Course – Analysis and Design of Algorithms

Course Instructor

Dr. Umadevi V

Department of CSE, BMSCE

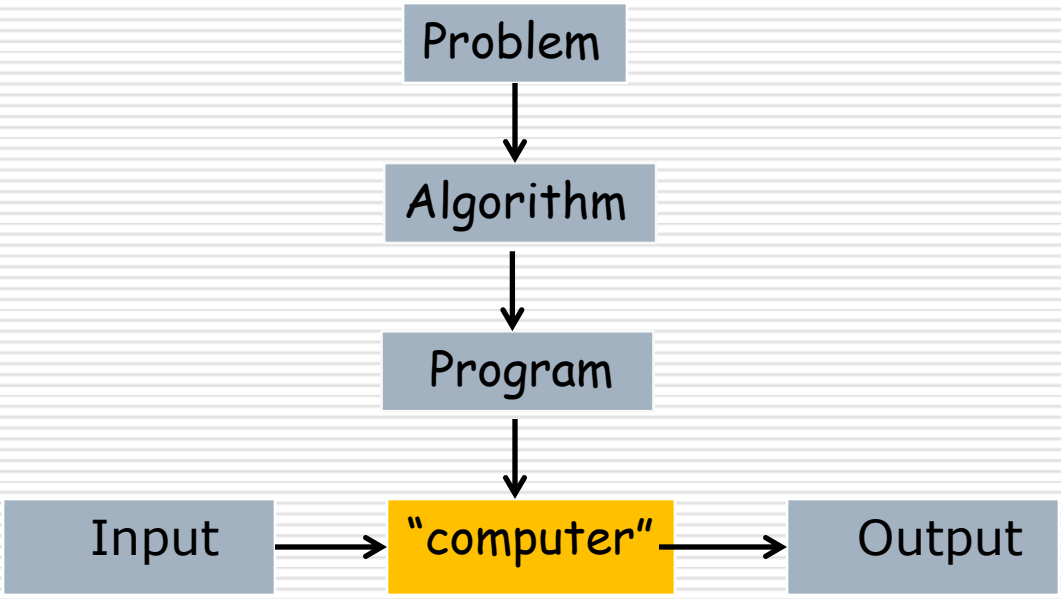
Webpage:https://sites.google.com/site/drvumadevi/

Unit 1: Introduction to Algorithms

- Fundamentals of Algorithmic Problem Solving
- Space and Time Complexity
- Order of Growth
- Asymptotic Notations

Algorithm

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



Examples of Algorithms

Computing **Greatest Common Divisor** of Two non-negative, not-both zero Integers

- gcd(m, n): the largest integer that divides both *m* and *n*
- First try - Euclid's Algorithm:
  - Idea: gcd(m, n) = gcd(n, m mod n)

Greatest Common Divisor (Euclid's Algorithm), **gcd(m, n)**

- **Step 1**: If n = 0, return value of m as the answer and stop; otherwise, proceed to Step 2.
- **Step 2**: Divide m by n and assign the value of the remainder to r.
- **Step 3:** Assign the value of n to m and the value of r to n. Go to Step 1.

Pseudocode for (Euclid's Algorithm), **gcd(m, n)**

ALGORITHM Euclid(m, n)

// Computes gcd(m, n) by Euclid's algorithm

// Input: Two nonnegative, not-both-zero integers m and n

//Output: Greatest common divisor of m and n

while n ≠ 0 **do**

    r = m **mod** n

    m = n

    n = r

return m

Question:
GCD(36,48) how many division Operations are required to compute GCD using Euclid algorithm ?

Second try: Middle-school procedure, **gcd(m, n)**

- **Step 1**: Find prime factors of m.
- **Step 2**: Find prime factors of n.
- **Step 3**: Identify all common prime factors of m and n
- **Step 4**: Compute product of all common factors and return product as the answer.

Question:
GCD(36,48) how many division Operations are required to compute GCD using Middle-School procedure ?

Third try: Consecutive Integer Checking, **gcd(m, n)**

- **Step 1:** Assign the value of min{m, n} to q.
- **Step 2:** Divide m by q. If the remainder is 0, go to Step 3; otherwise, go to Step 4.
- **Step 3:** Divide n by q. If the remainder is 0, return the value of q as the answer and stop; otherwise, proceed to Step 4.
- **Step 4:** Decrease the value of q by 1. Go to Step 2.

Question:
Try computing GCD(36,48) using Consecutive Integer Checking method ?

What can we learn from the three examples of gcd(m, n) ?

- Each step must be basic and unambiguous
- Same algorithm, but different representations (different pseudocodes)
- Same problem, but different algorithms, based on different ideas and having dramatically different speeds.
  - gcd(31415, 14142) = 1; Euclid takes ~0.08 ms whereas Consecutive Integer Checking takes ~0.55 ms, about 7 times speedier

Method 1: Greatest Common Divisor (Euclid's Algorithm), **gcd(m, n)**

- **Step 1**: If n = 0, return value of m as the answer and stop; otherwise, proceed to Step 2.
- **Step 2**: Divide m by n and assign the value of the remainder to r.
- **Step 3:** Assign the value of n to m and the value of r to n. Go to Step 1.

Method 2: Consecutive Integer Checking, **gcd(m, n)**

- **Step 1:** Assign the value of min{m, n} to q.
- **Step 2:** Divide m by q. If the remainder is 0, go to Step 3; otherwise, go to Step 4.
- **Step 3:** Divide n by q. If the remainder is 0, return the value of q as the answer and stop; otherwise, proceed to Step 4.
- **Step 4:** Decrease the value of q by 1. Go to Step 2.

C++ Program - Analysis of the methods to find the GCD of two numbers

```
#include<iostream.h>
#include<conio.h>
#include<time.h>

long int euclid(long int m,long int n)
{
        clock_t start,end;
        start=clock();
        long int r;
        while(n!=0)
        {
                r=m%n;
                m=n;
                n=r;
        }
        end=clock();
        cout<<endl<<"Time taken:"<<(end-start)/CLK_TCK<<" sec";
        return m;
}
```

C++ Program - Analysis of the methods to find the GCD of two numbers

```
long int con(long int m,long int n)
{
        clock_t start,end;
        start=clock();
        long int t,r,g;
        if(m>n)
        {  t=n; }
        else
        {   t=m; }

      a:do
        {
                r=m%t;
                if(r!=0)
                t--;
        } while(r!=0);
        if(r==0)
        {
                r=n%t;
                if(r==0)
                g=t;
                else
                {
                t--;
                goto a;
                }
        }
        end=clock();
        cout<<"Time taken :"<<(end-start)/CLK_TCK<<" sec";
        return g;
} /*End of the function con*/
```

C++ Program - Analysis of the methods to find the GCD of two numbers

```
void main()
{
        long int x,y;
        clrscr();

        cout<<"\t\t\tANALYSIS OF THE TWO ALGORITHMS"<<endl<<endl;
        cout<<"GCD - EUCLID'S ALG : "<<endl;
        cout<<"enter two numbers:";
        cin>>x>>y;
        cout<<endl<<endl<<"GCD : "<<euclid(x,y);
        cout<<endl<<endl<<"-----------------------------------------------";
        cout<<endl<<endl<<"GCD - CONSECUTIVE INTEGER CHECKING ALG :
         "<<endl<<endl;
        cout<<endl<<endl<<"GCD : "<<con(x,y);
        getch();
}
```

