

#### **COMPUTER ALGORITHMS**

#### Athasit Surarerks





### Introduction



#### **EUCLID's GAME**

Two players move in turn. On each move, a player has to write on the board a positive integer equal to the different from two numbers already in the board; this number must be new. The player who cannot move loses the game.



#### **ADDITION'S GAME**

Player I and II alternately choose integers, each choice being one of the integers 1, 2, ..., *k*, and each choice being made with knowledge of all preceding choices. As soon as the sum of the chosen integers exceeds *N*, the last player to choose loses.



### WORLD PUZZLE

► our people want to cross a bridge; they are all begin in the same side. You must have 17 minutes to get them all across to the other side. It is night, and they have one flashlight. A maximum of two people can cross the bridge at one time. Any party that crosses must have the flashlight with them. Each person walks at a different speed: 1, 2, 5, and 10. A pair must walk together at the rate of the slower person.



## **Problem solving**





# **Problem solving**





# **Complex number**

Given two complex numbers,

X = a + bjY = c + dj

4.02 USD

Find an algorithm to perform the product XY.

Cost of a multiplier is 1USD.Cost of an adder is 0.01USD.

(*ac-bd*) + (*ad+bc*) j



# **Complex number**

Given two complex numbers,

X = a + bjY = c + dj

Find an algorithm to perform the product XY.

Cost of a multiplier is 1USD.Cost of an adder is 0.01USD.

3.05 USD(a+b)(c+d) = ac+ad+bc+bd























## **Problem solving**

Communication Language translation Function Composition function Algorithms











month











then each pair produces another pair each month.

month 5

Assuming that no rabbits ever die, how many pairs of rabbits after *n* months.



### **Problem size**

Traveling salesperson Minimum spanning tree String matching Independent set Sum of subset Linear partition



### Introduction Analysis of algorithm Design of algorithm Complexity of algorithm



### Introduction Problems Asymptotic notations



Introduction Analysis of algorithm Worst-case analysis Average-case analysis Amortized analysis



Introduction Analysis of algorithm Design of algorithm **Divide-and-conquer Dynamic programming** Approximation algorithm



### Introduction Analysis of algorithm Design of algorithm Complexity of algorithm

### ALGOL statement



program-name(variables listed) begin statement(s) end variable ← expression if condition then statement(s) else statement(s) endif while condition do statement(s) enddo for variable ← initial-value step size until final-value do statement(s) enddo goto label return variables listed input variables listed output variables listed

any other miscellaneous statement



### Maximum-Minimum search





### Maximum-Minimum search

- 011 max-search (A[1..*n*])
- 012 begin
- 013 if n = 1 then return A[1] endif
- 014 answer  $\leftarrow$  max-search(A[1..*n*-1])
- 015 if A[n] > answer then return A[n] else return answer endif 016 end

Time used :  $T_{A2}(n) = T_{A2}(n-1) + t_{013} + t_{014} + t_{015} = nt_{013} + (n-1)(t_{014} + t_{015}).$ 



### Efficiency

he efficiency of an algorithm is the resources used to find an answer. It is usually measured in terms of the theoretical computations, such as comparisons or data moves, the memory used, the number of messages passed, the number of disk accesses, etc.


#### Efficiency

Most often we shall be interested in the rate of growth of the time or space required to solve larger and larger instances of problem. We would like to associate with a problem and integer, called the size of the problem, which is measure of the quantity of input data.



#### Efficiency

The time needed by an algorithm expressed as a function of the size of a problem is called the time complexity of the algorithm. The limiting behavior of the complexity as size increases is called the asymptotic time complexity. Analogous definitions can be made for space complexity and asymptotic space complexity.



