

Announcement

- Any question should be sent via email
 - Please use [ADA2018] in the subject
 - Please write your 學號 姓名 in the email
- Registration codes were sent out
 - Register (or drop?) the course ASAP
- Slides are available before the lecture starts
- Mini-HW 1 released
 - Due on 9/27 (Thu) 14:20
 - Submit to NTU COOL
- Judge system available
 - Programming part: submit to Online Judge <u>http://ada18-judge.csie.org</u>

			>_		
ļ	Algorit	:hm De	esign an	d Anal	ysis
		n University, Con Lectured by	nputer Science and Yun-Nung (Vivian) C	Information Engi <u>Chen</u>	
About	Syllabus	Homework	Teaching Team	NTU COOL	Judge Syster



2018



g



Mini-HW #1



Mini HW #1

Due Time: 2018/9/27 (Thu.) 14:20

Contact TAs: ada-ta@csie.ntu.edu.tw

Problem 1

Let f(n) = g(n) - h(n). Given $g(n) = \Theta(F(n))$ and h(n) = o(F(n)), prove or disprove $f(n) = \Omega(F(n))$. (Use the definitions of Θ , o and Ω given in textbook.)



M The

Outline

- Terminology
 - Problem (問題)
 - Problem instance (個例)
 - Computation model (計算模型)
 - Algorithm (演算法)
 - The hardness of a problem (難度)
- Algorithm Design & Analysis Process
- Review: Asymptotic Analysis
- Algorithm Complexity
- Problem Complexity



Efficiency Measurement = Speed

Why we care?

- Computers may be fast, but they are not infinitely fast
- Memory may be inexpensive, but it is not free





Terminology

Textbook Ch. 1 – The Role of Algorithms in Computing





The champion problem

- Input: n distinct integers $A[1], A[2], \ldots, A[n]$.
- Output: the index i with $1 \le i \le n$ such that

$$A[i] = \max_{1 \le j \le n} A[j].$$



Problem Instance (個例)

An instance of the champion problem

5 distinct integers 7, 4, 2, 9, 8.





Computation Model (計算模型)

- Each problem must have its rule (遊戲規則)
- Computation model (計算模型) = rule (遊戲規則)
- The problems with different rules have different hardness levels





Hardness (難易程度)

How difficult to solve a problem

- Example: how hard is the champion problem?
- Following the comparison-based rule





Problem Solving (解題)

Definition of "solving" a problem

 Giving an algorithm (演算法) that produces a correct output for any instance of the problem.



Algorithm (演算法)

- Algorithm: a detailed step-by-step instruction
 - Must follow the game rules
 - Like a step-by-step recipe
 - Programming language doesn't matter
 - \rightarrow problem-solving recipe (technology)
- If an algorithm produces a correct output for <u>any instance</u> of the problem
 - \rightarrow this algorithm "solves" the problem



"A well-defined computational procedure that transforms some input to some output"



Hardness (難度)

- Hardness of the problem
 - How much effort the best algorithm needs to solve any problem instance
- 防禦力
 - 看看最厲害的賽亞人要花多少攻擊力才能打贏對手





Algorithm Design & Analysis Process

Algorithm Design & Analysis Process

- 1) Formulate a **problem**
- 2) Develop an algorithm
- 3) Prove the correctness
- 4) Analyze **running time/space** requirement







1. Problem Formulation

The champion problem

- Input: n distinct integers $A[1], A[2], \ldots, A[n]$.
- Output: the index i with $1 \le i \le n$ such that

$$A[i] = \max_{1 \le j \le n} A[j].$$



2. Algorithm Design

- Create a detailed recipe for solving the problem
 - Follow the comparison-based rule
 - 不准偷看信封的內容
 - 請別人幫忙「比大小」
- Algorithm: 擂台法
- 1. int *i, j*;

Q2: Does it solve the champion problem?

Q1: Is this a comparison-based algorithm?

