

Algorithms

Chapter 3

With Question/Answer
Animations

Chapter Summary

- Algorithms
 - Example Algorithms
 - Algorithmic Paradigms
- Growth of Functions
 - Big- O and other Notation
- Complexity of Algorithms

Algorithms

- Section 3.1

Section Summary₁

- Properties of Algorithms
- Algorithms for Searching and Sorting
- Greedy Algorithms
- Halting Problem

Problems and Algorithms

- In many domains there are key general problems that ask for output with specific properties when given valid input.
- The first step is to precisely state the problem, using the appropriate structures to specify the input and the desired output.
- We then solve the general problem by specifying the steps of a procedure that takes a valid input and produces the desired output. This procedure is called an algorithm.

Algorithms



Abu Ja'far
Mohammed Ibin
Musa Al-Khowarizmi
(780-850)

Definition: An algorithm is a finite set of precise instructions for performing a computation or for solving a problem.

Example: Describe an algorithm for finding the maximum value in a finite sequence of integers.

Solution: Perform the following steps:

1. Set the temporary maximum equal to the first integer in the sequence.
2. Compare the next integer in the sequence to the temporary maximum.
 - If it is larger than the temporary maximum, set the temporary maximum equal to this integer.
3. Repeat the previous step if there are more integers. If not, stop.
4. When the algorithm terminates, the temporary maximum is the largest integer in the sequence.

Specifying Algorithms

- Algorithms can be specified in different ways. Their steps can be described in English or in pseudocode.
- Pseudocode is an intermediate step between an English language description of the steps and a coding of these steps using a programming language.
- The form of pseudocode we use is specified in Appendix 3. It uses some of the structures found in popular languages such as C++ and Java.
- Programmers can use the description of an algorithm in pseudocode to construct a program in a particular language.
- Pseudocode helps us analyze the time required to solve a problem using an algorithm, independent of the actual programming language used to implement algorithm.

Properties of Algorithms

- **Input:** An algorithm has input values from a specified set.
- **Output:** From the input values, the algorithm produces the output values from a specified set. The output values are the solution.
- **Correctness:** An algorithm should produce the correct output values for each set of input values.
- **Finiteness:** An algorithm should produce the output after a finite number of steps for any input.
- **Effectiveness:** It must be possible to perform each step of the algorithm correctly and in a finite amount of time.
- **Generality:** The algorithm should work for all problems of the desired form.

Finding the Maximum Element in a Finite Sequence

The algorithm in pseudocode:

```
procedure max ( $a_1, a_2, \dots, a_n$ : integers)
  max :=  $a_1$ 
  for  $i$  := 2 to  $n$ 
    if max <  $a_i$  then max :=  $a_i$ 
  return max {max is the largest element}
```

Does this algorithm have all the properties listed on the previous slide?

Some Example Algorithm Problems

- Three classes of problems will be studied in this section.
 1. Searching Problems: finding the position of a particular element in a list.
 2. Sorting problems: putting the elements of a list into increasing order.
 3. Optimization Problems: determining the optimal value (maximum or minimum) of a particular quantity over all possible inputs.

Searching Problems

- **Definition:** The general searching problem is to locate an element x in the list of distinct elements a_1, a_2, \dots, a_n , or determine that it is not in the list.
- The solution to a searching problem is the location of the term in the list that equals x (that is, i is the solution if $x = a_i$) or 0 if x is not in the list.
- For example, a library might want to check to see if a patron is on a list of those with overdue books before allowing him/her to checkout another book.
- We will study two different searching algorithms: linear search and binary search.

Linear Search Algorithm

- The linear search algorithm locates an item in a list by examining elements in the sequence one at a time, starting at the beginning.
 - First compare x with a_1 . If they are equal, return the position 1.
 - If not, try a_2 . If $x = a_2$, return the position 2.
 - Keep going, and if no match is found when the entire list is scanned, return 0.

```
procedure linear search ( $x$  : integer,  
                         $a_1, a_2, \dots, a_n$  : distinct integers)  
   $i := 1$   
  while ( $i \leq n$  and  $x \neq a_i$ )  
     $i := i + 1$   
  if  $i \leq n$  then location :=  $i$   
  else location := 0  
  return location {location is the subscript of the term  
                   that equals  $x$ , or is 0 if  $x$  is not found}
```

Binary Search

- Assume the input is a list of items in increasing order.
- The algorithm begins by comparing the element to be found with the middle element.
 - If the middle element is lower, the search proceeds with the upper half of the list.
 - If it is not lower, the search proceeds with the lower half of the list (through the middle position).
- Repeat this process until we have a list of size 1.
 - If the element we are looking for is equal to the element in the list, the position is returned.
 - Otherwise, 0 is returned to indicate that the element was not found.
- In Section 3.3, we show that the binary search algorithm is much more efficient than linear search.

Binary Search₂

- Here is a description of the binary search algorithm in pseudocode.

```
procedure binary search ( $x$  : integer,  $a_1, a_2, \dots, a_n$  : increasing integers)
   $i := 1$  { $i$  is the left endpoint of interval}
   $j := n$  { $j$  is right endpoint of interval}
  while  $i < j$ 
     $m := \lfloor (i + j) / 2 \rfloor$ 
    if  $x > a_m$  then  $i := m + 1$ 
    else  $j := m$ 
  if  $x = a_i$  then location :=  $i$ 
  else location := 0
  return location {location is the subscript  $i$  of the term  $a_i$  equal to  $x$ ,
                    or 0 if  $x$  is not found}
```

Binary Search₃

- **Example:** The steps taken by a binary search for 19 in the list:

1 2 3 5 6 7 8 10 12 13 15 16 18 19 20 22

1. The list has 16 elements, so the midpoint is 8. The value in the 8th position is 10. Since $19 > 10$, further search is restricted to positions 9 through 16.

1 2 3 5 6 7 8 10 12 13 15 16 18 19 20 22

2. The midpoint of the list (positions 9 through 16) is now the 12th position with a value of 16. Since $19 > 16$, further search is restricted to the 13th position and above.

1 2 3 5 6 7 8 10 12 13 15 16 18 19 20 22

3. The midpoint of the current list is now the 14th position with a value of 19. Since $19 \neq 19$, further search is restricted to the portion from the 13th through the 14th positions .

1 2 3 5 6 7 8 10 12 13 15 16 18 19 20 22

4. The midpoint of the current list is now the 13th position with a value of 18. Since $19 > 18$, search is restricted to the portion from the 14th position through the 14th.

1 2 3 5 6 7 8 10 12 13 15 16 18 19 20 22

5. Now the list has a single element and the loop ends. Since $19=19$, the location 14 is returned.

Sorting

- To sort the elements of a list is to put them in increasing order (numerical order, alphabetic, and so on).
- Sorting is an important problem because:
 - A nontrivial percentage of all computing resources are devoted to sorting different kinds of lists, especially applications involving large databases of information that need to be presented in a particular order (e.g., by customer, part number etc.).
 - An amazing number of fundamentally different algorithms have been invented for sorting. Their relative advantages and disadvantages have been studied extensively.
 - Sorting algorithms are useful to illustrate basic concepts in CS
- A variety of sorting algorithms are studied in this book; binary, insertion, bubble, selection, merge, quick, and tournament.
- Section 3.3 shows time required to sort a list using sorting algorithms covered in this section.