# Complexity

Algorithm	Growth rate	Time complexity
A <sub>1</sub>	1000 <i>n</i>	n
A <sub>2</sub>	100 <i>n</i> log <i>n</i>	n log n
$A_3$	10 <i>n</i> ²	n²
A <sub>4</sub>	n <sup>3</sup>	n <sup>3</sup>
A <sub>5</sub>	2 <sup>n</sup>	2 <sup>n</sup>



#### **Growth rate**





#### **Growth rate**

Let *f* and *g* be two functions, and if f(n) < g(n)

if and only if,  $\lim_{n\to\infty} f(n)/g(n) = 0.$ 



### Exercises

#### Compare the growth rate of

- 0.5<sup>n</sup>
- 1
- log<sub>10</sub> n
- n
- n 10<sup>n</sup>  $0.5^n < 1 < \log n < n < 10^n$



# Exercises

- Compare the growth rate of
  - $\ln^9 n \text{ and } n^{0.1}$ ,
  - *n*<sup>100</sup> and 2<sup>*n*</sup>.



# L'hÔpital's rules

Let *f* and *g* be two functions, and if  $\lim_{n\to\infty} f(n) = \infty,$   $\lim_{n\to\infty} g(n) = \infty$ 

Then

 $\lim_{n\to\infty} f(n)/g(n) = \lim_{n\to\infty} f(n)'/g(n)'.$ 



# Asymptotic notations

It is the asymptotic complexity of an algorithm which ultimately determines the size of problems that can be solved by the algorithm.



#### ω

A theoretical measure of the execution of an algorithm, usually the time or memory needed, given the problem size n, which is usually the number of items. Informally, saying some equation f(n) = (g(n)) means g(n) becomes insignificant relative to f(n) as n goes to infinity. More formally, it means that for any positive constant c, there exists a constant k, such that  $0 \le cg(n) < f(n)$  for all  $n \ge k$ . The value of k must not depend on n, but may depend on c. That is

 $\omega(g(n)) = \{f(n) \mid \lim_{n \to \infty} (f(n)/g(n)) = \infty \}.$ 



#### Θ

A theoretical measure of the execution of an algorithm, usually the time or memory needed, given the problem size n, which is usually the number of items. Informally, saying some equation  $f(n) = \Theta(g(n))$  means it is within a constant multiple of g(n). More formally, it means there are positive constants  $c_1$ ,  $c_2$ , and k, such that  $0 \le c_1 g(n) \le f(n) \le c_2 g(n)$ 

#### $\Theta(g(n)) = \{f(n) \mid \lim_{n \to \infty} (f(n)/g(n)) = c, c \neq 0, c \neq \infty\}.$

for all  $n \ge k$ . The values of  $c_1$ ,  $c_2$ , and k must be fixed for the function f and must not depend on n.



# **Big-O**

A theoretical measure of the execution of an algorithm, usually the time or memory needed, given the problem size n, which is usually the number of items. Informally, saying some equation f(n) = O(g(n)) means it is less than some constant multiple of g(n). More formally it means there are positive constants c and k, such that  $0 \le f(n) \le cg(n)$ 

 $O(g(n)) = \{f(n) \mid \lim_{n \to \infty} (f(n)/g(n)) = c, c \neq \infty\}.$ 

$$O(g(n)) = o(g(n)) \cup \Theta(g(n)).$$

for all  $n \ge k$ . The values of *c* and *k* must be fixed for the function *f* and must not depend on *n*.



# **Big-O**

The importance of this measure can be seen in trying to decide whether an algorithm is adequate, but may just need a better implementation, or the algorithm will always be too slow on a big enough input. For instance, quicksort, which is  $O(n \log n)$  on average, running on a small desktop computer can beat bubble sort, which is  $O(n^2)$ , running on a supercomputer if there are a lot of numbers to sort. To sort 1,000,000 numbers, the quicksort takes 6,000,000 steps on average, while the bubble sort takes 1,000,000,000,000 steps!



# $\Omega$

A theoretical measure of the execution of an algorithm, usually the time or memory needed, given the problem size n, which is usually the number of items. Informally, saying some equation  $f(n) = \Omega(g(n))$  means it is more than some constant multiple of g(n). More formally, it means there are positive constants c and k, such that  $0 \le cg(n) \le f(n)$ 

$$\Omega(g(n)) = \omega(g(n)) \cup \Theta(g(n)).$$

for all  $n \ge k$ . The values of *c* and *k* must be fixed for the function *f* and must not depend on *n*.