ANALYSIS OF THE TWO LGORITHMS

GCD - EUCLID'S ALG :
enter two numbers:7896543 345678

Time taken: 0.08 millisecond
GCD : 3

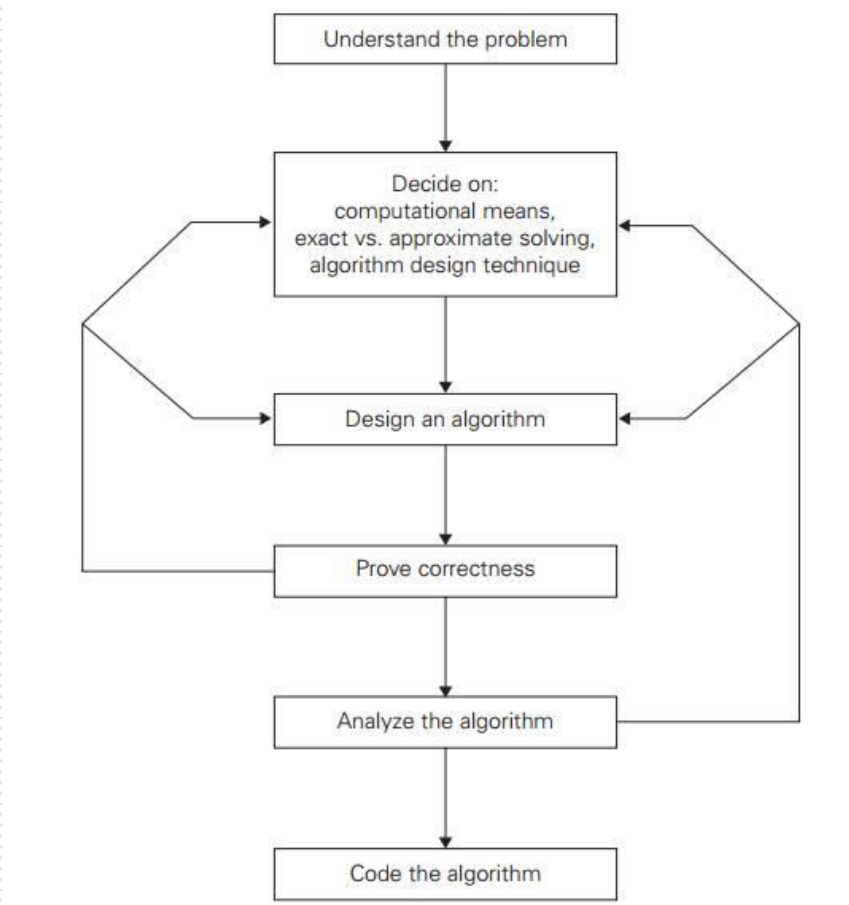
GCD - CONSECUTIVE INTEGER CHECKING ALG :

Time taken :0.55 millisecond
GCD : 3

INFERENCE:
The euclid's method takes less time than the consecutive integer checking method and hence euclid's method is better.

Fundamentals of Algorithmic Problem Solving

- Sequence of steps in the process of design and analysis of algorithms



Question

With the help of a flow chart, explain the various steps of algorithm design and analysis process.

Fundamentals of Algorithmic Problem Solving (Contd….)

- Understanding the problem
  - Ask questions, do a few small examples by hand, think about special cases, etc.
  - An input is an instance of the problem the algorithm solves
  - Specify exactly the set of instances the algorithm needs to handle
  - Example: gcd(m, n)

Fundamentals of Algorithmic Problem Solving (Contd….)

- Decide on
  - Exact vs. approximate solution
    - Approximate algorithm: Cannot solve exactly, e.g., extracting square roots, solving nonlinear equations, etc.
  - Appropriate Data Structure

Fundamentals of Algorithmic Problem Solving (Contd….)

- Design algorithm
- Prove correctness of the algorithm
  - Yields required output for every legitimate input in finite time
  - E.g., Euclid's: gcd(m, n) = gcd(n, m mod n)
    - Second integer gets smaller on every iteration, because (m mod n) can be 0, 1, …, n-1 thus less than n
    - The algorithm terminates when the second integer is 0

Fundamentals of Algorithmic Problem Solving (Contd….)

- Analyze algorithm
  - **Time efficiency:** How fast it runs
  - **Space efficiency:** How much extra memory it uses
  - **Simplicity:** Easier to understand, usually contains fewer bugs, sometimes simpler is more efficient, but not always!
  - **Generality:** Example, whether two integers are relatively prime, use gcd(m, n)

Fundamentals of Algorithmic Problem Solving (Contd….)

- Coding algorithm
  - Write in a programming language for a real machine
  - Standard tricks:
    - Compute loop invariant (which does not change value in the loop) outside loop
    - Replace expensive operation by cheap ones

Discussion: Algorithms in your Life

What algorithms do you use in every day life? Do you think you could write a program to make them more efficient?

What algorithms do you think are used by your favorite Games and Apps?

Have you ever made an algorithm for a program? What did it do? Was it correct and efficient?

Test Your Analytical Skill

Problem 1

There are n lockers in a hallway numbered sequentially from 1 to n. Initially, all the locker doors are closed. You make n passes by the lockers, each time starting with locker #1. On the ith pass, i = 1, 2, ..., n, you toggle the door of every ith locker: if the door is closed, you open it, if it is open, you close it. For example, after the first pass every door is open; on the second pass you only toggle the even-numbered lockers (#2, #4, ...) so that after the second pass the even doors are closed and the odd ones are opened; the third time through you close the door of locker #3 (opened from the first pass), open the door of locker #6 (closed from the second pass), and so on. After the last pass, which locker doors are open and which are closed? How many of them are open ?

Easy but inefficient

Algorithm 1 Locker(A[0..$n-1$])

```
for i ← 0 to n − 1 do
    A[i] ← 0
end for
for i ← 1 to n do
    j ← i − 1
    while j < n do
        A[j] ← 1 − A[j]
        j ← j + i
    end while
end for
openCount ← 0
for i ← 0 to n − 1 do
    if A[i] = 1 then
        print (i + 1)-th locker door is open
        openCount ← openCount + 1
    else
        print (i + 1)-th locker door is closed
    end if
end for
print openCount locker doors are open
```

Better

Since all the doors are initially closed, a door will be open after the last pass if and only if it is toggled an odd number of times. Door i $(1 \leq i \leq n)$ is toggled on pass j $(1 \leq j \leq n)$ if and only if j divides i. Hence, the total number of times door i is toggled is equal to the number of its divisors. Note that if j divides i, i.e. $i = jk$, then k divides i too. Hence all the divisors of i can be paired (e.g., for $i = 12$, such pairs are 1 and 12, 2 and 6, 3 and 4) unless i is a perfect square (e.g., for $i = 16$, 4 does not have another divisor to be matched with). This implies that i has an odd number of divisors if and only if it is a perfect square, i.e., $i = j^2$. Hence doors that are in the positions that are perfect squares and only such doors will be open after the last pass. The total number of such positions not exceeding n is equal to $\lfloor \sqrt{n} \rfloor$: these numbers are the squares of the positive integers between 1 and $\lfloor \sqrt{n} \rfloor$ inclusively.

Better

Algorithm 2 LockerBetter($A[0..n-1]$)

```
for i ← 0 to n − 1 do
    A[i] ← 0
end for
c ← ⌊√n⌋
for i ← 0 to c − 1 do
    A[i · i] ← 1
end for
for i ← 0 to n − 1 do
    if A[i] = 1 then
        print (i + 1)-th locker door is open
    else
        print (i + 1)-th locker door is closed
    end if
end for
print c locker doors are open
```

Problem 1

- Design an algorithm for swapping two 3 digit non-zero integers n, m. Besides using arithmetic operations, your algorithm should not use any temporary variable.
- Solution

1) use 2 variables (say) a & b
2) ask user to input values
3) read values
4) execute the Exclusive-Or (XOR) operation like this to swap a and b:

```
a ^= b;
b ^= a;
a ^= b;
```

5) display the swapped values

Similarly you can try with Addition-Subtraction or Multiplication-Divition

Analysis of Algorithms

Space Complexity
Time Complexity

Reasons to Analyze Algorithms

- Predict Performance
- Compare Algorithms
- Provide Guarantees
- Understand theoretical basis.

- Primary Practical Reason: Avoid Performance Bugs

client gets poor performance because programmer did not understand performance characteristics

Performance measure of the algorithm

Two kinds of efficiency:
Space Efficiency or Space Complexity
Time Efficiency or Time Complexity

Two kinds of Algorithm Efficiency

- Analyzing the efficiency of an algorithm (or the complexity of an algorithm) means establishing the amount of computing resources needed to execute the algorithm. There are two types of resources:

• Memory space. It means the amount of space used to store all data processed by the algorithm.

• Running time. It means the time needed to execute all the operations specified in the algorithm.

Space efficiency: Deals with the space required by the algorithm

Time efficiency: It indicates how fast an algorithm runs.

What is Space complexity?

For any algorithm, memory is required for the following purposes...

- Memory required to store program instructions
- Memory required to store constant values
- Memory required to store variable values

Space complexity of an algorithm can be defined as follows...

Total amount of computer memory required by an algorithm to complete its execution is called as space complexity of that algorithm

What is Space complexity?

For any algorithm, memory is required for the following purposes...

- Memory required to store program instructions
- Memory required to store constant values
- Memory required to store variable values

Space complexity of an algorithm can be defined as follows...

Total amount of computer memory required by an algorithm to complete its execution is called as space complexity of that algorithm

Generally, when a program is under execution it uses the computer memory for THREE reasons. They are as follows...