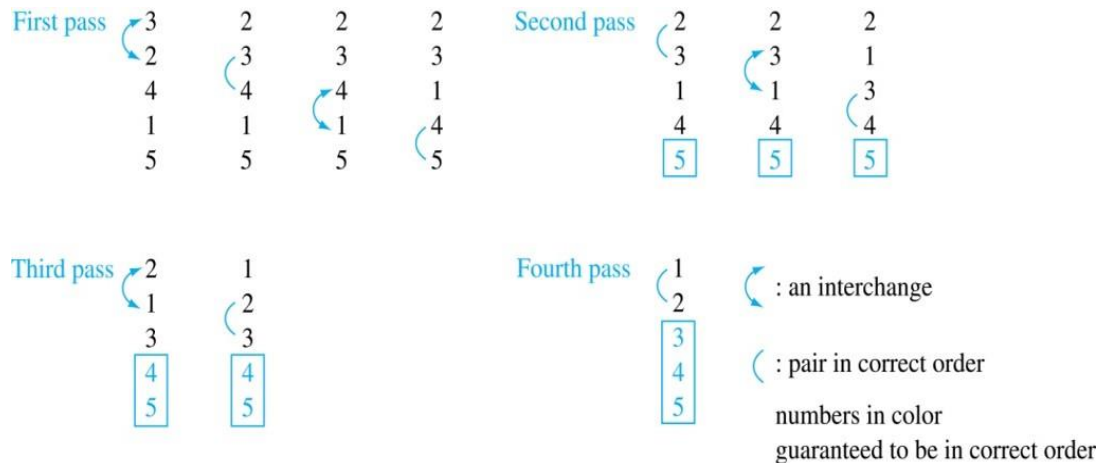
Bubble Sort

- Bubble sort makes multiple passes through a list. Every pair of elements that are found to be out of order are interchanged.

```
procedure bubblesort ( $a_1, \dots, a_n$ : real numbers with  $n \geq 2$ )  
  for  $i := 1$  to  $n - 1$   
    for  $j := 1$  to  $n - i$   
      if  $a_j > a_{j+1}$  then interchange  $a_j$  and  $a_{j+1}$   
  { $a_1, \dots, a_n$  is now in increasing order}
```

Bubble Sort

Example: Show the steps of bubble sort with 3 2 4 1 5



At the first pass the largest element has been put into the correct position

At the end of the second pass, the 2nd largest element has been put into the correct position.

In each subsequent pass, an additional element is put in the correct position.

[Jump to long description](#)

Insertion Sort

Insertion sort begins with the 2nd element. It compares the 2nd element with the 1st and puts it before the first if it is not larger.

Next the 3rd element is put into the correct position among the first 3 elements.

In each subsequent pass, the $n+1^{\text{st}}$ element is put into its correct position among the first $n+1$ elements.

Linear search is used to find the correct position.

```
procedure insertion sort ( $a_1, \dots, a_n$  : real  
                           numbers with  $n \geq 2$ )  
  
  for  $j := 2$  to  $n$   
     $i := 1$   
    while  $a_j > a_i$   
       $i := i + 1$   
     $m := a_j$   
    for  $k := 0$  to  $j - i - 1$   
       $a_{j-k} := a_{j-k-1}$   
     $a_i := m$   
  {Now  $a_1, \dots, a_n$  is in increasing order}
```

Insertion Sort

- **Example:** Show all the steps of insertion sort with the input: 3 2 4 1 5
 - i. 2 3 4 1 5 (*first two positions are interchanged*)
 - ii. 2 3 4 1 5 (*third element remains in its position*)
 - iii. 1 2 3 4 5 (*fourth is placed at beginning*)
 - iv. 1 2 3 4 5 (*fifth element remains in its position*)

Greedy Algorithms



- Optimization problems minimize or maximize some parameter over all possible inputs.
- Among the many optimization problems we will study are:
 - Finding a route between two cities with the smallest total mileage.
 - Determining how to encode messages using the fewest possible bits.
 - Finding the fiber links between network nodes using the least amount of fiber.
- Optimization problems can often be solved using a greedy algorithm, which makes the "best" choice at each step. Making the "best choice" at each step does not necessarily produce an optimal solution to the overall problem, but in many instances, it does.
- After specifying what the "best choice" at each step is, we try to prove that this approach always produces an optimal solution, or find a counterexample to show that it does not.
- The greedy approach to solving problems is an example of an algorithmic paradigm, which is a general approach for designing an algorithm. We return to algorithmic paradigms in Section 3.3.

Greedy Algorithms: Making Change

- **Example:** Design a greedy algorithm for making change (in U.S. money) of n cents with the following coins: quarters (25 cents), dimes (10 cents), nickels (5 cents), and pennies (1 cent), using the least total number of coins.
- **Idea:** At each step choose the coin with the largest possible value that does not exceed the amount of change left.
 1. If $n = 67$ cents, first choose a quarter leaving $67 - 25 = 42$ cents. Then choose another quarter leaving $42 - 25 = 17$ cents
 2. Then choose 1 dime, leaving $17 - 10 = 7$ cents.
 3. Choose 1 nickel, leaving $7 - 5 = 2$ cents.
 4. Choose a penny, leaving one cent. Choose another penny leaving 0 cents.



Greedy Change-Making Algorithm

Solution: Greedy change-making algorithm for n cents. The algorithm works with any coin denominations c_1, c_2, \dots, c_r .

```
procedure change( $c_1, c_2, \dots, c_r$  : values of coins, where  $c_1 > c_2 > \dots > c_r$ ;  
                   $n$  : a positive integer)
```

```
  for  $i := 1$  to  $r$ 
```

```
     $d_i := 0$  [ $d_i$  counts the coins of denomination  $c_i$ ]
```

```
    while  $n \geq c_i$ 
```

```
       $d_i := d_i + 1$  [add a coin of denomination  $c_i$ ]
```

```
       $n = n - c_i$ 
```

```
  [ $d_i$  counts the coins  $c_i$ ]
```

For the example of U.S. currency, we may have quarters, dimes, nickels and pennies: $c_1 = 25$, $c_2 = 10$, $c_3 = 5$, and $c_4 = 1$.

Proving Optimality for U.S. Coins

- Show that the change making algorithm for U.S. coins is optimal.
- **Lemma 1:** If n is a positive integer, then n cents in change using quarters, dimes, nickels, and pennies, using the fewest coins possible has at most 2 dimes, 1 nickel, 4 pennies, and cannot have 2 dimes and a nickel. The total amount of change in dimes, nickels, and pennies must not exceed 24 cents.
- **Proof:** By contradiction
 - If we had 3 dimes, we could replace them with a quarter and a nickel.
 - If we had 2 nickels, we could replace them with 1 dime.
 - If we had 5 pennies, we could replace them with a nickel.
 - If we had 2 dimes and 1 nickel, we could replace them with a quarter.
 - The allowable combinations, have a maximum value of 24 cents; 2 dimes and 4 pennies.

Proving Optimality for U.S. Coins

- **Theorem:** The greedy change-making algorithm for U.S. coins produces change using the fewest coins possible.
- **Proof:** By contradiction.
 1. Assume there is a positive integer n such that change can be made for n cents using quarters, dimes, nickels, and pennies, with a fewer total number of coins than given by the algorithm.
 2. Then, $q' \leq q$ where q' is the number of quarters used in this optimal way and q is the number of quarters in the greedy algorithm's solution. But this is not possible by Lemma 1, since the value of the coins other than quarters cannot be greater than 24 cents.
 3. Similarly, by Lemma 1, the two algorithms must have the same number of dimes, nickels, and quarters.

Greedy Change-Making Algorithm

- Optimality depends on the denominations available.
- For U.S. coins, optimality still holds if we add half dollar coins (50 cents) and dollar coins (100 cents).
- But if we allow only quarters (25 cents), dimes (10 cents), and pennies (1 cent), the algorithm no longer produces the minimum number of coins.
- Consider the example of 31 cents. The optimal number of coins is 4, i.e., 3 dimes and 1 penny. What does the algorithm output?