- *2. j* = 1;
- 3. for (*i* = 2; *i* <= *n*; *i*++)
- 4. if (A[i] > A[j])
- 5. j = i;
- 6. return *j*;





3. Correctness of the Algorithm

Prove by contradiction (反證法)

The algorithm solves the champion problem.

Proof Let j^* be the correct answer. That is, $A[j^*] = \max\{A[1], \ldots, A[n]\}.$

- If $j^* = 1$, then Step 5 is never reached. Therefore, 1 is correctly returned.
- If j* > 1, then in the iteration of the for-loop with i = j*, j becomes j*. By definition of j*, A[j*] > A[i] holds for each i = j* + 1,...,n. Therefore, in the remaining iterations of the for-loop, the value of j does not change. Hence, at the end of the algorithm, j* is correctly returned.

1	int <i>i i</i> .	
±.	iiic 7, j,	
2.	<i>j</i> = 1;	
3.	for (<i>i</i> = 2; <i>i</i> <= <i>n</i> ; <i>i</i> ++)	
4.	if $(A[i] > A[j])$	
5.	j = i;	
6.	return <i>j</i> ;	



- How much effort the best algorithm needs to solve any problem instance
 - Follow the comparison-based rule
 - 不准偷看信封的內容
 - 請別人幫忙「比大小」
- Effort: we first use the times of comparison for measurement



- The hardness of the champion problem is (n 1) comparisons
 - a) There is an algorithm that can solve the problem using at most (n-1) comparisons
 - This can be proved by 擂臺法, which uses (n 1) comparisons for any problem instance
 - b) For any algorithm, there exists a problem instance that requires (*n* 1) comparisons
 - Why?



- Q: Is there an algorithm that only needs (*n* 2) comparisons?
- A: Impossible!
- Reason
 - A single comparison only decides a loser
 - If there are only (n 2) comparisons, the most number of losers is (n - 2)
 - There exists a least 2 integers that did not lose
 - \rightarrow any algorithm cannot tell who the champion is



Finding Hardness

Use the upper bound and the lower bound

 When they meet each other, we know the <u>hardness of the</u> problem





Upper bound

- how many comparisons are <u>sufficient</u> to solve the champion problem
- Each algorithm provides an upper bound
- The smarter algorithm provides tighter, lower, and better upper bound

多此一舉擂臺法

i = 1;

1. int *i, j*;

2.

5.

 \rightarrow (2*n* - 2) comparisons

- 3. for (*i* = 2; *i* <= *n*; *i*++)
- 4. if ((A[i] > A[j]) && (A[j] < A[i]))
 - j = i;
- 6. return *j*;

When upper bound = lower bound, the problem is solved. \rightarrow We figure out the hardness of the problem



- how many comparisons <u>in</u> <u>the worst case are necessary</u> to solve the champion problem
- Some arguments provide different lower bounds
- Higher lower bound is better

Every integer needs to be in the comparison once \rightarrow (*n*/2) comparisons



4. Algorithm Analysis

- The majority of researchers in algorithms studies the <u>time</u> and <u>space</u> required for solving problems in two directions
 - Upper bounds: designing and analyzing algorithms
 - Lower bounds: providing arguments
- When the upper and lower bounds match, we have an optimal algorithm and the problem is completely resolved





$\xrightarrow{\text{O}} Asymptotic Analysis$





Donald E. Knuth (1938-)

Motivation

- The hardness of the champion problem is exactly n-1 comparisons
- Different problems may have different 「難度量尺」
 - cannot be interchangeable
- Focus on the standard growth of the function to ignore the unit and coefficient effects



- For a problem P, we want to figure out
 - The hardness (complexity) of this problem P is $\Theta(f(n))$
 - n is the instance size of this problem P
 - *f*(*n*) is a function
 - $\Theta(f(n))$ means that "it exactly equals to the growth of the function"
- Then we can argue that under the comparison-based computation model
 - The hardness of the champion problem is $\Theta(n)$
 - The hardness of the sorting problem is $\Theta(n \log n)$



- Use the upper bound and the lower bound
- When they match, we know the hardness of the problem