- **Instruction Space:** It is the amount of memory used to store compiled version of instructions.
- **Data Space:** It is the amount of memory used to store all the variables and constants.
- **Environmental Stack:** It is the amount of memory used to store information of partially executed functions at the time of function call.

Space Complexity

Instruction space + Data space + Stack space

Calculating Space Complexity

To calculate the space complexity, we must know the memory required to store different datatype values (according to the compiler). For example, the C Programming Language compiler requires the following...

- 1 byte to store **Character** value,
- 2 bytes to store **Integer** value,
- 4 bytes to store **Floating** Point value,
- 6 or 8 bytes to store **double** value

Calculating Space Complexity

Example, Calculating the Data Space required for the following given code

```
int square(int a)
{
return a*a;
}
```

Calculating Space Complexity

Example, Calculating the Data Space required for the following given code

```
int square(int a)
{
return a*a;
}
```

	Data Space Required
For int a	2 Bytes
For returning a*a	2 Bytes
Total	**4 Bytes**

Calculating Space Complexity

Example:

```
int square(int a)
{
return a*a;
}
```

	Data Space Required
For int a	2 Bytes
For returning a*a	2 Bytes
Total	**4 Bytes**

Data Space Required:

- This code requires 2 bytes of memory to store variable '**a**' and another 2 bytes of memory is used for **return value**.
- That means, totally it requires **4 bytes of memory** to complete its execution. And this 4 bytes of memory is fixed for any input value of 'a'. This space complexity is said to be *Constant Space Complexity*.

- If any **algorithm requires a fixed amount of space** for all input values then that space complexity is said to be **Constant Space Complexity**

Calculating Space Complexity

Example, Calculating the Data Space required for the following given code

```
int sum(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)
sum = sum + A[i];
return sum;
}
```

Calculating Space Complexity

Example, Calculating the Data Space required for the following given code

```
int sum(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)
sum = sum + A[i];
return sum;
}
```

	Data Space Required
For parameter int A[]	n *2 Bytes
For parameter n	2 Bytes
For local variable sum	2 Bytes
For local variable i	2 Bytes
Total	**2n+6 Bytes**

Calculating Space Complexity

```
int sum(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)
sum = sum + A[i];
return sum; }
```

Data Space Required:

- '**n*2**' bytes of memory to store array variable '**a[]**'
  2 bytes of memory for integer parameter '**n**'
  4 bytes of memory for local integer variables '**sum**' and '**i**' (2 bytes each)

- That means, totally it requires **'2n+6' bytes of memory** to complete its execution. Here, the amount of memory depends on the input value of 'n'. This space complexity is said to be *Linear Space Complexity*.

If the **amount of space required by an algorithm is increased with the increase of input value**, then that space complexity is said to be **Linear Space Complexity**

Test your Knowledge

- Find Data Space required for the following code:

```
int sum(int x, int y, int z) {
 int r = x + y + z;
return r;
 }
```

Is the Space Complexity of this code is "**Constant Space Complexity**" or "**Linear Space Complexity" ?**

Test your Knowledge

- Find Data Space required for the following code:

```
void matrixAdd(int a[], int b[], int c[], int n) {
 for (int i = 0; i < n; ++i) {
 c[i] = a[i] + b[j]
 }
 }
```

Is the Space Complexity of this code is
"**Constant Space Complexity"**
or "**Linear Space Complexity" ?**

Performance measure of the algorithm

Two kinds of efficiency:

Space Efficiency or Space Complexity

Time Efficiency or Time Complexity

What is Time complexity?

- Every algorithm requires some amount of computer time to execute its instruction to perform the task. This computer time required is called time complexity.

Time complexity of an algorithm can be defined as follows...

- **The time complexity of an algorithm is the total amount of time required by an algorithm to complete its execution.**

What is Time complexity?

- Every algorithm requires some amount of computer time to execute its instruction to perform the task. This computer time required is called time complexity.

Time complexity of an algorithm can be defined as follows...

- **The time complexity of an algorithm is the total amount of time required by an algorithm to complete its execution.**

Generally, running time of an algorithm depends upon the following...

- Whether it is running on **Single** processor machine or **Multi** processor machine.
- Whether it is a **32 bit** machine or **64 bit** machine
- **Read** and **Write** speed of the machine.
- The time it takes to perform **Arithmetic** operations, **logical** operations, **return** value and **assignment** operations etc.,
- **Input** data

Calculating Time Complexity

- When we calculate time complexity of an algorithm, we consider only input data and ignore the remaining things, as they are machine dependent.

Example, Calculating the Time Complexity required for the following given code

```
int sum(int a, int b) {
return a+b;
}
```

	Time Required	
To **calculate** a+b	1 Unit of time	
For **returning** a+b	1 Unit of time	
Total	**2 Units of time**	

Calculating Time Complexity

- When we calculate time complexity of an algorithm, we consider only input data and ignore the remaining things, as they are machine dependent.

Example, Calculating the Time Complexity required for the following given code

```
int sum(int a, int b) {
return a+b;
}
```

	Time Required	*
To **calculate** a+b	1 Unit of time	1 Secs
For **returning** a+b	1 Unit of time	1 Secs
Total	**2 Units of time**	**2 Secs**

*Hypothetical approximation of time

Calculating Time Complexity

- When we calculate time complexity of an algorithm, we consider only input data and ignore the remaining things, as they are machine dependent.

Example:

```
int sum(int a, int b) {
return a+b; }
```

This Code requires 1 unit of time to calculate a+b and 1 unit of time to return the value. That means, totally it takes 2 units of time to complete its execution. And it does not change based on the input values of a and b. That means for all input values, it requires same amount of time i.e. 2 units.

If any program requires fixed amount of time for all input values then its time complexity is said to be **Constant Time Complexity.**

Calculating Time Complexity

Example, Calculate Time complexity for the following given code:

```
int fun(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)
        sum = sum + A[i];
return sum;
}
```

Calculating Time Complexity

```
int sumOfList(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)  sum = sum + A[i];
 return sum;  }
```

For the above code, time complexity can be calculated as follows...

	Cost or Number Operations in the Statement		
int sum = 0, i;	1 (sum=0 initializing sum with zero)		
for(i = 0; i < n; i++)			
sum = sum + A[i];			
return sum;			

Calculating Time Complexity

```
int sumOfList(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)  sum = sum + A[i];
 return sum;  }
```

For the above code, time complexity can be calculated as follows...

	Cost or Number Operations in the Statement		
int sum = 0, i;	1 (initializing zero to sum)		
for(i = 0; i < n; i++)	1 + 1 + 1 (i=0, i<n, i++)		
sum = sum + A[i];			
return sum;			

Calculating Time Complexity

```
int sumOfList(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)  sum = sum + A[i];
 return sum;  }
```

For the above code, time complexity can be calculated as follows...

	Cost or Number Operations in the Statement		
int sum = 0, i;	1 (initializing zero to sum)		
for(i = 0; i < n; i++)	1+1+1 (i=0, i<n, i++)		
sum = sum + A[i];	1+ 1 (Addition and Assigning result to sum)		
return sum;	1 (returning sum)		

Calculating Time Complexity

```
int sumOfList(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)  sum = sum + A[i];
 return sum;  }
```

For the above code, time complexity can be calculated as follows...

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	
int sum = 0, i;	1	**1**	
for(i = 0; i < n; i++)	1+1+1		
sum = sum + A[i];	1+ 1		
return sum;	1		

Calculating Time Complexity

```
int sumOfList(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)  sum = sum + A[i];
 return sum;  }
```

For the above code, time complexity can be calculated as follows...

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	
int sum = 0, i;	1	**1**	
for(i = 0; i < n; i++)	1+1+1	**1+(n+1)+n** (i=0 gets executed one time, i<n gets executed (n+1) times, i++ gets executed n times	
sum = sum + A[i];	1+ 1		
return sum;	1		
			Total

Calculating Time Complexity

```
int sumOfList(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)  sum = sum + A[i];
 return sum;  }
```

For the above code, time complexity can be calculated as follows...

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	
int sum = 0, i;	1	1	
for(i = 0; i < n; i++)	1+1+1	1+(n+1)+n	
sum = sum + A[i];	1+ 1	n + n	
return sum;	1	1	

Calculating Time Complexity

```
int sumOfList(int A[], int n) {
int sum = 0, i;
for(i = 0; i < n; i++)  sum = sum + A[i];
 return sum;  }
```

For the above code, time complexity can be calculated as follows...