Greedy Scheduling

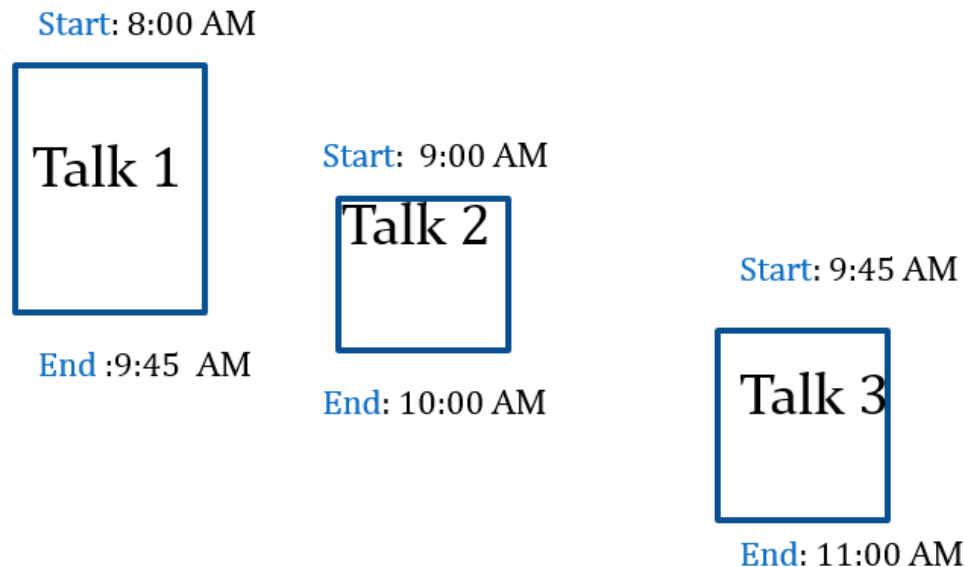
- **Example:** We have a group of proposed talks with start and end times. Construct a greedy algorithm to schedule as many as possible in a lecture hall, under the following assumptions:
 - When a talk starts, it continues till the end.
 - No two talks can occur at the same time.
 - A talk can begin at the same time that another ends.
 - Once we have selected some of the talks, we cannot add a talk which is incompatible with those already selected because it overlaps at least one of these previously selected talks.

How should we make the “best choice” at each step of the algorithm? That is, which talk do we pick?

- The talk that starts earliest among those compatible with chosen talks?
- The talk that is shortest among those already compatible?
- The talk that ends earliest among compatible talks?

Greedy Scheduling₂

Picking the shortest talk doesn't work.



Can you find a counterexample here?

But picking the one that ends soonest does work. The algorithm is specified on the next page.

Greedy Scheduling algorithm

Solution: At each step, choose the talks with the earliest ending time among the talks compatible with those selected.

```
procedure schedule( $s_1 \leq s_2 \leq \dots \leq s_n$  : start times,  $e_1 \leq e_2 \leq \dots \leq e_n$  : end times)

    sort talks by finish time and reorder so that  $e_1 \leq e_2 \leq \dots \leq e_n$ 
     $S := \emptyset$ 
    for  $j := 1$  to  $n$ 
        if talk  $j$  is compatible with  $S$  then
             $S := S \cup \{\text{talk } j\}$ 
    return  $S$  [  $S$  is the set of talks scheduled]
```

Will be proven correct by induction in Chapter 5.

Halting Problem

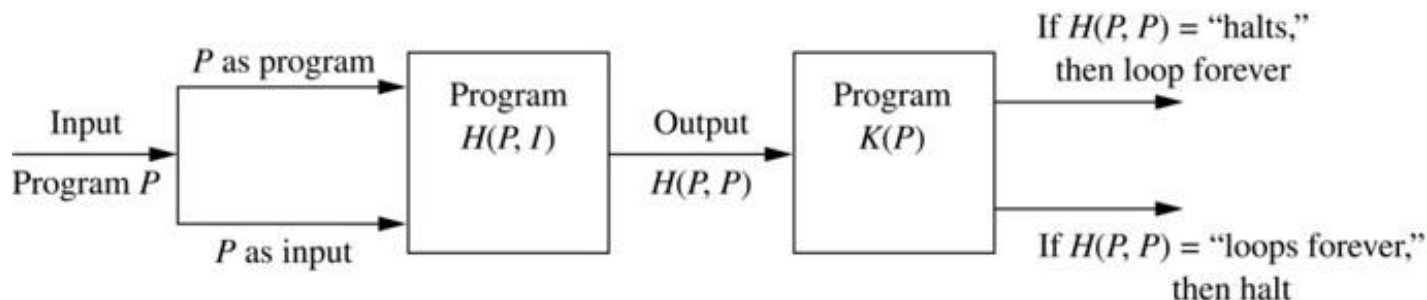
- **Example:** Can we develop a procedure that takes as input a computer program along with its input and determines whether the program will eventually halt with that input.
- **Solution:** Proof by contradiction.

Assume that there is such a procedure and call it $H(P,I)$. The procedure $H(P,I)$ takes as input a program P and the input I to P .

- H outputs "halt" if it is the case that P will stop when run with input I .
- Otherwise, H outputs "loops forever."

Halting Problem

- Since a program is a string of characters, we can call $H(P, P)$. Construct a procedure $K(P)$, which works as follows.
- If $H(P, P)$ outputs "loops forever" then $K(P)$ halts.
- If $H(P, P)$ outputs "halt" then $K(P)$ goes into an infinite loop printing "ha" on each iteration.



[Jump to long description](#)

Halting Problem

- Now we call K with K as input, i.e. $K(K)$.
- If the output of $H(K,K)$ is "loops forever" then $K(K)$ halts. **A Contradiction.**
- If the output of $H(K,K)$ is "halts" then $K(K)$ loops forever. **A Contradiction.**
- Therefore, there cannot be a procedure that can decide whether or not an arbitrary program halts. The halting problem is unsolvable.

The Growth of Functions

- Section 3.2

Section Summaries



Donald E. Knuth
(Born 1938)

Big-O Notation

Big-O Estimates for Important Functions

Big-Omega and Big-Theta Notation



Edmund Landau
(1877-1938)



Paul Gustav Heinrich Bachmann
(1837-1920)

The Growth of Functions

- In both computer science and in mathematics, there are many times when we care about how fast a function grows.
- In computer science, we want to understand how quickly an algorithm can solve a problem as the size of the input grows.
 - We can compare the efficiency of two different algorithms for solving the same problem.
 - We can also determine whether it is practical to use a particular algorithm as the input grows.
 - We'll study these questions in Section 3.3.
- Two of the areas of mathematics where questions about the growth of functions are studied are:
 - number theory (covered in Chapter 4)
 - combinatorics (covered in Chapters 6 and 8)

