(n))$$

$$\Omega(f(n)) \times \Omega(g(n)) = \Omega(f(n) \times g(n))$$

$$\Theta(f(n)) \times \Theta(g(n)) = \Theta(f(n) \times g(n))$$

### **Division of Functions**

$$\frac{O(f(n))}{O(g(n))} = O\left(\frac{f(n)}{g(n)}\right)$$

$$\frac{\Omega(f(n))}{\Theta(g(n))} = \Omega\left(\frac{f(n)}{g(n)}\right)$$

$$\frac{\Theta(f(n))}{\Theta(g(n))} = \Theta\left(\frac{f(n)}{g(n)}\right)$$

$$\frac{O(f(n))}{\Omega(g(n))} = O\left(\frac{f(n)}{g(n)}\right) \qquad \qquad \frac{\Omega(f(n))}{O(g(n))} = \Omega\left(\frac{f(n)}{g(n)}\right)$$

$$\frac{\Theta(f(n))}{\Omega(g(n))} = O\left(\frac{f(n)}{g(n)}\right) \qquad \qquad \frac{\Theta(f(n))}{O(g(n))} = \Omega\left(\frac{f(n)}{g(n)}\right)$$

$$\frac{\Omega(f(n))}{O(g(n))} = \Omega\left(\frac{f(n)}{g(n)}\right)$$

$$\frac{\Theta(f(n))}{\Omega(g(n))} = O\left(\frac{f(n)}{g(n)}\right)$$

$$\frac{\Theta(f(n))}{O(g(n))} = \Omega\left(\frac{f(n)}{g(n)}\right)$$

102

### **Growth Rates of Common Functions**

The following table shows how much time an algorithm that performs f(n) operations on an input of size n takes assuming each operation takes one nanosecond ( $10^{-9}$  seconds) to execute.

n $f(n)$	$\lg n$	n	$n \lg n$	$n^2$	$2^n$	n!
10	$0.003~\mu s$	$0.01~\mu s$	$0.033~\mu s$	$0.1~\mu s$	1 μs	3.63 ms
20	$0.004~\mu s$	$0.02 \ \mu s$	$0.086~\mu s$	$0.4~\mu s$	1 ms	77.1 years
30	$0.005~\mu s$	$0.03~\mu s$	$0.147 \ \mu s$	$0.9~\mu s$	1 sec	$8.4 \times 10^{15} \text{ yrs}$
40	$0.005~\mu s$	$0.04~\mu s$	$0.213~\mu s$	$1.6~\mu s$	18.3 min	
50	$0.006~\mu s$	$0.05~\mu s$	$0.282~\mu s$	$2.5~\mu s$	13 days	
100	$0.007~\mu s$	$0.1~\mu s$	$0.644~\mu s$	10 μs	$4 \times 10^{13} \text{ yrs}$	
1,000	$0.010 \ \mu s$	$1.00 \ \mu s$	$9.966~\mu s$	1 ms		
10,000	$0.013~\mu s$	$10 \mu s$	$130~\mu s$	100 ms		
100,000	$0.017~\mu s$	0.10  ms	$1.67 \mathrm{\ ms}$	10 sec		
1,000,000	$0.020 \ \mu s$	1  ms	19.93  ms	16.7 min		
10,000,000	$0.023~\mu s$	$0.01  \mathrm{sec}$	0.23 sec	1.16 days		
100,000,000	$0.027~\mu s$	0.10 sec	2.66 sec	115.7 days		
1,000,000,000	$0.030~\mu s$	1 sec	29.90 sec	31.7 years		

Source: "The Algorithm Design Manual" (2nd Edition, Springer, 2008) by Steven Skiena







104