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**
int sum = 0, i;	1	1	1
for(i = 0; i < n; i++)	1+1+1	1+(n+1)+n	2n+2
sum = sum + A[i];	1+ 1	n + n	2n
return sum;	1	1	1
		Running Time T(n)	4n+4

Calculating Time Complexity (Contd….

- For the calculation done in previous slide
  Cost is the amount of computer time required for a single operation in each line.
  Repetition is the amount of computer time required by each operation for all its repetitions.
  Total is the amount of computer time required by each operation to execute.

  So above code requires **'4n+4' Units** of computer time to complete the task. Here the exact time is not fixed. And it changes based on the **n** value. If we increase the **n** value then the time required also increases linearly.

  Totally it takes '4n+4' units of time to complete its execution and it is *Linear Time Complexity*.

- If the amount of time required by an algorithm is increased with the increase of input value then that time complexity is said to be **Linear Time Complexity**

Test Your Knowledge

Find Time Complexity of
the given Algorithm

```
sum(n)
1:    S ← 0
2:    i ← 1
3:    while i <= n do
4:        S ← S + i
5:        i ← i + 1
6:    endwhile
7: return S
```

Test Your Knowledge

Find Time Complexity of the given Algorithm

```
sum(n)
1:    S ← 0
2:    i ← 1
3:    while i <= n do
4:        S ← S + i
5:        i ← i + 1
6:    endwhile
7: return S
```

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**
1	1	1	1
2	1	1	1
3	1	n+1	n+1
4	1+1	n+n	2n
5	1+1	n+n	2n
7	1	1	1
		Running Time T(n)	5n+4

Test Your Knowledge

- Find Time Complexity of the following Algorithm

```
product(a[1..m, 1..n], b[1..n, 1..p])
1:    for i = 1, m do
2:        for j = 1, p do
3:            c[i, j] ← 0
4:            for k = 1, n do
5:                c[i, j] ← c[i, j] + a[i, k] * b[k, j]
6:            endfor
7:        endfor
8:    endfor
9: return c[1..m, 1..p]
```

```
int product(int a[m][n], int b[n][p]){
for(i=1;i<=m;i++){

for(j=1;j<=p;j++){
    c[i][j]=0;

     for(k=1;k<=n;k++){
     c[i][j]=c[i][j]+a[i][k]*b[k][j]

     }

    }

   }
return c
}
```

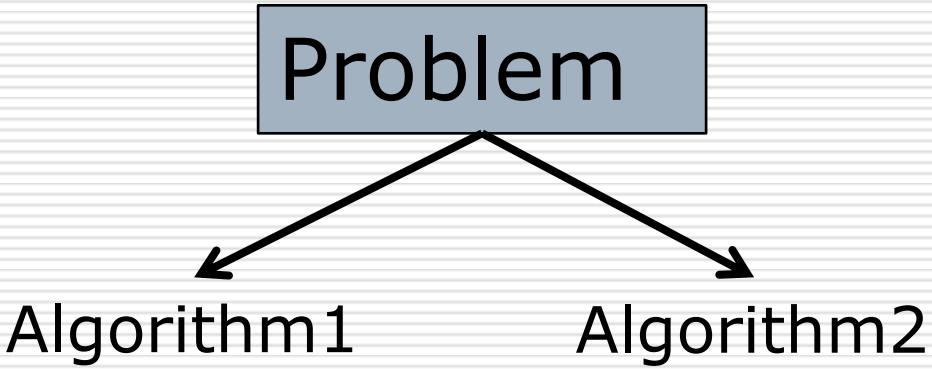
Test Your Knowledge

Find Time Complexity for the given Algorithm

```
product(a[1..m, 1..n], b[1..n, 1..p])
1:    for i = 1, m do
2:        for j = 1, p do
3:            c[i, j] ← 0
4:            for k = 1, n do
5:                c[i, j] ← c[i, j] + a[i, k] * b[k, j]
6:            endfor
7:        endfor
8:    endfor
9: return c[1..m, 1..p]
```

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**	
1	1+1+1	1+(m+1)+m	2m+2	2m+2
2	1+1+1	(1+(p+1)+p)m	(2p+2)m	2pm+2m
3	1	(p)m	pm	pm
4	1+1+1	((1+(n+1)+n)p)m	((2n+2)p)m	2npm+2pm
5	1+1+1	((n+n+n)p)m	((3n)p)m	3npm
9	1	1	1	1
			Running Time T(n)	5npm+5pm+4m+3

Given two algorithms for a task, how do we find out which one is better?



One naive way of doing this is – implement both the algorithms and run the two programs on your computer for different inputs and see which one takes less time. There are many problems with this approach for analysis of algorithms.

1) It might be possible that for some inputs, first algorithm performs better than the second. And for some inputs second performs better.

2) It might also be possible that for some inputs, first algorithm perform better on one machine and the second works better on other machine for some other inputs.

Question

What is the meaning of the notation **T(n)** w.r.t analysis of algorithms ?

Answer

- $T_A(\mathbf{n})$ = Maximum time taken (or Number of Machine operations needed) by the algorithm A to solve **input of size n**.
- Input size refers to number of values in the data set. Example: Say ten lakh Aadhar card numbers has to be sorted then **input size n** refers to 10,00,000
- $T_A(\mathbf{n})$ is the measure of **Goodness** of Algorithm A

T(n)

Expression we get for T(n) may not be of great consequence for real Computers /Computations because it various from one machine architecture to another machine architecture

Example: Program 1

```
#include <stdio.h>
main(){
int n,temp;

scanf(“%d”,&n)

temp=10*30;
}
```

Example: Program 1

```
#include <stdio.h>
main(){
int n,temp;

scanf("%d",&n)

temp=10*30;
}
```

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**
temp=10*30;	1+1	**1+1**	**2**
		Running Time T(n)	**2**

Example: Program 1

```
#include <stdio.h>
main(){
int n,temp;

scanf("%d",&n)

temp=10*30;
}
```

Running Time T(n)=2

Constant Time

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**
temp=10*30;	1+1	**1+1**	**2**
		Running Time T(n)	**2**

Example: Program 2

```
#include <stdio.h>
main(){
int i, n,temp;
scanf("%d",&n)
for(i=0; i < n; i++) {
                    temp=10*30;
                     }
}
```

Example: Program 2

```
#include <stdio.h>
main(){
int i, n,temp;
scanf("%d",&n)
for(i=0; i < n; i++) {
                    temp=10*30;
                     }
}
```

	Cost or Number of Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**
for(i=0; i < n; i++)	1+1+1	**1+(n+1)+n**	**2n+2**
temp=10*30;	1+1	**n+n**	**2n**
		Running Time T(n)	**4n+2**

T(n) for different values of n

```
#include <stdio.h>
main(){
int i, n,temp;
scanf("%d",&n)
for(i=0; i < n; i++) {
                    temp=10*30;
                     }   }
```

Running Time T(n)=4n+2

n	T(n)=4n+2	On Computer 1 Time taken(in Secs)	On Computer 2 Time taken (in Secs)
10	42	??	??
20	82	??	??
1000	4002	??	??
50000	20002	??	??

T(n) for different values of n

```
#include <stdio.h>
#include <time.h>
main(){
long int n,i; int temp;
clock_t start, end;

scanf("%ld",&n)
start=clock();
for(i=0; i < n; i++) {
                    temp=10*30;
                     }
end=clock();
printf("Time take %f in Secs",(((double)(end-start))/CLOCKS_PER_SEC));
}
```

Running Time T(n)=4n+2

T(n) for different values of n

```
#include <stdio.h>
main(){
int i, n,temp;
scanf("%d",&n)
for(i=0; i < n; i++) {
                    temp=10*30;
                     }   }
```

n	T(n)=4n+2	On Computer 1 Time taken(in Secs)	On Computer 2 Time taken (in Secs)
10	42	0.000002	
20	82	0.000003	
1000	4002	0.000010	
50000	200002	0.000395	

T(n) for different values of n

```
#include <stdio.h>
main(){
int i, n,temp;
scanf("%d",&n)
for(i=0; i < n; i++) {
                    temp=10*30;
                     }   }
```

n	T(n)=4n+2	On Computer 1 Time taken(in Secs)	On Computer 2 Time taken (in Secs)
10	42	**0.000002**	**0.000002**
20	82	**0.000003**	**0.000002**
1000	4002	**0.000010**	**0.000006**
50000	200002	**0.000395**	**0.000252**