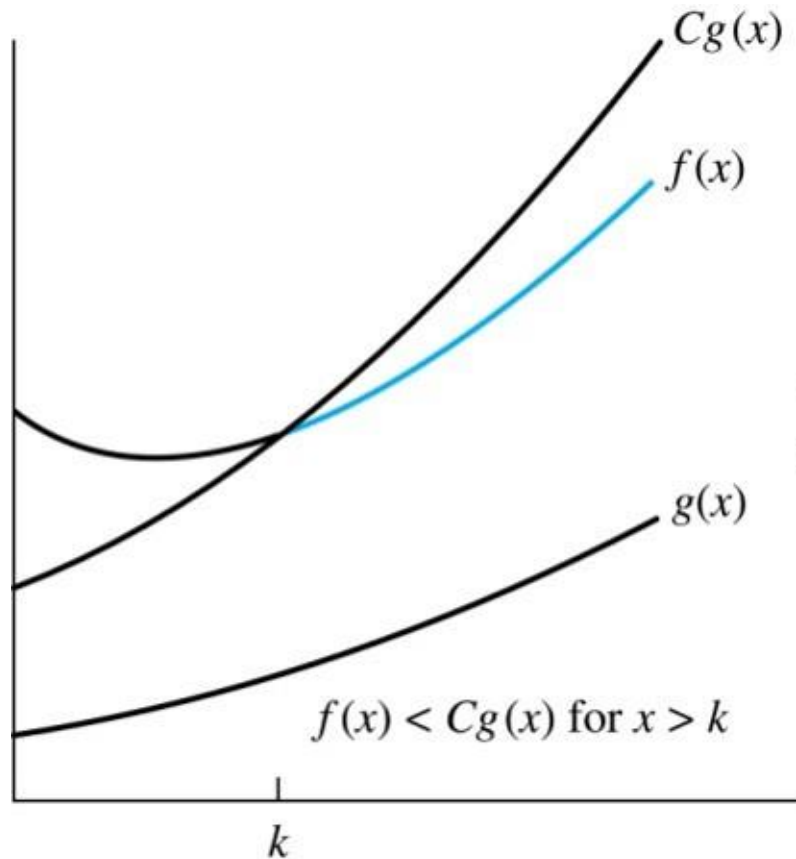
Big-O Notation

- **Definition:** Let f and g be functions from the set of integers or the set of real numbers to the set of real numbers. We say that $f(x)$ is $O(g(x))$ if there are constants C and k such that

$$|f(x)| \leq C|g(x)| \quad \text{when } x > k$$

- read as
 - $f(x)$ is big-O of $g(x)$
 - g asymptotically dominates f
- constants C and k are called witnesses or witness points to the relationship $f(x)$ is $O(g(x))$
 - only one pair of witnesses is needed

Illustration of Big-O Notation₁



$$f(x) \text{ is } O(g(x))$$

The part of the graph of $f(x)$ that satisfies $f(x) < Cg(x)$ is shown in color.

[Jump to long description](#)

Some Important Points about Big-O Notation

- if one pair of witnesses is found, then there are infinitely many pairs
- we can always make the k or the C larger and still maintain the inequality $|f(x)| \leq C|g(x)|$.
- sometimes this is written as $f(x) = O(g(x))$
 - may be confusing since an inequality relates f and g
 - better to write $f(x) \in O(g(x))$
- usually, we will drop the absolute value sign since we will always deal with functions of positive values

Using the Definition of Big-O Notation

- **Example:** Show that $f(x) = x^2 + 2x + 1$ is $O(x^2)$
- **Solution:** when $x > 1$, $x < x^2$ and $1 < x^2$, so

$$0 \leq x^2 + 2x + 1 \leq x^2 + 2x^2 + x^2 = 4x^2$$

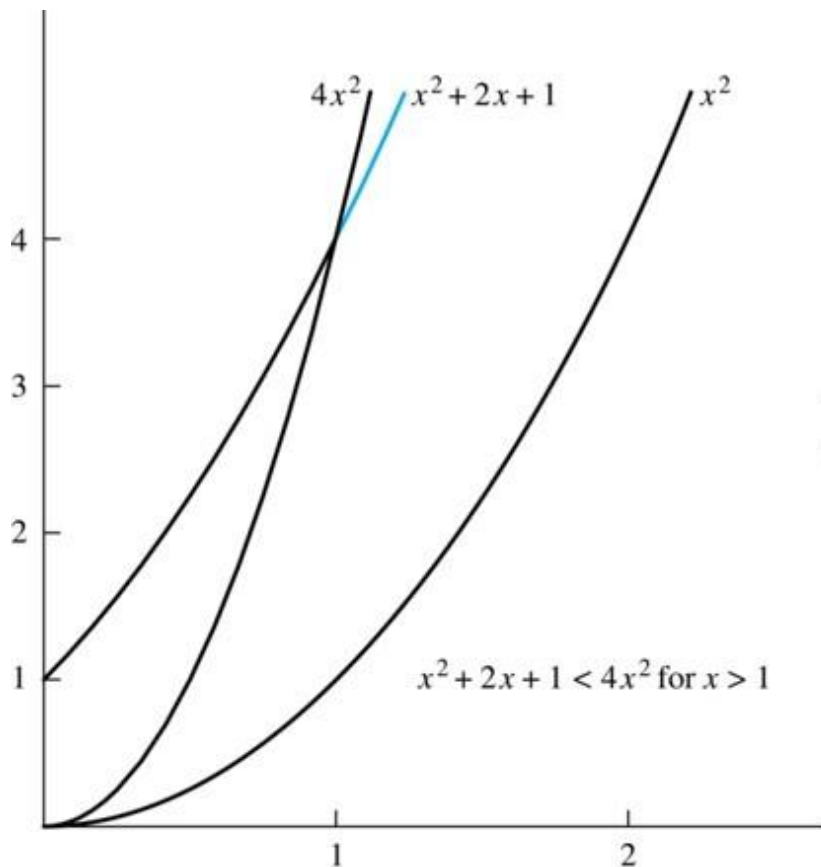
- take $C = 4$ and $k = 1$ as witnesses to show that $f(x)$ is $O(x^2)$
- alternatively, when $x > 2$, we have $2x \leq x^2$ and $1 < x^2$, so

$$0 \leq x^2 + 2x + 1 \leq x^2 + x^2 + x^2 = 3x^2 \text{ when } x > 2$$

- take $C = 3$ and $k = 2$ as witnesses instead

Illustration of Big-O Notation

$$f(x) = x^2 + 2x + 1 \text{ is } O(x^2)$$



The part of the graph of $f(x) = x^2 + 2x + 1$ that satisfies $f(x) < 4x^2$ is shown in blue.

[Jump to long description](#)

Big-O Notation

- if $f(x) = x^2 + 2x + 1$ and $g(x) = x^2$
 - $f(x)$ and $g(x)$ are of the same order
 - $f(x)$ is $O(g(x))$ and $g(x)$ is $O(f(x))$
- if $f(x)$ is $O(g(x))$ and $h(x)$ is larger than $g(x)$ for all positive real numbers, then $f(x)$ is $O(h(x))$
- in practice, the goal is to select the function $g(x)$ in $O(g(x))$ as small as possible (up to multiplication by a constant, of course)

Using the Definition of Big-O Notation

- **Example:** Show that $7x^2$ is $O(x^3)$.
- **Solution:** when $x > 7$, $7x^2 < x^3$, take $C = 1$ and $k = 7$ as witnesses to establish that $7x^2$ is $O(x^3)$
 - would $C = 7$ and $k = 1$ work?
- **Example:** Show that n^2 is not $O(n)$.
- **Solution:** suppose there are constants C and k for which $n^2 \leq Cn$, whenever $n > k$
 - by dividing both sides of $n^2 \leq Cn$ by n , $n \leq C$ must hold for all $n > k$: a contradiction!

Big-O Estimates for Polynomials

Example: Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$

where a_0, a_1, \dots, a_n are real numbers with $a_n \neq 0$. Then $f(x)$ is $O(x^n)$.

Proof: $|f(x)| = |a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0|$

Uses triangle inequality,
an exercise in Section 1.8.

Assuming $x > 1$

$$\begin{aligned} &\leq |a_n| x^n + |a_{n-1}| x^{n-1} + \cdots + |a_1| x + |a_0| \\ &= x^n \left(|a_n| + |a_{n-1}|/x + \cdots + |a_1|/x^{n-1} + |a_0|/x^n \right) \\ &\leq x^n \left(|a_n| + |a_{n-1}| + \cdots + |a_1| + |a_0| \right) \end{aligned}$$

Take $C = |a_n| + |a_{n-1}| + \cdots + |a_0|$ and $k = 1$. Then $f(x)$ is $O(x^n)$.