- First learn how to analyze / measure the effort an algorithm needs
 - Time complexity
 - Space complexity
- Focus on worst-case complexity
 - "average-case" analysis requires the assumption about the probability distribution of problem instances

Турез	Description
Worst Case	Maximum running time for any instance of size <i>n</i>
Average Case	Expected running time for a random instance of size n
Amortized	Worse-case running time for a series of operations



Review of Asymptotic Notation (Textbook Ch. 3.1)

- f(n) = time or space of an algorithm for an input of size n
- Asymptotic analysis: focus on the **growth** of f(n) as $n \to \infty$



Review of Asymptotic Notation (Textbook Ch. 3.1)

- f(n) = time or space of an algorithm for an input of size n
- Asymptotic analysis: focus on the **growth** of f(n) as $n \to \infty$
- O, or Big-Oh: upper bounding function
- Ω, or Big-Omega: lower bounding function
- Θ, or Big-Theta: tightly bounding function





Formal Definition of Big-Oh (Textbook Ch. 3.1)

• For any two functions f(n) and g(n),

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

if there exist positive constants n_0 and c s.t.

$$0 \le f(n) \le c \cdot g(n)$$

for all $n \ge n_0$.

g(n)的某個常數倍 $c \cdot g(n)$ 可以在n夠大時壓得住f(n)



$f(n) = O\bigl(g(n)\bigr)$

- Intuitive interpretation
 - f(n) does not grow faster than g(n)
- Comments
 - 1) f(n) = O(g(n)) roughly means $f(n) \le g(n)$ in terms of rate of growth
 - 2) "=" is not "equality", it is like " ϵ (belong to)" The equality is $\{f(n)\} \subseteq O(g(n))$
 - 3) We do not write O(g(n)) = f(n)
- Note
 - f(n) and g(n) can be negative for some integers n
 - In order to compare using asymptotic notation O, both have to be <u>non-negative</u> for sufficiently large n
 - This requirement holds for other notations, i.e. Ω , Θ , o, ω



Review of Asymptotic Notation (Textbook Ch. 3.1)

- Benefit
 - Ignore the <u>low-order terms</u>, <u>units</u>, and <u>coefficients</u>
 - Simplify the analysis
- Example: $f(n) = 5n^3 + 7n^2 8$
 - Upper bound: f(n) = O(n³), f(n) = O(n⁴), f(n) = O(n³log₂n)
 - Lower bound: $f(n) = \Omega(n^3)$, $f(n) = \Omega(n^2)$, $f(n) = \Omega(n \log_2 n)$
 - Tight bound: f(n) = Θ(n³)

- "=" doesn't mean "equal to"
- Q: $f(n) = O(n^3)$ and $f(n) = O(n^4)$, so $O(n^3) = O(n^4)$?
 - O(n³) represents a set of functions that are upper bounded by cn³ for some constant c when n is large enough
 - In asymptotic analysis, "=" means "e (belong to)"

Exercise:
$$100n^2 = O(n^3 - n^2)$$
?

- Draft.
$$100n^2 \leq 100(n^3 - n^2)$$
$$\leftarrow 200n^2 \leq 100n^3$$
$$\leftarrow 2 \leq n$$

• Let $n_0 = 2$ and c = 100

$$100n^2 \le 100(n^3 - n^2)$$

holds for $n \ge 2$

$$100n^2 = O(n^3 - n^2)$$

Exercise: $n^2 = O(n)$?

- Disproof.
 - Assume for a contradiction that there exist positive constants c and n₀ s.t.

$$n^2 \le cn$$

holds for any integer n with $n \ge n_0$.

• Assume $n = 1 + \lceil \max(n_0, c) \rceil$

and because $\ n>n_0, n>c$, it follows that $n^2>cn$

Due to contradiction, we know that

$$n^2 \neq O(n)$$



Rules (Textbook Ch. 3.1)

The following statements hold for any real-valued functions f(n)and g(n), where there is a constant n_0 such that f(n) and g(n)are nonnegative for any integer $n \ge n_0$.