T(n) for different values of n

```
#include <stdio.h>
main(){
int i, n,temp;
scanf("%d",&n)
for(i=0; i < n; i++) {
                    temp=10*30;
                     }  }
```

Running Time T(n)=4n+2

Linear Time

n	T(n)=4n+2	On Computer 1 Time taken(in Secs)	On Computer 2 Time taken (in Secs)
10	42	0.000002	0.000002
20	82	0.000003	0.000002
1000	4002	0.000010	0.000006
50000	200002	0.000395	0.000252

Example: Program 3

```
#include <stdio.h>
main(){  int i,j,n,temp;
scanf("%d", &n)
for(i=0; i < n; i++) {
        for(j=0; j < n;j++) {
                temp=10*30;
                 } } }
```

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**
for(i=0; i < n; i++)	1+1+1	**1+(n+1)+n**	**2n+2**
for(j=0; j < n; j++)	1+1+1	**(1+(n+1)+n)n**	**$2n^2+2n$**
temp=10*30;	1+1	**(n+n)n**	**$2n^2$**
		Running Time T(n)	**$4n^{2+}4n+2$**

T(n) for different values of n

```
#include <stdio.h>
main(){  int i,j,n,temp;
scanf("%d", &n)
for(i=0; i < n; i++) {
        for(j=0; j < n;j++) {
                temp=10*30;
                } } }
```

Running Time $T(n)=4n^{2+}4n+2$

Quadratic Time

n	$T(n)=4n^2+4n+2$	On Computer 1 Time taken(in Secs)	On Computer 2 Time taken (in Secs)
10	442	0.000003	0.000002
20	1682	0.000010	0.000007
1000	4004002	0.008255	0.005062
50000	10000200002	5.766243	4.765878

Rate of Growth or Order of Growth

Order of growth in algorithm means how the time for computation increases when you increase the input size. It really matters when your input size is very large.

Order of growth provide only a crude description of the behavior of a process.

Algorithms analysis is all about understanding growth rates. That is as the amount of data gets bigger, how much more resource will my algorithm require? Typically, we describe the resource growth rate of a piece of code in terms of a function.

Order of Growth: Linear vs Quadratic

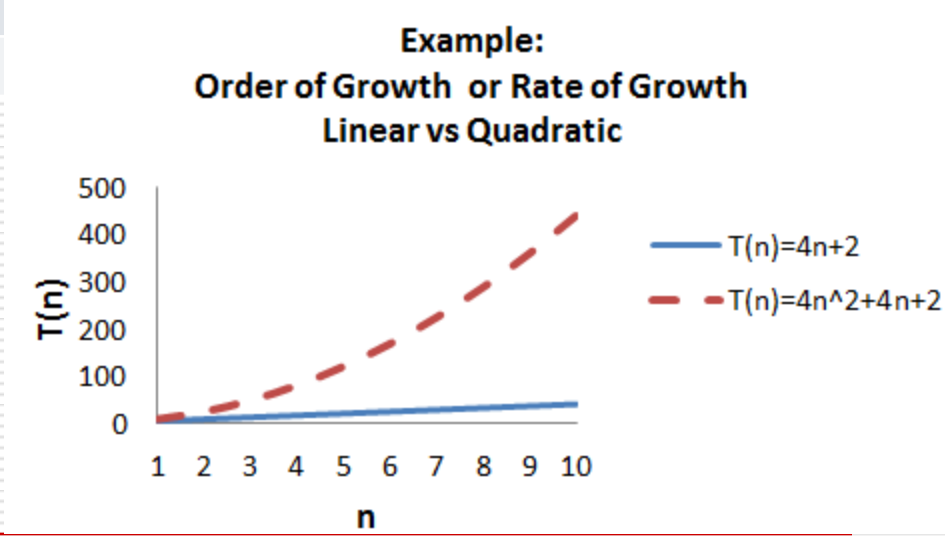
n	T(n)=4n+2	T(n)=4n^2+4n+2
1	6	10
2	10	26
3	14	50
4	18	82
5	22	122
6	26	170
7	30	226
8	34	290
9	38	362
10	42	442

Example

Order of Growth: Linear vs Quadratic

n	T(n)=4n+2	T(n)=4n^2+4n+2
1	6	10
2	10	26
3	14	50
4	18	82
5	22	122
6	26	170
7	30	226
8	34	290
9	38	362
10	42	442

Example



Example: Program 4

```
#include <stdio.h>
main(){  int i,j,k, n,temp;
scanf("%d", &n)
for(i=0; i < n; i++) {
        for(j=0; j < n;j++) {
                for(k=0; k < n;k++) {
                        temp=10*30;
                } } } }
```

Example: Program 4

```
#include <stdio.h>
main(){  int i,j,k, n,temp;
scanf("%d", &n)
for(i=0; i < n; i++) {
        for(j=0; j < n;j++) {
                for(k=0; k < n;k++) {
                        temp=10*30;
                } } } }
```

Running Time T(n)=$4n^3+4n^2+4n+2$

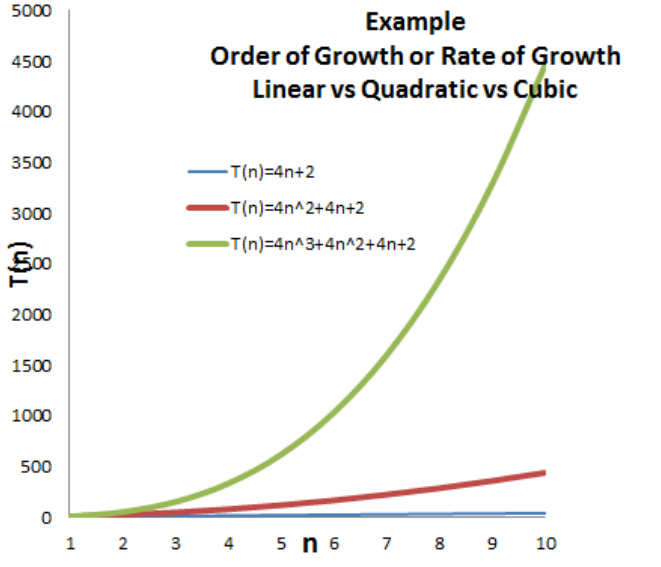
Cubic Time

	Cost or Number Operations in the Statement	**Repetitions or No. of Times of Execution**	**Total**
for(i=0; i < n; i++)	1+1+1	1+(n+1)+n	$2n+2$
for(j=0; j < n; j++)	1+1+1	(1+(n+1)+n)n	$2n^2+2n$
for(k=0; k < n;k++)	1+1+1	((1+(n+1)+n)n)n	$2n^3+2n^2$
temp=10*30;	1+1	((n+n)n)n	$2n^3$
		Running Time T(n)	$4n^3+4n^2+4n+2$

Order of Growth (or Rate of Growth): Linear vs Quadratic vs Cubic

n	T(n)=4n+2	T(n)=4n^2+4n+2	T(n)=4n^3+4n^2+4n+2
1	6	10	14
2	10	26	58
3	14	50	158
4	18	82	338
5	22	122	622
6	26	170	1034
7	30	226	1598
8	34	290	2338
9	38	362	3278
10	42	442	4442

Example



Interpretation of T(n)

- What is important is "**form (or shape)** of T(n)" i.e.., whether T(n) is Linear, Quadratic, Cubic..etc.
- Using the expression of T(n) we may not be able to give exact estimate but we can interpret the **behavior of the algorithm** when implemented on any computer.
- Analyzing the behavior of the algorithm for **LARGE n** is important.(i.e., as n tends to infinity n -> ∞)

Question

Consider, you are given with 10 Aadhaar card numbers and you are asked to sort this numbers in Ascending order. Assume Aadhaar card numbers are available in an Notepad file stored on computer memory. Which of the following strategy you will use:

a. Sort by hand (or mentally) and update the file

b. Sort by writing a program

Question

Consider, you are given with 1000 Aadhaar card numbers and you are asked to sort this numbers in Ascending order. Assume Aadhaar card numbers are available in an database file stored on computer memory. Which of the following strategy you will use:

a. Sort by hand (or mentally) and update the file

b. Sort by writing a program

Question

Consider, you are given with 10,000 Aadhaar card numbers and you are asked to sort this numbers in Ascending order. Assume, Aadhaar card numbers are available in an database file stored on computer memory.