The leading term $a_n x^n$ of a polynomial dominates its growth.

Big-O Estimates for Some Important Functions

- **Example:** Use big-O notation to estimate the sum of the first n positive integers.
- **Solution:** $1 + 2 + \cdots + n \leq n + n + \cdots + n = n^2$

$1 + 2 + \cdots + n$ is $O(n^2)$ taking $C = 1$ and $k = 1$.

- **Example:** Use big-O notation to estimate the factorial function
- **Solution:**

$$f(n) = n! = 1 \times 2 \times \cdots \times n .$$

$$n! = 1 \times 2 \times \cdots \times n \leq n \times n \times \cdots \times n = n^n$$

$n!$ is $O(n^n)$ taking $C = 1$ and $k = 1$.

Big-O Estimates for Some Important Functions

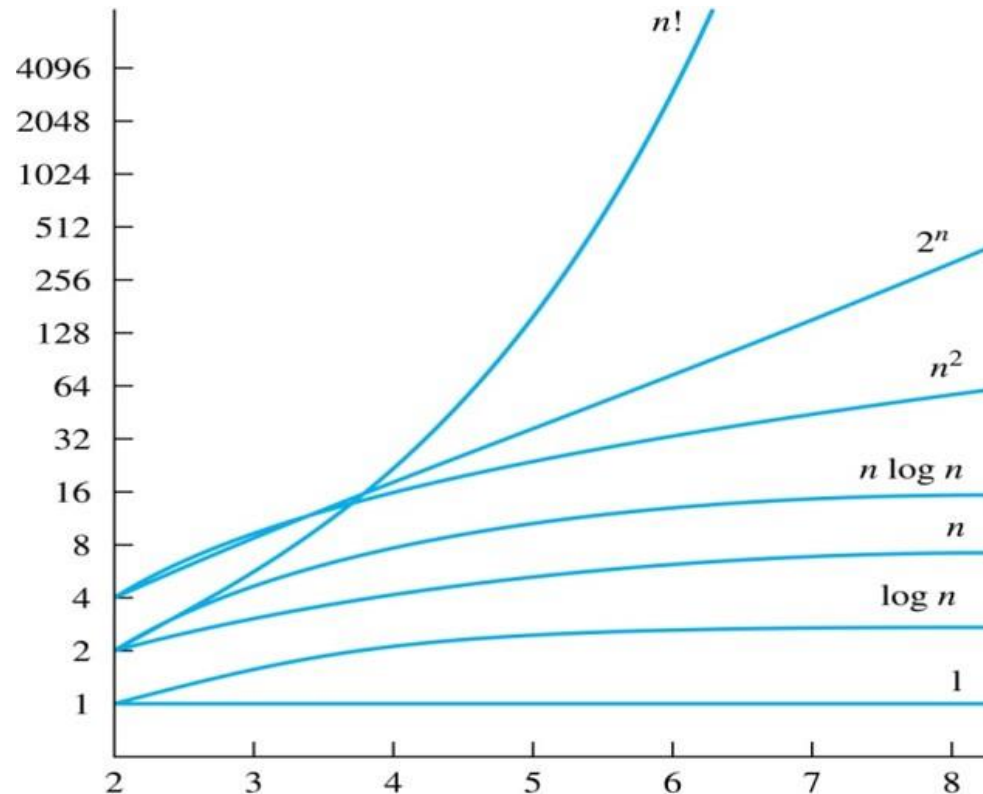
Example: Use big- O notation to estimate $\log n!$

Solution: Given that $n! \leq n^n$ (previous slide)

then $\log(n!) \leq n \cdot \log(n)$.

Hence, $\log(n!)$ is $O(n \cdot \log(n))$ taking $C = 1$ and $k = 1$.

Display of Growth of Functions



Note the difference in behavior of functions as n gets larger

[Jump to long description](#)

Useful Big-O Estimates Involving Logs, Powers, and Exponents

- if $d > c > 1$, then

n^c is $O(n^d)$, but n^d is not $O(n^c)$

- if $b > 1$ and c and d are positive, then

$(\log_b n)^c$ is $O(n^d)$, but n^d is not $O((\log_b n)^c)$

- if $b > 1$ and d is positive, then

n^d is $O(b^n)$, but b^n is not $O(n^d)$

- if $c > b > 1$, then

b^n is $O(c^n)$, but c^n is not $O(b^n)$

Combinations of Functions₁

- if $f_1(x)$ is $O(g_1(x))$ and $f_2(x)$ is $O(g_2(x))$ then
 $(f_1 + f_2)(x)$ is $O(\max(|g_1(x)|, |g_2(x)|))$
- if $f_1(x)$ and $f_2(x)$ are both $O(g(x))$ then
 $(f_1 + f_2)(x)$ is $O(g(x))$
- if $f_1(x)$ is $O(g_1(x))$ and $f_2(x)$ is $O(g_2(x))$ then
 $(f_1 \cdot f_2)(x)$ is $O(g_1(x) \cdot g_2(x))$

Ordering Functions by Order of Growth

- Put the functions below in order so that each function is big-O of the next function on the list.

$$f_1(n) = (1.5)^n$$

$$f_2(n) = 8n^3 + 17n^2 + 111$$

$$f_3(n) = (\log n)^2$$

$$f_4(n) = 2^n$$

$$f_5(n) = \log(\log n)$$

$$f_6(n) = n^2 (\log n)^3$$

$$f_7(n) = 2^n (n^2 + 1)$$

$$f_8(n) = n^3 + n(\log n)^2$$

$$f_9(n) = 10000$$

$$f_{10}(n) = n!$$

We solve this exercise by successively finding the function that grows slowest among all those left on the list.

$$f_9(n) = 10000 \quad (\text{constant, does not increase with } n)$$

$$f_5(n) = \log(\log n) \quad (\text{grows slowest of all the others})$$

$$f_3(n) = (\log n)^2 \quad (\text{grows next slowest})$$

$$f_6(n) = n^2 (\log n)^3 \quad (\text{next largest, } (\log n)^3 \text{ factor smaller than any power of } n)$$

$$f_2(n) = 8n^3 + 17n^2 + 111 \quad (\text{tied with the one below})$$

$$f_8(n) = n^3 + n(\log n)^2 \quad (\text{tied with the one above})$$

$$f_1(n) = (1.5)^n \quad (\text{next largest, an exponential function})$$

$$f_4(n) = 2^n \quad (\text{grows faster than one above since } 2 > 1.5)$$

$$f_7(n) = 2^n (n^2 + 1) \quad (\text{grows faster than above because of the } n^2 + 1 \text{ factor})$$

$$f_{10}(n) = n! \quad (n! \text{ grows faster than } cn^c \text{ for every } c)$$

Big-Omega Notation

Definition: Let f and g be functions from the set of integers or the set of real numbers to the set of real numbers. We say that $f(x)$ is $\Omega(g(x))$ if there are constants C and k such that

$$|f(x)| \geq C|g(x)| \quad \text{when } x > k.$$

Ω is the upper case version of the lower case Greek letter ω .

We say that " $f(x)$ is big-Omega of $g(x)$."

Big-O gives an upper bound on the growth of a function, while Big-Omega gives a lower bound. Big-Omega tells us that a function grows at least as fast as another.

$f(x)$ is $\Omega(g(x))$ if and only if $g(x)$ is $O(f(x))$. This follows from the definitions. See the text for details.

Big-Omega Notation

Example: Show that $f(x) = 8x^3 + 5x^2 + 7$ is $\Omega(g(x))$ where $g(x) = x^3$.