• Rule 1:
$$f(n) = O(f(n))$$
.

- Rule 2: If c is a positive constant, then $c \cdot O(f(n)) = O(f(n))$.
- Rule 3: If f(n) = O(g(n)), then O(f(n)) = O(g(n)).
- Rule 4: $O(f(n)) \cdot O(g(n)) = O(f(n) \cdot g(n)).$
- Rule 5: $O(f(n) \cdot g(n)) = f(n) \cdot O(g(n)).$



Other Notations (Textbook Ch. 3.1)

$$\begin{split} f(n) &= O(g(n)) \to f(n) \leq g(n) \text{ in rate of growth} \\ f(n) &= \Omega(g(n)) \to f(n) \geq g(n) \text{ in rate of growth} \\ f(n) &= \Theta(g(n)) \to f(n) = g(n) \text{ in rate of growth} \\ f(n) &= o(g(n)) \to f(n) < g(n) \text{ in rate of growth} \\ f(n) &= \omega(g(n)) \to f(n) > g(n) \text{ in rate of growth} \end{split}$$



- First learn how to analyze / measure the effort an algorithm needs
 - Time complexity
 - Space complexity
- Focus on worst-case complexity
 - "average-case" analysis requires the assumption about the probability distribution of problem instances

Using O to give upper bounds on the worst-case time complexity of algorithms





Rule 2

Algorithm Analysis

■ 擂台法			The worst-case time complexity is		
1.	int <i>i, j</i> ;	0(1) time	$O(1) + O(1) + O(n) \cdot (O$	(1) + O(1)) + O(1)	
2.	<i>j</i> = 1;	O(1) time	$3 \cdot O(1) + O(n) \cdot (2C)$	$\mathcal{P}(1))$	
3.	for (<i>i</i> = 2; <i>i</i> <= <i>n</i> ; <i>i</i> ++)	O(n) iterations	$=O(1) + O(n) \cdot O(1)$	Rule 2	
4.	if (A[<i>i</i>] > A[<i>j</i>])	O(1) time	=O(1)+O(n)	Rule 4	
5.	j = i;	O(1) time	=O(n)+O(n)	1 = O(n) & Rule 3	
6.	return <i>j</i> ;	O(1) time	$=2 \cdot O(n)$		

=O(n)

Adding everything together → an upper bound on the worst-case time complexity

40



Sorting Problem

• Input:

An array A of n distinct integers.

• Output:

Reorder A such that $A[1] < A[2] < \cdots < A[n]$.



Algorithm Analysis

- Bubble-Sort Algorithm
- O(1) time int *i*, *done*; 1. f(n) iterations do { 2. O(1) time done = 1;3. O(n) iterations for (*i* = 1; *i* < *n*; *i* ++) { 4. O(1) time if (A[i] > A[i + 1]) { 5. O(1) time exchange A[i] and A[i + 1]; 6. O(1) time done = 0;7. 8. $O(1) + f(n) \cdot (O(1) + O(n) \cdot O(1))$ } 9. $=O(1) + f(n) \cdot O(n)$ } while (*done* == 0) $=f(n) \cdot O(n)$ f(n) = O(n)10. $= O(n^2)$ prove by induction









- First learn how to analyze / measure the effort an algorithm needs
 - Time complexity
 - Space complexity
- Focus on worst-case complexity
 - "average-case" analysis requires the assumption about the probability distribution of problem instances

Using O to give **upper bounds** on the worst-case time complexity of algorithms Using Ω to give **lower bounds** on the worst-case time complexity of algorithms



Algorithm Analysis

▪ 擂台法

- 1. int *i*;
- 2. int *m* = *A*[1];
- 3. for (*i* = 2; *i* <= *n*; *i* ++) {
- 4. if (A[i] > m)
 - m = A[i];
- 6.

5.