You are given with two sorting algorithms, say the

efficiency of Algorithm1 is $T_{A1}(n)=4n+2$ and

efficiency of Algorithm2 is $T_{A2}(n)=4n^2+2$

Which of the following strategy you will use:

a. Write a program to sort by implementing Algorithm1
b. Write a program to sort by implementing Algorithm2

Answer

Consider, you are given with 10,000 Aadhaar card numbers and you are asked to this numbers in Ascending order. Assume, Aadhaar card numbers are available in an database file stored on computer memory.

You are given with two sorting algorithms, say the

efficiency of Algorithm1 is $T_{A1}(n)=4n+2$ and

efficiency of Algorithm2 is $T_{A2}(n)=4n^2+2$

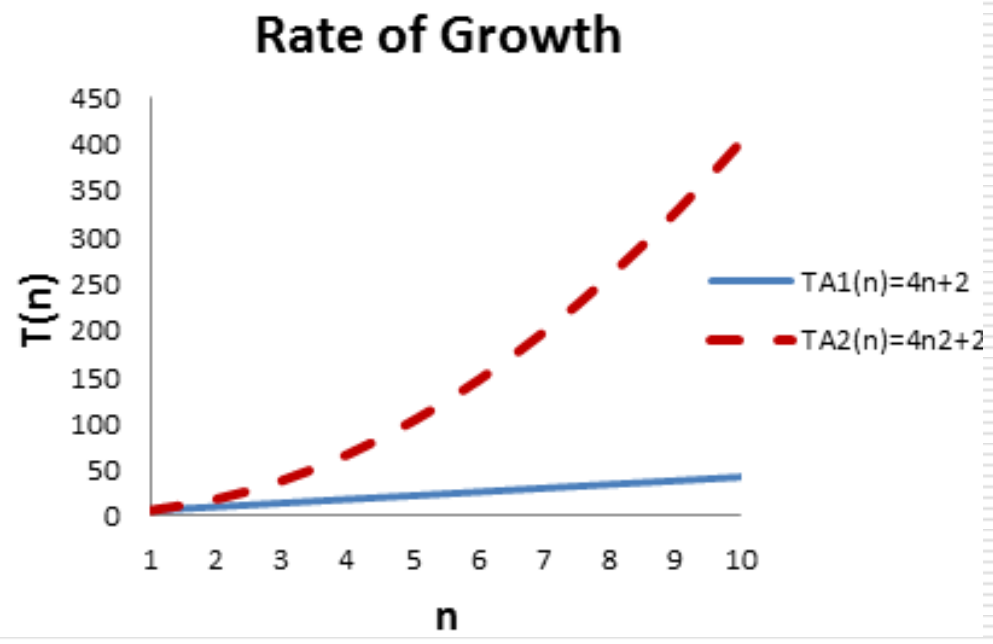
Which of the following strategy you will use:

a. Write a program to sort by implementing Algorithm1

b. Write a program to sort by implementing Algorithm2

We can choose eitherAlgo1 or Algo2 if it is one time sorting. And also on modern computer Sorting ten thousand numbers Will not take much time because it will be done in fraction of seconds

n	$T_{A1}(n)=4n+2$	$T_{A2}(n)=4n^2+2$
1	6	6
2	10	18
3	14	38
4	18	66
5	22	102
6	26	146
7	30	198
8	34	258
9	38	326
10	42	402



Question

Consider, you are given with 10,00,000 Aadhaar card numbers and you are asked sort to this numbers in Ascending order. Assume, Aadhaar card numbers are available in an database file stored on computer memory.

You are given with two sorting algorithms, say the

efficiency of Algorithm1 is $T_{A1}(n)=10^4n$ and

efficiency of Algorithm2 is $T_{A2}(n)=n^2$

Which of the following strategy you will use:

a. Write a program to sort by implementing Algorithm1
b. Write a program to sort by implementing Algorithm2

Answer

Consider, you are given with 10,00,000 Aadhaar card numbers and you are asked sort to this numbers in Ascending order. Assume, Aadhaar card numbers are available in an database file stored on computer memory.

You are given with two sorting algorithms, say the

efficiency of Algorithm1 is $T_{A1}(n)=10^4n$ and

efficiency of Algorithm2 is $T_{A2}(n)=n^2$

Which of the following strategy you will use:

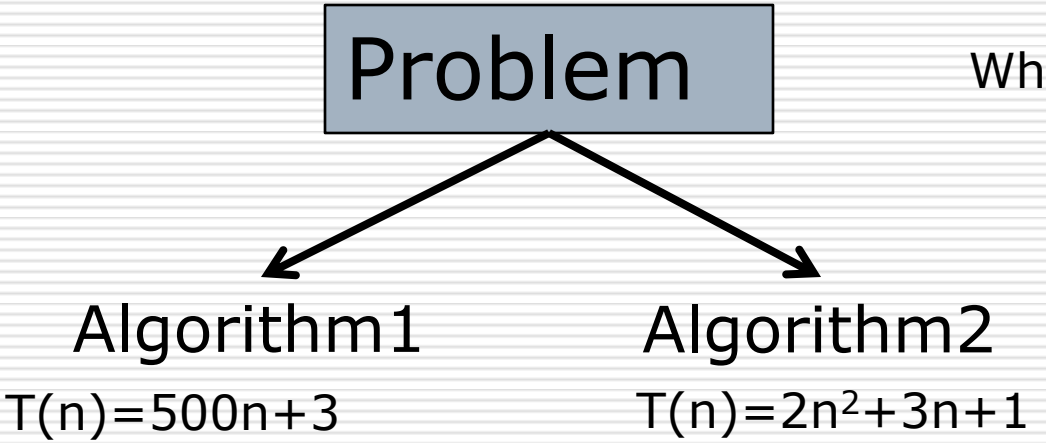
a. Write a program to sort by implementing Algorithm1

b. Write a program to sort by implementing Algorithm2

n	$T_{A1}(n)=(10\text{^}4)n$	$T_{A2}(n)=n\text{^}2$
1.0E+01	1.0E+05	1.0E+02
1.0E+02	1.0E+06	1.0E+04
1.0E+03	1.0E+07	1.0E+06
1.0E+04	**1.0E+08**	**1.0E+08**
1.0E+05	1.0E+09	1.0E+10
1.0E+06	1.0E+10	1.0E+12

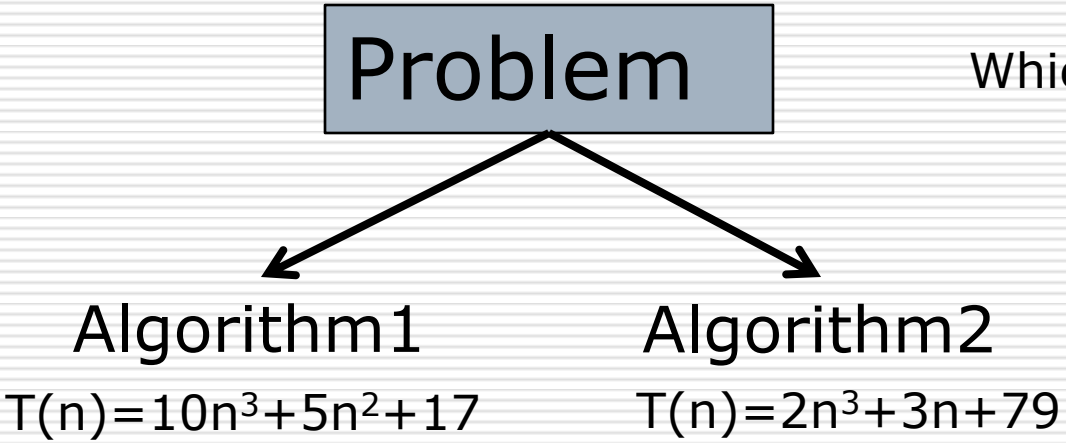
For $n<=10^4$,$10^4n > n^2$ *Algo2* is better
But for $n> 10^4$ *Algo1* is better

Question



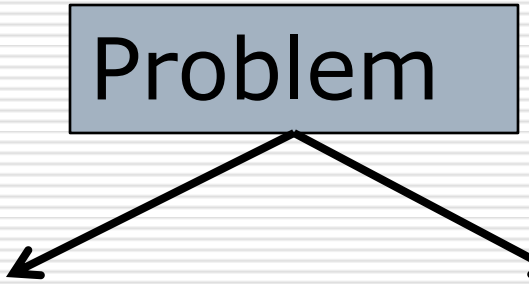
Which algorithm to use ?

Question



Which algorithm to use ?

Question



Which algorithm to use ?