Solution: $f(x) = 8x^3 + 5x^2 + 7 \geq 8x^3$ for all positive real numbers x .

- Is it also the case that $g(x) = x^3$ is $O(8x^3 + 5x^2 + 7)$?

Big-Theta Notation

Definition: Let f and g be functions from the set of integers or the set of real numbers to the set of real numbers. The function

$f(x)$ is $\Theta(g(x))$ if $f(x)$ is $O(g(x))$ and $f(x)$ is $\Omega(g(x))$.

We say that " f is big-Theta of $g(x)$ " and also that " $f(x)$ is of order $g(x)$ " and also that " $f(x)$ and $g(x)$ are of the same order."

$f(x)$ is $\Theta(g(x))$ if and only if there exists constants C_1, C_2 and k such that $C_1g(x) < f(x) < C_2g(x)$ if $x > k$. This follows from the definitions of big-O and big-Omega.

Θ is the upper case version of the lower case Greek letter θ .

Big-Theta Notation

Example: Show that the sum of the first n positive integers is $\Theta(n^2)$.

Solution: Let $f(n) = 1 + 2 + \cdots + n$

- We have already shown that $f(n)$ is $O(n^2)$.
- To show that $f(n)$ is $\Omega(n^2)$, we need a positive constant C such that $f(n) > Cn^2$ for sufficiently large n . Summing only the terms greater than $n/2$ we obtain the inequality

$$\begin{aligned} 1 + 2 + \cdots + n &\geq \lceil n/2 \rceil + (\lceil n/2 \rceil + 1) + \cdots + n \\ &\geq \lceil n/2 \rceil + \lceil n/2 \rceil + \cdots + \lceil n/2 \rceil \\ &= (n - \lceil n/2 \rceil + 1) \lceil n/2 \rceil \\ &\geq (n/2)(n/2) = n^2 / 4 \end{aligned}$$

- Taking $C = \frac{1}{4}$, $f(n) > Cn^2$ for all positive integers n . Hence, $f(n)$ is $\Omega(n^2)$, and we can conclude that $f(n)$ is $\Theta(n^2)$.

Big-Theta Notation

- **Example:** Show that $f(x) = 3x^2 + 8x \log x$ is $\Theta(x^2)$.

- **Solution:**

$$3x^2 + 8x \log x \leq 11x^2 \text{ for } x > 1,$$

$$\text{since } 0 \leq 8x \log x \leq 8x^2.$$

$$\text{Hence, } 3x^2 + 8x \log x \text{ is } O(x^2).$$

$$x^2 \text{ is clearly } O(3x^2 + 8x \log x)$$

$$\text{Hence, } 3x^2 + 8x \log x \text{ is } \Theta(x^2).$$

Big-Theta Notation

When $f(x)$ is $\Theta(g(x))$ it must also be the case that $g(x)$ is $\Theta(f(x))$.

Note that $f(x)$ is $\Theta(g(x))$ if and only if it is the cases that $f(x)$ is $O(g(x))$ and $g(x)$ is $O(f(x))$.

Sometimes writers are careless and write as if big- O notation has the same meaning as big-Theta.

Big-Theta Estimates for Polynomials

Theorem: Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$
where a_0, a_1, \dots, a_n are real numbers with $a_n \neq 0$.

Then $f(x)$ is of order x^n (or $\Theta(x^n)$).

(The proof is an exercise.)

Example:

The polynomial $f(x) = 8x^5 + 5x^2 + 10$ is order of x^5 (or $\Theta(x^5)$).

The polynomial $f(x) = 8x^{199} + 7x^{100} + x^{99} + 5x^2 + 25$
is order of x^{199} (or $\Theta(x^{199})$).

Complexity of Algorithms

- Section 3.3

Section Summary

- Time Complexity
- Worst-Case Complexity
- Algorithmic Paradigms
- Understanding the Complexity of Algorithms

The Complexity of Algorithms₁

- Given an algorithm, how efficient is this algorithm for solving a problem given input of a particular size? To answer this question, we ask:
 - How much time does this algorithm use to solve a problem?
 - How much computer memory does this algorithm use to solve a problem?
- When we analyze the time the algorithm uses to solve the problem given input of a particular size, we are studying the time complexity of the algorithm.
- When we analyze the computer memory the algorithm uses to solve the problem given input of a particular size, we are studying the space complexity of the algorithm.

The Complexity of Algorithms₂

- In this course, we focus on time complexity. The space complexity of algorithms is studied in later courses.
- We will measure time complexity in terms of the number of operations an algorithm uses and we will use big-O and big-Theta notation to estimate the time complexity.
- We can use this analysis to see whether it is practical to use this algorithm to solve problems with input of a particular size. We can also compare the efficiency of different algorithms for solving the same problem.
- We ignore implementation details (including the data structures used and both the hardware and software platforms) because it is extremely complicated to consider them.

Time Complexity

- To analyze the time complexity of algorithms, we determine the number of operations, such as comparisons and arithmetic operations (addition, multiplication, etc.). We can estimate the time a computer may actually use to solve a problem using the amount of time required to do basic operations.
- We ignore minor details, such as the "house keeping" aspects of the algorithm.
- We will focus on the worst-case time complexity of an algorithm. This provides an upper bound on the number of operations an algorithm uses to solve a problem with input of a particular size.
- It is usually much more difficult to determine the average case time complexity of an algorithm. This is the average number of operations an algorithm uses to solve a problem over all inputs of a particular size.

Complexity Analysis of Algorithms

Example: Describe the time complexity of the algorithm for finding the maximum element in a finite sequence.

```
procedure max( $a_1, a_2, \dots, a_n$ : integers)
  max :=  $a_1$ 
  for  $i := 2$  to  $n$ 
    if  $\text{max} < a_i$  then  $\text{max} := a_i$ 
  return max{max is the largest element}
```

Solution: Count the number of comparisons.

- The $\text{max} < a_i$ comparison is made $n - 1$ times.
- Each time i is incremented, a test is made to see if $i \leq n$.
- One last comparison determines that $i > n$.
- Exactly $2(n - 1) + 1 = 2n - 1$ comparisons are made.

Hence, the time complexity of the algorithm is $\Theta(n)$.

Worst-Case Complexity of Linear Search

Example: Determine the time complexity of linear search algorithm.

```
procedure linear search ( $x$  : integer,  $a_1, a_2, \dots, a_n$  : distinct integers)
   $i := 1$ 
  while ( $i \leq n$  and  $x \neq a_i$ )
     $i := i + 1$ 
  if  $i \leq n$  then  $location := i$ 
  else  $location := 0$ 
  return  $location$  { $location$  is the subscript of the term that equals  $x$ , or is 0 if  $x$  is not found}
```

Solution: Count the number of comparisons.

- At each step two comparisons are made; $i \leq n$ and $x \neq a_i$.
- To end the loop, one comparison $i \leq n$ is made.
- After the loop, one more $i \leq n$ comparison is made.

If $x = a_i$, $2i + 1$ comparisons are used. If x is not on the list, $2n + 1$ comparisons are made and then an additional comparison is used to exit the loop. So, in the worst case $2n + 2$ comparisons are made.

Hence, the complexity is $\Theta(n)$.

Average-Case Complexity of Linear Search

- **Example:** Describe the average case performance of the linear search algorithm. (Although usually it is very difficult to determine average-case complexity, it is easy for linear search.)
- **Solution:** Assume the element is in the list and that the possible positions are equally likely. By the argument on the previous slide, if $x = a_i$, the number of comparisons is $2i + 1$.

$$\frac{3 + 5 + 7 + \dots + (2n + 1)}{n} = \frac{2(1 + 2 + 3 \dots + n) + n}{n} = \frac{2 \left[\frac{n(n+1)}{2} \right] + n}{n} = n + 2$$

Hence, the average-case complexity of linear search is $\Theta(n)$.