### Some properties

Asymptotic	Transitivity	Reflexivity	Symmetry	
Little-o	Yes	No	No	
ω	Yes	No	No	
Θ	Yes	Yes	Yes	
Big-O	Yes	Yes	No	
Ω	Yes	Yes	No	



### Some properties

Asymptotic	Transitivity	Reflexivity	Symmetry	
Little-o	Yes	No	No	
ω	Yes	No	No	
Θ	Yes	Yes	Yes	
Big-O	Yes	Yes	No	
Ω	Yes	Yes	No	



### Transpose Symmetry

f(n) = O(g(n)) if and only if  $g(n) = \Omega(f(n))$ f(n) = o(g(n)) if and only if  $g(n) = \omega(f(n))$ .



#### **Asymptotic Notations**



 $f(x) = \Omega(g(x))$ For all  $x \ge k_{2'}$ 

 $f(x) \ge c_1 g(x).$ 

f(x) = O(g(x))For all  $x \ge k_1$ ,

 $f(x) \leq c_2 g(x).$ 

$$\begin{split} f(x) &= \Theta(g(x)) \\ \text{For all } x \geq k_1, \ k_2, \\ &\quad c_1 g(x) \leq f(x) \leq c_2 g(x). \end{split}$$



#### **Asymptotic Notations**

#### SPECIAL ORDERS OF GROWTH

constant logarithmic :  $\Theta(\log n)$ linear quadratic :  $\Theta(n^2)$ 

: Θ(1) polylogarithmic :  $\Theta(\log^c n)$  ,  $c \ge 1$ sublinear :  $\Theta(n^a)$  , 0 < a < 1 $: \Theta(n)$ polynomial :  $\Theta(n^c)$  ,  $c \ge 1$ exponential :  $\Theta(c^n)$  , c > 1



# Exercises

- Show that
  - $\log n! = \Theta(n \log n),$
  - $\log_a n = \Theta(\log_b n)$ .





# Analysis of algorithm



A detail unambiguous sequence of actions to perform to accomplish some task. Named after a Persian mathematician Abu Ja'far Mohammed ibn Mûsâ Al-Khawarizmi who wrote a book with arithmetic rules dating from about 825 A.D.

#### Al-Khawarizmi



### CRITERIAS

- Correctness
- Amount of work done
- Amount of space used
  - Simplicity
  - Optimality



#### Maximum-Minimum search





#### Maximum-Minimum search

- 011 max-search (A[1..*n*])
- 012 begin
- 013 if n = 1 then return A[1] endif
- 014 answer  $\leftarrow$  max-search(A[1..*n*-1])
- if A[n] > answer then return A[n] else return answer endifend

Time used :  $T_{A2}(n) = T_{A2}(n-1) + t_{013} + t_{014} + t_{015} = nt_{013} + (n-1)(t_{014} + t_{015}).$   $T_{A2}(n) = n\Theta(t_{013}) + (n-1)(\Theta(t_{014}) + \Theta(t_{015}))$  $T_{A2}(n) = \Theta(n)$ 



#### RUNNING TIME ANALYSIS

#### To simplify running time analysis,

- Count only "barometer" instructions
  - Use asymptotic analysis



### **Basic instruction**

To analyze total time used of an algorithm, we usually translate each operation into basic operation such that it has a constant time complexity,  $\Theta(1)$ .

**Example:** Which statements are basic?

STATEMENT	Y/N
if a < b then a $\leftarrow$ a+1 else a $\leftarrow$ a-1 endif.	YES
if b < c then e $\leftarrow$ c! endif	NO
$d \leftarrow c^n$	NO
goto 005	YES
call max(A[1n])	NO
return $\leftarrow$ subroutine(A,B)	NO



# **Sorting Algorithm**

001 002 003 004 005 006 007	be foi	for i	n]) - n step –1 until 2 do for i ← 2 step 1 until last do if A[i] < A[i-1] thenbuffer ← A[i] A[i] ← A[i-1] A[i-1] ← buffer endif enddo						
008	er	d	Operation	Times	Complexity				
$\langle \rangle$			003	n	Θ( <i>n</i> )				
			004	$\sum_{2 \le \text{last} \le n} \text{last}$	Θ( <i>n</i> ( <i>n</i> +1)/2 - 1)				
			005	<u>∑</u> 2≤last≤ <i>n</i> last-1	Θ( <i>n</i> ( <i>n</i> -1)/2)				
				– Total	$\Theta(n^2)$				



# Sequencing

Give two operations,  $S_1$  and  $S_2$ , such that  $S_2$ will be processed after  $S_1$  sequentially. Total time complexity depends on the complexity of each operation.