Algorithm1

$T(n)=10n^3+5n^2+17$

Algorithm2

$T(n)=2n^3+3n+79$

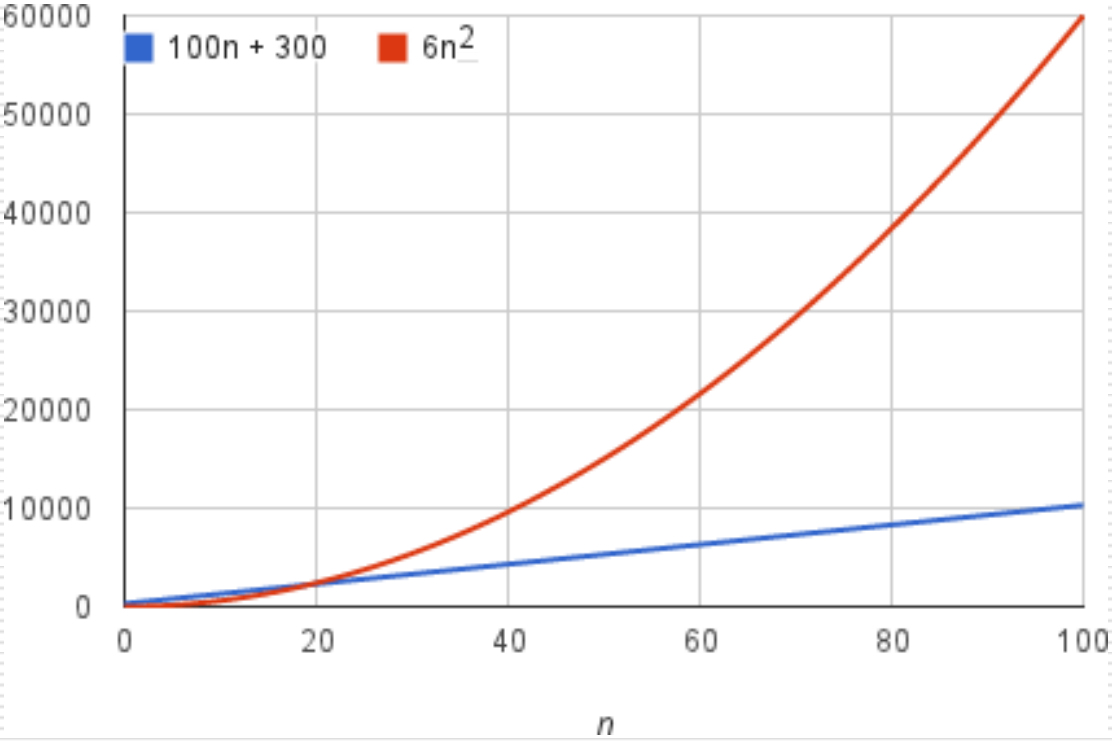
Answer:
The above two time complexities are tedious to be judged.
Hence we will go with **approximating the time complexities** i.e., finding Out the class to which the algorithm belongs because as n tends to infinity (n -> ∞) i.e., when n takes large values the value of $(5n^2+17)$ and the value of $(3n+79)$ will go out. Therefore we will be worrying about $10n^3$ and $2n^3$

Example

Consider T(n)= $\mathbf{6n^2 + 100n + 300}$

Example

Consider T(n)= **$6n^2 + 100n + 300$**



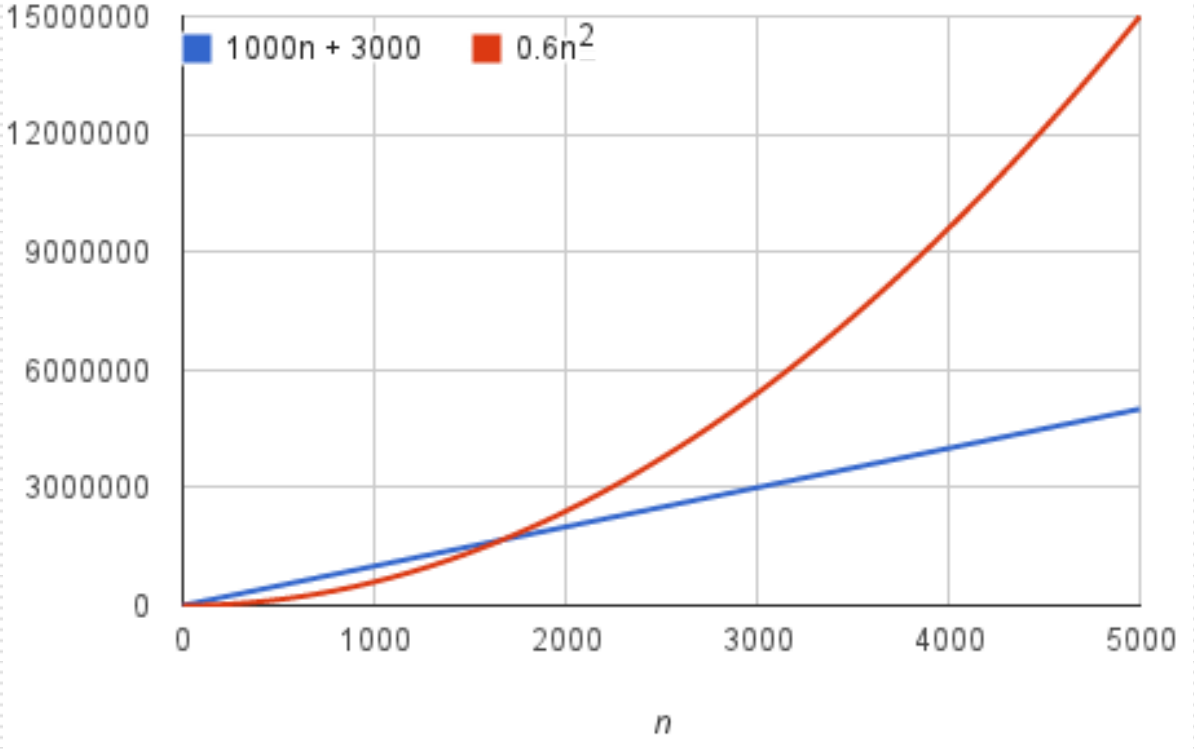
The $6n^2$ term becomes larger than the remaining terms, 100n + 300, once n becomes large enough, 20 in this case.

Example

Consider T(n)= $\mathbf{0.6n^2 + 1000n + 3000}$

Example

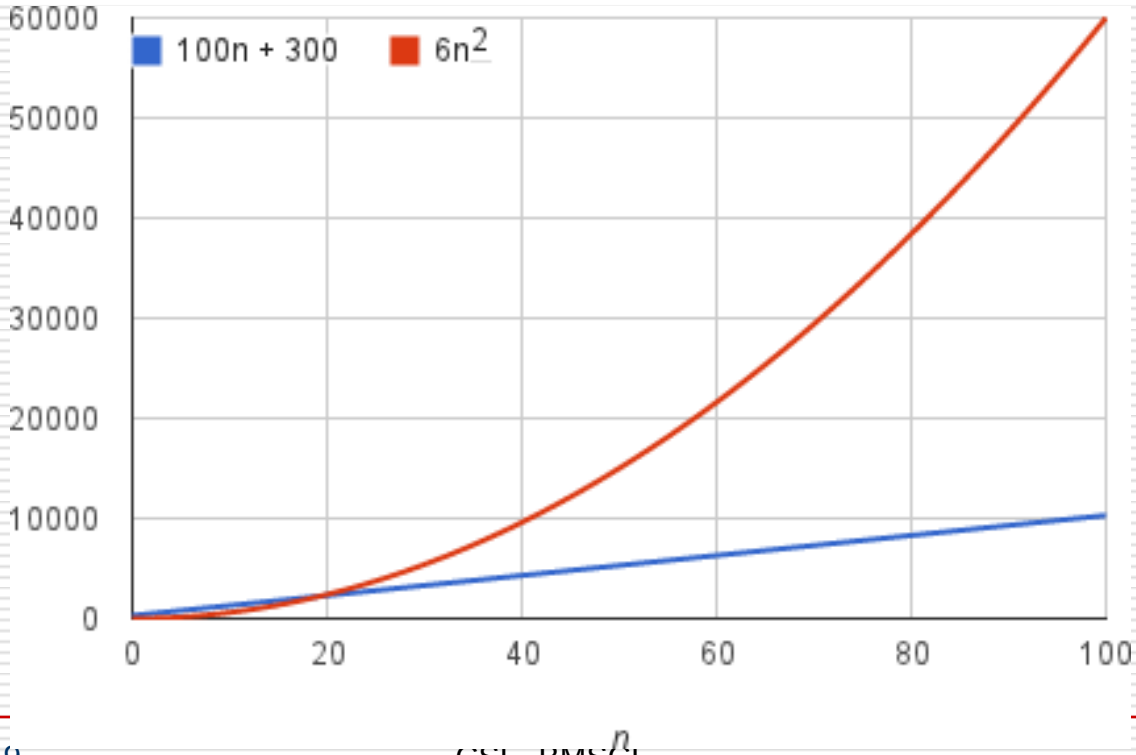
Consider T(n)= **$0.6n^2 + 1000n + 3000$**



The $0.6n^2$ term becomes larger than the remaining terms, 1000n + 3000, once n becomes large enough, 1700 in this case.