Worst-Case Complexity of Binary Search

Example: Describe the time complexity of binary search in terms of the number of comparisons used.

```
procedure binary search( $x$ : integer,  $a_1, a_2, \dots, a_n$ : increasing integers)
   $i := 1$  { $i$  is the left endpoint of interval}
   $j := n$  { $j$  is right endpoint of interval}
  while  $i < j$ 
     $m := \lfloor (i + j)/2 \rfloor$ 
    if  $x > a_m$  then  $i := m + 1$ 
    else  $j := m$ 
  if  $x = a_i$  then  $location := i$ 
  else  $location := 0$ 
  return  $location$  {location is the subscript  $i$  of the term  $a_i$  equal to  $x$ , or 0 if  $x$  is not found}
```

Solution: Assume (for simplicity) $n = 2^k$ elements. Note that $k = \log n$.

- Two comparisons are made at each stage; $i < j$, and $x > a_m$.
- At the first iteration the size of the list is 2^k and after the first iteration it is 2^{k-1} . Then 2^{k-2} and so on until the size of the list is $2^1 = 2$.
- At the last step, a comparison tells us that the size of the list is the size is $2^0 = 1$ and the element is compared with the single remaining element.
- Hence, at most $2k + 2 = 2 \log n + 2$ comparisons are made.
- Therefore, the time complexity is $\Theta(\log n)$, better than linear search.

Worst-Case Complexity of Bubble Sort

Example: What is the worst-case complexity of bubble sort in terms of the number of comparisons made?

```
procedure bubblesort ( $a_1, \dots, a_n$ : real numbers with  $n \geq 2$ )  
  for  $i := 1$  to  $n - 1$   
    for  $j := 1$  to  $n - i$   
      if  $a_j > a_{j+1}$  then interchange  $a_j$  and  $a_{j+1}$   
    { $a_1, \dots, a_n$  is now in increasing order}
```

Solution: A sequence of $n-1$ passes is made through the list. On each pass $n - i$ comparisons are made.

$$(n-1) + (n-2) + \dots + 2 + 1 = \frac{n(n-1)}{2}$$

The worst-case complexity of bubble sort is $\Theta(n^2)$ since

$$\frac{n(n-1)}{2} = \frac{1}{2}n^2 - \frac{1}{2}n.$$

Worst-Case Complexity of Insertion Sort

Example: What is the worst-case complexity of insertion sort in terms of the number of comparisons made?

Solution: The total number of comparisons are

$$2 + 3 + \dots + n = \frac{n(n-1)}{2} - 1$$

Therefore the complexity is $\Theta(n^2)$.

```
procedure insertion sort ( $a_1, \dots, a_n$  :  
    real numbers with  $n \geq 2$ )  
    for  $j := 2$  to  $n$   
         $i := 1$   
        while  $a_j > a_i$   
             $i := i + 1$   
         $m := a_j$   
        for  $k := 0$  to  $j - i - 1$   
             $a_{j-k} := a_{j-k-1}$   
         $a_i := m$ 
```

Matrix Multiplication Algorithm

- The definition for matrix multiplication can be expressed as an algorithm; $\mathbf{C} = \mathbf{A} \mathbf{B}$ where \mathbf{C} is an $m \times n$ matrix that is the product of the $m \times k$ matrix \mathbf{A} and the $k \times n$ matrix \mathbf{B} .
- This algorithm carries out matrix multiplication based on its definition.

```
procedure matrix multiplication(A,B: matrices)
  for  $i := 1$  to  $m$ 
    for  $j := 1$  to  $n$ 
       $c_{ij} := 0$ 
      for  $q := 1$  to  $k$ 
         $c_{ij} := c_{ij} + a_{iq} b_{qj}$ 
  return  $\mathbf{C}$  { $\mathbf{C} = [c_{ij}]$  is the product of  $\mathbf{A}$  and  $\mathbf{B}$ }
```

$\mathbf{A} = [a_{ij}]$ is a $m \times k$ matrix

$\mathbf{B} = [b_{ij}]$ is a $k \times n$ matrix

Complexity of Matrix Multiplication

- **Example:** How many additions of integers and multiplications of integers are used by the matrix multiplication algorithm to multiply two $n \times n$ matrices.
- **Solution:** There are n^2 entries in the product. Finding each entry requires n multiplications and $n - 1$ additions. Hence, n^3 multiplications and $n^2(n - 1)$ additions are used.
- Hence, the complexity of matrix multiplication is $O(n^3)$.

Boolean Product Algorithm

- The definition of Boolean product of zero-one matrices can also be converted to an algorithm.

```
procedure Boolean product (A, B: zero-one matrices)
  for i := 1 to m
    for j := 1 to n
       $c_{ij} := 0$ 
      for q := 1 to k
         $c_{ij} := c_{ij} \vee (a_{iq} \wedge b_{qj})$ 
      return C {C = [cij] is the Boolean product of A and B}
```

Complexity of Boolean Product Algorithm

- **Example:** How many bit operations are used to find $A \odot B$, where A and B are $n \times n$ zero-one matrices?
- **Solution:** There are n^2 entries in the $A \odot B$. A total of n ORs and n ANDs are used to find each entry. Hence, each entry takes $2n$ bit operations. A total of $2n^3$ operations are used.
- Therefore the complexity is $O(n^3)$

Matrix-Chain Multiplication

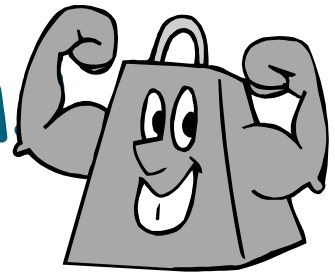
- How should the *matrix-chain* $A_1 A_2 \cdots A_n$ be computed using the fewest multiplications of integers, where A_1, A_2, \dots, A_n are $m_1 m_2, m_2 m_3, \dots, m_n m_{n+1}$ integer matrices. Matrix multiplication is associative (exercise in Section 2.6).
- **Example:** In which order should the integer matrices $A_1 A_2 A_3$ - where A_1 is 30×20 , A_2 20×40 , A_3 40×10 - be multiplied to use the least number of multiplications.
- **Solution:** There are two possible ways to compute $A_1 A_2 A_3$.
 - $A_1(A_2 A_3)$: $A_2 A_3$ takes $20 \cdot 40 \cdot 10 = 8000$ multiplications. Then multiplying A_1 by the 20×10 matrix $A_2 A_3$ takes $30 \cdot 20 \cdot 10 = 6000$ multiplications. So the total number is $8000 + 6000 = 14,000$.
 - $(A_1 A_2) A_3$: $A_1 A_2$ takes $30 \cdot 20 \cdot 40 = 24,000$ multiplications. Then multiplying the 30×40 matrix $A_1 A_2$ by A_3 takes $30 \cdot 40 \cdot 10 = 12,000$ multiplications. So the total number is $24,000 + 12,000 = 36,000$.
- So the first method is best.

An efficient algorithm for finding the best order for matrix-chain multiplication can be based on the algorithmic paradigm known as *dynamic programming*. (see Ex. 57 in Sec 8.1)

Algorithmic Paradigms

- An algorithmic paradigm is a general approach based on a particular concept for constructing algorithms to solve a variety of problems.
- Greedy algorithms were introduced in Section 3.1.
- We discuss brute-force algorithms in this section.
- We will see divide-and-conquer algorithms (Chapter 8), dynamic programming (Chapter 8), backtracking (Chapter 11), and probabilistic algorithms (Chapter 7). There are many other paradigms that you may see in later courses.