Upper bound of  $S_1S_2 = \max(\text{ upper bound } S_1, \text{ upper bound } S_2)$ Lower bound of  $S_1S_2 = \max(\text{ lower bound } S_1, \text{ lower bound } S_2)$ 



# Sequencing

	S <sub>1</sub>	Time low/up	S <sub>2</sub>	Time low/up	Time low/up	$S_1S_2$
	Θ( <i>n</i> )		Θ( <i>n</i> )			
	Θ(n)		O( <i>n</i> )			
	Ð(Ŋ)		O( <i>n</i> <sup>2</sup> )			
	Θ( <i>n</i> ²)		Θ( <i>n</i> )			



# Sequencing

	S <sub>1</sub>	Time low/up	S <sub>2</sub>	Time low/up	Time low/up	$S_1S_2$
	Θ( <i>n</i> )	Ω( <i>n</i> ) Ο( <i>n</i> )	Θ( <i>n</i> )	Ω( <i>n</i> ) Ο( <i>n</i> )	Ω( <i>n</i> ) Ο( <i>n</i> )	Θ( <i>n</i> )
	Θ(n)	Ω( <i>n</i> ) Ο( <i>n</i> )	O( <i>n</i> )	unknown O( <i>n</i> )	Ω( <i>n</i> ) Ο( <i>n</i> )	Θ( <i>n</i> )
	Θ(Ŋ)	Ω( <i>n</i> ) Ο( <i>n</i> )	O( <i>n</i> <sup>2</sup> )	unknown O( <i>n</i> ²)	Ω( <i>n</i> ) Ο( <i>n</i> ²)	O( <i>n</i> <sup>2</sup> )
	Θ( <i>n</i> ²)	$\Omega(n^2)$ $O(n^2)$	Θ( <i>n</i> )	Ω( <i>n</i> ) Ο( <i>n</i> )	Ω( <i>n</i> ²) Ο( <i>n</i> ²)	Θ( <i>n</i> ²)



### If statement

Given three commands,  $S_1$ ,  $S_2$  and  $S_3$ .

Let  $t_1$ ,  $t_2$  and  $t_3$  be time complexity of  $S_1$ ,  $S_2$  and  $S_3$  respectively.

Consider a statement

#### if $S_1$ then $S_2$ else $S_3$ endif

et *T* be time complexity of this statement, we have that:

Lower bound of  $T = t_1 + \min(t_2, t_3)$ 

Upper bound of  $T = t_1 + \max(t_2, t_3)$ 



### If statement

#### Example:

Find the time complexity of if  $S_1$  then  $S_2$  else  $S_3$  endif if  $t_1 = \Theta(2^n)$ ,  $t_2 = O(n)$  and  $t_3 = O(n \log n)$ .

Upper bound of *T* 

 $= \Omega(2^n) + \min(\text{ unknown, unknown})$  $= \Omega(2^n).$ Lower bound of  $T = O(2^n) + max(O(n), O(n \log n))$  $= O(2^n) + O(n \log n)$ 

$$= \mathcal{O}(2^n).$$

Time complexity is  $\Theta(2^n)$ .



# If statement

#### Exercise

Find the time complexity of if  $\Theta(n \log n)$  then  $O(n^2)$  else  $\Theta(\log n)$  endif

O(*n*<sup>2</sup>)

if  $O(n^3)$  then  $O(n \log n)$  else  $\Theta(2^n)$  endif

 $O(2^{n})$ 

if O(n!) then  $\Theta(n^2)$  else  $\Theta(n)$  endif

O(*n*!)

if  $\Theta(n)$  then  $O(n \log n)$  else  $\Theta(n \log n)$  endif

O(*n* log *n*)



#### **For-loop Statement**

Time complexity can be computed by summary of time of each loop.