7. return *m*;

}

 $\Omega(1)$ time $\Omega(1)$ time $\Omega(n)$ iterations $\Omega(1)$ time $\Omega(1)$ time

 $\Omega(1)$ time

 Adding everything together
 → a lower bound on the worstcase time complexity? $3 \cdot \Omega(1) + \Omega(n) \cdot (2 \cdot \Omega(1))$ $= \Omega(1) + \Omega(n) \cdot \Omega(1)$ $= \Omega(1) + \Omega(n)$ $= \Omega(n)$

45



Algorithm Analysis

• 百般無聊擂台法

 $\Omega(1)$ time int *i*; 1. $\Omega(1)$ time int m = A[1];2. $\Omega(n)$ iterations for (*i* = 2; *i* <= *n*; *i* ++) { 3. if (A[i] > m) $\Omega(1)$ time 4. $\Omega(1)$ time m = A[i];5. if (i == n) $\Omega(1)$ time 6. $\Omega(n)$ time do *i*++ *n* times 7. } 8. $\Omega(1)$ time return *m*; 9.

$$3 \cdot \Omega(1) + \Omega(n) \cdot (3 \cdot \Omega(1) + \Omega(n))$$

= $\Omega(1) + \Omega(n) \cdot \Omega(n)$
= $\Omega(1) + \Omega(n^2)$
= $\Omega(n^2)$

Adding together may result in errors. The safe way is to analyze using **problem instances**.

e.g. try A[i] = i or A[i]=2(n-i) to check the time complexity $\rightarrow \Omega(1)$





Bubble-Sort Algorithm

```
int i, done;
1.
                                     f(n) iterations
      do {
2.
         done = 1;
3.
         for (i = 1; i < n; i + +) \{ \Omega(n) \text{ time } \}
4.
           if (A[i] > A[i + 1]) {
5.
              exchange A[i] and A[i + 1];
6.
              done = 0;
7.
8.
9.
      } while (done == 0)
10.
```







Algorithm Complexity

In the worst case, what is the growth of function an algorithm takes

Time Complexity of an Algorithm

- We say that the (worst-case) time complexity of Algorithm A is $\Theta(f(n))$ if
- 1. Algorithm A runs in time O(f(n)) &
- 2. Algorithm A runs in time $\Omega(f(n))$ (in the worst case)
 - An input instance I(n) s.t. Algorithm A runs in $\Omega(f(n))$ for each n

Tightness of the Complexity

- If we say that the time complexity analysis about O(f(n)) is tight
- = the algorithm runs in time $\Omega(f(n))$ in the worst case
- = (worst-case) time complexity of the algorithm is $\Theta(f(n))$
 - Not over-estimate the worst-case time complexity of the algorithm
- If we say that the time complexity analysis of Bubble-Sort algorithm about ${\cal O}(n^2)$ is tight
- = Time complexity of Bubble-Sort algorithm is $\Omega(n^2)$
- = Time complexity of Bubble-Sort algorithm is $\Theta(n^2)$

Algorithm Analysis

• 百般無聊擂台法

- 1. int *i*;
- 2. int m = A[1];
- 3. for (*i* = 2; *i* <= *n*; *i* ++) {
- 4. if (A[i] > m)
 - m = A[i];
- 6. if (*i* == *n*)
- f (*i == n*) do *i*++ *n* times
- 8.

5.

7.

9. return *m*;

}

O(1) time O(n) iterations O(1) time O(1) time O(1) time O(n) time O(1) time

O(1) time

The worst-case time complexity of 「百般無聊擂臺法」is $\Theta(n)$.

non-tight analysis

 $3 \cdot O(1) + O(n) \cdot (3 \cdot O(1) + O(n))$ $= O(1) + O(n) \cdot O(n)$ $= O(1) + O(n^2)$ $= O(n^2)$

tight analysis

Step 3 takes O(n) iterations for the forloop, where only last iteration takes O(n)time and the rest take O(1) time. The steps 3-8 take time

 $O(n) \cdot O(1) + 1 \cdot O(n) = O(n)$

The same analysis holds for $\Omega(n)$



Algorithm Comparison

- Q: can we say that Algorithm 1 is a better algorithm than Algorithm 2 if
 - Algorithm 1 runs in O(n) time
 - Algorithm 2 runs in $O(n^2)$ time
- A: No! The algorithm with a lower upper bound on its worst-case time does not necessarily have a lower time complexity.