Brute-Force Algorithm



- A brute-force algorithm is solved in the most straightforward manner, without taking advantage of any ideas that can make the algorithm more efficient.
- Brute-force algorithms we have previously seen are sequential search, bubble sort, and insertion sort.

Computing the Closest Pair of Points by Brute-Force

Example: Construct a brute-force algorithm for finding the closest pair of points in a set of n points in the plane and provide a worst-case estimate of the number of arithmetic operations.

Solution: Recall that the distance between (x_i, y_i) and (x_j, y_j) is $\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$. A brute-force algorithm

simply computes the distance between all pairs of points and picks the pair with the smallest distance.

Note: There is no need to compute the square root, since the square of the distance between two points is smallest when the distance is smallest.

Computing the Closest Pair of Points by Brute-Force

Algorithm for finding the closest pair in a set of n points.

```
procedure closest pair  $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n): x_i, y_i \text{ real numbers})$ 
```

```
   $min = \infty$ 
```

```
  for  $i := 2$  to  $n$ 
```

```
    for  $j := 1$  to  $i-1$ 
```

```
      if  $(x_j - x_i)^2 + (y_j - y_i)^2 < min$ 
```

```
        then  $min := (x_j - x_i)^2 + (y_j - y_i)^2$ 
```

```
           $closest\ pair := (x_i, y_i), (x_j, y_j)$ 
```

```
  return closest pair
```

The algorithm loops through $n(n-1)/2$ pairs of points, computes the value

$(x_j - x_i)^2 + (y_j - y_i)^2$ and compares it with the minimum, etc. So, the algorithm uses $\Theta(n^2)$ arithmetic and comparison operations.

We will develop an algorithm with $O(n \log n)$ worst-case complexity in Section 8.3.

Understanding the Complexity of Algorithms

- **TABLE 1** Commonly Used Terminology for the Complexity of Algorithms.

<i>Complexity</i>	<i>Terminology</i>
$\Theta(1)$	Constant complexity
$\Theta(\log n)$	Logarithmic complexity
$\Theta(n)$	Linear complexity
$\Theta(n \log n)$	Linearithmic complexity
$\Theta(n^b)$	Polynomial complexity
$\Theta(b^n)$, where $b > 1$	Exponential complexity
$\Theta(n!)$	Factorial complexity

Understanding the Complexity of Algorithms

• **TABLE 2** The Computer Time Used by Algorithms.

<i>Problem Size</i>	<i>Bit Operations Used</i>					
<i>n</i>	$\log n$	<i>n</i>	$n \log n$	n^2	2^n	$n!$
10	3×10^{-11} s	10^{-10} s	3×10^{-10} s	10^{-9} s	10^{-8} s	3×10^{-7} s
10^2	7×10^{-11} s	10^{-9} s	7×10^{-9} s	10^{-7} s	4×10^{11} yr	*
10^3	1.0×10^{-10} s	10^{-8} s	1×10^{-7} s	10^{-5} s	*	*
10^4	1.3×10^{-10} s	10^{-7} s	1×10^{-6} s	10^{-3} s	*	*
10^5	1.7×10^{-10} s	10^{-6} s	2×10^{-5} s	0.1 s	*	*
10^6	2×10^{-10} s	10^{-5} s	2×10^{-4} s	0.17 min	*	*

Times of more than 10^{100} years are indicated with an *.

Complexity of Problems

- Tractable Problem: There exists a polynomial time algorithm to solve this problem. These problems are said to belong to the Class P.
- Intractable Problem: There does not exist a polynomial time algorithm to solve this problem
- Unsolvable Problem: No algorithm exists to solve this problem, e.g., halting problem.
- Class NP: Solution can be checked in polynomial time. But no polynomial time algorithm has been found for finding a solution to problems in this class.
- NP Complete Class: If you find a polynomial time algorithm for one member of the class, it can be used to solve all the problems in the class.

P Versus NP Problem



Stephen
Cook
(Born
1939)

The P versus NP problem asks whether the class $P = NP$? Are there problems whose solutions can be checked in polynomial time, but can not be solved in polynomial time?

- Note that saying no one has yet found a polynomial time algorithm is different from showing that the problem cannot be solved by a polynomial time algorithm.

If a polynomial time algorithm for any of the problems in the NP complete class were found, then that algorithm could be used to obtain a polynomial time algorithm for every problem in the NP complete class.

- Satisfiability (in Section 1.3) is an NP complete problem.

It is generally believed that $P \neq NP$ since no one has been able to find a polynomial time algorithm for any of the problems in the NP complete class.

The problem of P versus NP remains one of the most famous unsolved problems in mathematics (including theoretical computer science). The Clay Mathematics Institute has offered a prize of \$1,000,000 for a solution.

Appendix of Image Long Descriptions

Bubble Sort - Appendix

- There are four passes of using a bubble sort for a column of numbers. There are 5 columns in the first pass. In the first column, the numbers from the top to the bottom are 3, 2, 4, 1, and 5. Numbers 3 and 2 are interchanged. In the second column, the numbers are 2, 3, 4, 1, and 5. Numbers 3 and 4 are in the correct order. In the third column, numbers 4 and 1 are interchanged. In the fourth column, the numbers are 2, 3, 1, 4, and 5. Numbers 4 and 5 are in the correct order. There are 3 columns in the second pass. In the first column, the numbers are 2, 3, 1, 4, and 5. Numbers 2 and 3 are in the correct order, number 5 is guaranteed to be in the correct order. In the second column, numbers 3 and 1 are interchanged, number 5 is guaranteed to be in the correct order. In the third column, the numbers are 2, 1, 3, 4, and 5. Numbers 3 and 4 are in the correct order, and number 5 is guaranteed to be in the correct order. There are 2 columns in the third pass. In the first column, the numbers are 2, 1, 3, 4, and 5. Numbers 2 and 1 are interchanged, numbers 4 and 5 are guaranteed to be in the correct order. In the second column, the numbers are 1, 2, 3, 4, and 5. Numbers 2 and 3 are in the correct order, numbers 4 and 5 are guaranteed to be in the correct order. There is one column in the fourth pass. The numbers are 1, 2, 3, 4, and 5. Numbers 1 and 2 are in the correct order, numbers 3, 4 and 5 are guaranteed to be in the correct order.s

[Jump to the image](#)

Halting Problem - Appendix

- There is an input Program P . It divides into 2 branches. P as a program and P as an input. They both go to a box named Program H left parenthesis P , I right parenthesis. An output of this box is H left parenthesis P , P right parenthesis. It goes to a box called Program K left parenthesis P right parenthesis. The box has two outputs. The first output is if H left parenthesis P , P right parenthesis equals "halts", then loop forever. The second output is if H left parenthesis P , P right parenthesis equals "loops forever" then halt.

[Jump to the image](#)

Illustration of Big-O Notation - Appendix

- The first curve is given by the equation x squared. The second curve is given by the equation x squared plus two x plus one. The third curve is given by the equation 4 times x squared. The first and the second curves intersect at the point 1, 4. After the intersection point, the second curve is between the first and the third curves and highlighted.

[Jump to the image](#)

Illustration of Big-O Notation - Appendix

- The first curve is $Cg(x)$. The second curve is $f(x)$. The third curve is $g(x)$. The first and the second curves intersect at $x = k$. After the intersection point, the second curve is between the first and the third curves and highlighted.

[Jump to the image](#)

Display of Growth of Functions - Appendix

- The vertical scale is logarithmic, and it ranges from 0 to 4096. From the top to the bottom the curves are: n factorial starting from $y = 2$, two raised to the n power starting from $y = 4$, n squared starting from $y = 4$, n logarithm n starting from $y = 2$, n starting from $y = 2$, logarithm n starting from $y = 1$, one starting from $y = 1$.

[Jump to the image](#)