Let  $t_i$  be time used in the  $i^{th}$  iteration. Then total time used is equal to

$$\Sigma_{\text{all }i} t_{i}$$



#### **For-loop Statement**

**Example:** Find time complexity of

001for last  $\leftarrow n$  step -1 until 2 do002for  $i \leftarrow 2$  step 1 until *last* do003if A[i] < A[i-1] then swap(A[i],A[i-1]) endif</td>004enddo005enddo

Time complexity

$$= \sum_{2 \le last \le n} \sum_{2 \le i \le last+1} \Theta(1)$$
  
=  $\sum_{2 \le last \le n} \Theta(last)$   
=  $\Theta(\sum_{2 \le last \le n} last)$   
=  $\Theta(n(n+1)/2 - 1)$   
=  $\Theta(n^2)$ .


## **For-loop Statement**

#### **Example:** Find time complexity of

```
001for i \leftarrow 1 step 1 until m do002for j \leftarrow m step -1 until 1 do003for k \leftarrow 1 step 1 until j do004sum \leftarrow sum + i + j + k005enddo006enddo007enddo
```

Time complexity

$$= \sum_{1 \le i \le m} \left( \sum_{1 \le j \le m} \left( \sum_{1 \le k \le j} \left( \Theta(1) \right) \right) \right)$$
  
$$= \sum_{1 \le i \le m} \left( \sum_{1 \le j \le m} \Theta(\sum_{1 \le k \le j} k) \right)$$
  
$$= \sum_{1 \le i \le m} \left( \sum_{1 \le j \le m} \Theta(j) \right)$$
  
$$= \sum_{1 \le i \le m} \left( \Theta(m^2) \right)$$
  
$$= \Theta(m^3)$$



### **While-loop Statement**

By the same way, we have to count the number of iterations.

Example:Find the time complexity of001while (n > 0) do002 $n \leftarrow \lfloor n/2 \rfloor$ 003end do

Consider the value of n in each loop. For the *i*-th iteration, the value of n is equal to  $n/2^i$ . Since the number of iterations is equal to  $\log_2 n$ , time complexity =  $\Theta(\log_2 n)$ .



### **While-loop Statement**

Example: Fine the complexity of





### **While-loop Statement**

**Definition:** An algorithm to compute the greatest common divisor of two integers. It is Euclid(a,b){if (b=0) then return a; else return Euclid(b, a mod b);}.

#### **Algorithm Euclid's GCD**



GCD(a,b) 001 b > abegin 002 while (a > o) do 003 004 q ← a 005 a ← b mod a 006  $b \leftarrow q$ 007 enddo 800 return b 009 end



It is important to see that the size of problem should be decreased for each iteration of recursive called. We can count all instructions by

1.Set run time complexity = T(n) where *n* is the size of problem.

2.Time complexity of recursive call statement = T(m) where m is the input size and m < n.

3.Let T(n) = T(m) + time complexity of other instructions.

4. Find T(n) in term of asymptotic complexity function.



### **Algorithm Selection Sort**

001 SelectionSort(A{1..*n*])

002 begin

- if ( $n \le 1$ ) then return endif 003
- maxindex  $\leftarrow$  Max(A[1..*n*]) 004 buffer  $\leftarrow A[n]$

005

006  $A[n] \leftarrow A[maxindex]$ 

A[maxindex]  $\leftarrow$  buffer 007

SelectionSort(A[1..*n*-1]) 800 009

end

Time complexity T(n)

$$= T(n-1) + \Theta(n)$$
  
= T(n-2) +  $\Theta(n-1) + \Theta(n)$   
= ...  
=  $\sum_{1 \le i \le n} \Theta(i) = \Theta(n^2).$ 



#### **Algorithm Binary Search**

001 BinarySearch(A[1..*n*],x)

002 | begin

- 003 | if (A[1] > A[n]) then return –1 endif
- 004 | mid  $\leftarrow$  (A[1] + A[*n*])/2
- 005 | if ( x = A[mid] ) then return mid endif

006 if (x < A[mid]) then return BinarySearch(A[1..mid-1],x)</li>
007 else return BinarySearch(A[mid+1..n],x) endif
008 end