Comparing A and B



- Algorithm A is no worse than Algorithm B in terms of worst-case time complexity if there exists a positive function f(n) s.t.
 - Algorithm A runs in time O(f(n)) &
 - Algorithm *B* runs in time $\Omega(f(n))$ in the worst case
- Algorithm A is (strictly) better than Algorithm B in terms of worst-case time complexity if there exists a positive function f(n) s.t.
 - Algorithm A runs in time O(f(n)) &
 - Algorithm *B* runs in time $\omega(f(n))$ in the worst case

or

- Algorithm A runs in time o(f(n)) &
- Algorithm *B* runs in time $\Omega(f(n))$ in the worst case





Problem Complexity

In the worst case, what is the growth of the function the optimal algorithm of the problem takes

Time Complexity of a Problem

- We say that the (worst-case) time complexity of Problem P is $\Theta(f(n))$ if
- The time complexity of Problem P is O(f(n)) &
 There exists an O(f(n))-time algorithm that solves Problem P
- 2. The time complexity of Problem *P* is $\Omega(f(n))$
 - Any algorithm that solves Problem P requires $\Omega(f(n))$ time
- The time complexity of the champion problem is $\Theta(n)$ because
- 1. The time complexity of the champion problem is O(n) &
 - 。「擂臺法」is the O(n)-time algorithm
- 2. The time complexity of the champion problem is $\Omega(n)$
 - \circ Any algorithm requires $\Omega(n)$ time to make each integer in comparison at least once



Optimal Algorithm

- If Algorithm A is an optimal algorithm for Problem P in terms of worst-case time complexity
 - Algorithm A runs in time O(f(n)) &
 - The time complexity of Problem P is $\Omega(f(n))$ in the worst case
- Examples (the champion problem)
 - 擂台法 → optimal algorithm
 - It runs in O(n) time &
 - Any algorithm solving the problem requires time $\Omega(n)$ in the worst case
 - 百般無聊擂台法 → optimal algorithm
 - It runs in O(n) time &
 - Any algorithm solving the problem requires time $\Omega(n)$ in the worst case



Comparing *P* and *Q*



- Problem P is no harder than Problem Q in terms of (worst-case) time complexity if there exists a function f(n) s.t.
 - The (worst-case) time complexity of Problem P is O(f(n)) &
 - The (worst-case) time complexity of Problem Q is $\Omega(f(n))$
- Problem P is (strictly) easier than Problem Q in terms of (worst-case) time complexity if there exists a function f(n) s.t.
 - The (worst-case) time complexity of Problem P is O(f(n)) &
 - The (worst-case) time complexity of Problem *Q* is $\omega(f(n))$ or
 - The (worst-case) time complexity of Problem P is o(f(n)) &
 - The (worst-case) time complexity of Problem Q is $\Omega(f(n))$

Concluding Remarks

Algorithm Design and Analysis Process

- 1) Formulate a **problem**
- 2) Develop an algorithm
- 3) Prove the correctness
- 4) Analyze **running time/space** requirement
- Usually brute force (暴力法) is not very efficient
- Analysis Skills
 - Prove by contradiction
 - Induction
 - Asymptotic analysis
 - Problem instance
- Algorithm Complexity
 - In the worst case, what is the growth of function <u>an algorithm</u> takes
- Problem Complexity
 - In the worst case, what is the growth of the function the optimal algorithm of the problem takes

Design Step
Analysis Step



Reading Assignment

Textbook Ch. 3 – Growth of Function





Question?

Important announcement will be sent to @ntu.edu.tw mailbox & post to the course website

Course Website: http://ada.miulab.org

Email: ada-ta@csie.ntu.edu.tw