Time complexity  $T(n) \leq T(n/2) + \Theta(1)$ 

```
\leq T(n/2) + \Theta(1)

\leq T(n/2^{2}) + \Theta(1) + \Theta(1)

\leq T(n/2^{k}) + \sum_{1 \leq i \leq k} \Theta(i)

\leq \Theta(1) + \sum_{1 \leq i \leq \log n} \Theta(1)

\leq \Theta(\log_{2} n)

= O(\log_{2} n).
```



The recursive call in a given algorithm can be represented by using recursive tree. For example, the run time complexity of a given algorithm is

 $T(n) = 2T(n/2) + \Theta(n)$ , for any n > 1 and  $T(1) = \Theta(1)$ .





T(n) = T(n/2) + T(n/3), for any n > 1 and  $T(1) = \Theta(1)$ .





T(n) = T(n/a) + T(n/b), for any  $n \ge 1$  and  $T(1) = \Theta(1)$ .











T(n) = aT(n/b) + f(n) where  $a \ge 1$  and  $b \ge 1$ 





#### $T(n) = aT(n/b) + f(n) \text{ where } a \ge 1 \text{ and } b \ge 1$

This complexity depends on three variables, a, b and f(n).

f(n)	T( <i>n</i> )
$O(n^{(\log b a)-\epsilon})$ , for fixed $\epsilon > 0$	O(n <sup>(logb a)</sup> )

This means that the growth rate of f(n) is less than  $O(n^{(\log b a)})$ . Then total time complexity is  $O(n^{(\log b a)})$ .



#### T(n) = aT(n/b) + f(n) where $a \ge 1$ and $b \ge 1$

This complexity depends on three variables, a, b and f(n).

<i>f</i> ( <i>n</i> )	T( <i>n</i> )
$O(n^{(\log b a)-\varepsilon})$ , for fixed $\varepsilon > 0$	O(n <sup>(logb a)</sup> )
Θ(n <sup>(logb a)</sup> )	$\Theta(n^{(\log b a)} \log n)$

It is clear that total time complexity is equal to  $\Theta(n^{(\log b a)} \log n)$ .



#### T(n) = aT(n/b) + f(n) where $a \ge 1$ and $b \ge 1$

This complexity depends on three variables, a, b and f(n).

This means that the growth rate of f(n) is more than  $\Omega(n^{(\log b a)})$ , and the condition  $af(n/b) \le cf(n)$  for c < 1 and  $n > n_0$  means that the growth rate of f(n) decreases where *n* increases. Then total growth rate is  $\Theta(f(n))$ .

$$\begin{array}{l} \Omega(n^{(\log b \ a)+\varepsilon}), \ \text{for fixed } \varepsilon > 0\\ \text{and}\\ af(n/b) \leq cf(n)\\ \text{for } c < 1 \ \text{and} \ n > n_0 \end{array} \qquad \Theta(f(n)) \end{array}$$



### Example: Find the solution of

- T(n) = 16T(n/4) + n
- $T(n) = 27T(n/3) + n^3$
- $T(n) = 3T(n/4) + n \log n$ .

T(n) = 16T(n/4) + nSince  $\log_4 16 - 0.1 > 1$ , we can conclude that time complexity = O(n<sup>2</sup>).



### Example: Find the solution of

- T(n) = 16T(n/4) + n
- $T(n) = 27T(n/3) + n^3$
- $T(n) = 3T(n/4) + n \log n$ .

 $T(n) = 27T(n/3) + n^3$ Since  $\log_3 27 = 3$ , it is clear that time complexity =  $\Theta(n^3 \log n)$ .



### Example: Find the solution of

- T(n) = 16T(n/4) + n
- $T(n) = 27T(n/3) + n^3$
- $T(n) = 3T(n/4) + n \log n$ .

 $T(n) = 3T(n/4) + n \log n.$ Since  $\log_4 3 < 0.8$ , and  $f(n) = n \log n = \Omega(n^{0.8+0.2}).$ We also have that  $3f(n/4) \le (3n/4)\log n.$ Then time complexity =  $\Theta(n \log n).$ 



# **Amortized analysis**



## Three methods

Aggregate method Accounting method Potential method