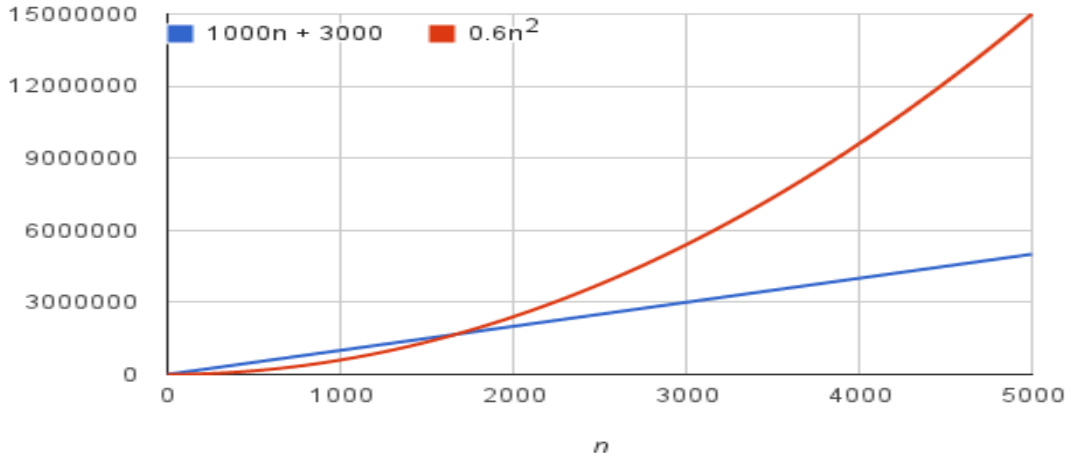
Explanation

For example, suppose that an algorithm, running on an input of size n, takes **$6n^2 + 100n + 300$** machine instructions. The $6n^2$ term becomes larger than the remaining terms, 100n + 300, once n becomes large enough, 20 in this case. Here's a chart showing values of **$6n^2$** and **$100n + 300$** for values of *n* from 0 to 100:



Explanation

We would say that the running time of this algorithm grows as n^2, dropping the coefficient 6 and the remaining terms $100n + 300$. It doesn't really matter what coefficients we use; as long as the running time is $an^2 + bn + c$, for some numbers $a > 0$, b, and c, there will always be a value of n for which an^2 is greater than $bn + c$, and this difference increases as n increases. For example, here's a chart showing values of $0.6n^2$ and $1000n + 3000$ so that we've reduced the coefficient of n^2 by a factor of 10 and increased the other two constants by a factor of 10:



The value of n at which $0.6n^2$ becomes greater than $1000n + 3000$ has increased, but there will always be such a crossover point, no matter what the constants.

What is a Time Complexity/Order of Growth?

- Time Complexity/Order of Growth defines the amount of time taken by any program with respect to the size of the input.
- Time Complexity specifies how the program would behave as the order of size of input is increased. So, Time Complexity is just a function of size of its input.

Some of basic and most common time complexities such as:

- Constant Time Complexity: Constant running time
- Linear Time Complexity *(n)* : Linear running time
- Logarithmic Time Complexity *(log n)* : Logarithmic running time
- Log-Linear Time Complexity *(n log n)* : Log--linear running time
- Polynomial Time Complexity *(n^c)* : Polynomial running time (c is a constant)
- Exponential Time Complexity *(c^n)* : Exponential running time (c is a constant being raised to a power based on size of input)

What is Constant Time Complexity?

- The code that runs in fixed amount of time or has fixed number of steps of execution no matter what is the size of input has constant time complexity. For instance, let's try and derive a Time Complexity for following code:

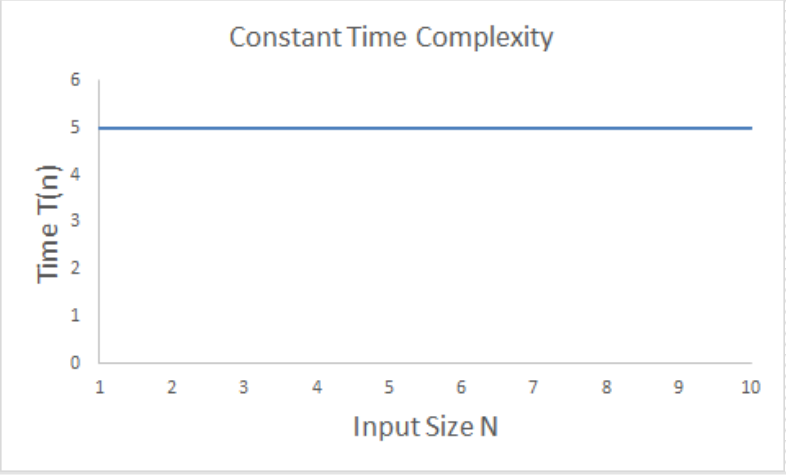
What is Constant Time Complexity?

- The code that runs in fixed amount of time or has fixed number of steps of execution no matter what is the size of input has constant time complexity. For instance, let's try and derive a Time Complexity for following code:

```
def my_sum(a, b):
        return a+b
```

If we call this function by my_sum(2, 5) it will return 7 in 1 step. That single step of computation is summing a and b. No matter how large is the size of input i.e. a and b is, it will always return the sum in 1 step.

So, the Time Complexity of the above code is a Constant Time Complexity.



What is Linear Time Complexity?

The code whose Time Complexity or Order of Growth increases linearly as the size of the input is increased has Linear Time Complexity.

For instance, let's see this code which returns the sum of a list.

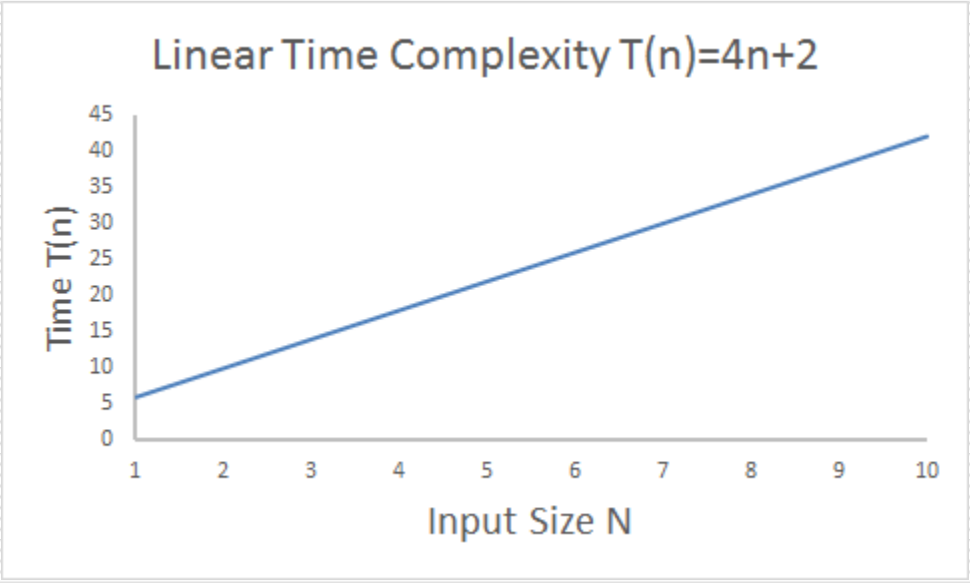
What is Linear Time Complexity?

The code whose Time Complexity or Order of Growth increases linearly as the size of the input is increased has Linear Time Complexity.

For instance, let's see this code which returns the sum of a list.

```
for(i=0; i < n; i++)
        temp=10*30;
```

T(n)= 4n+2



What is Logarithmic Time Complexity?

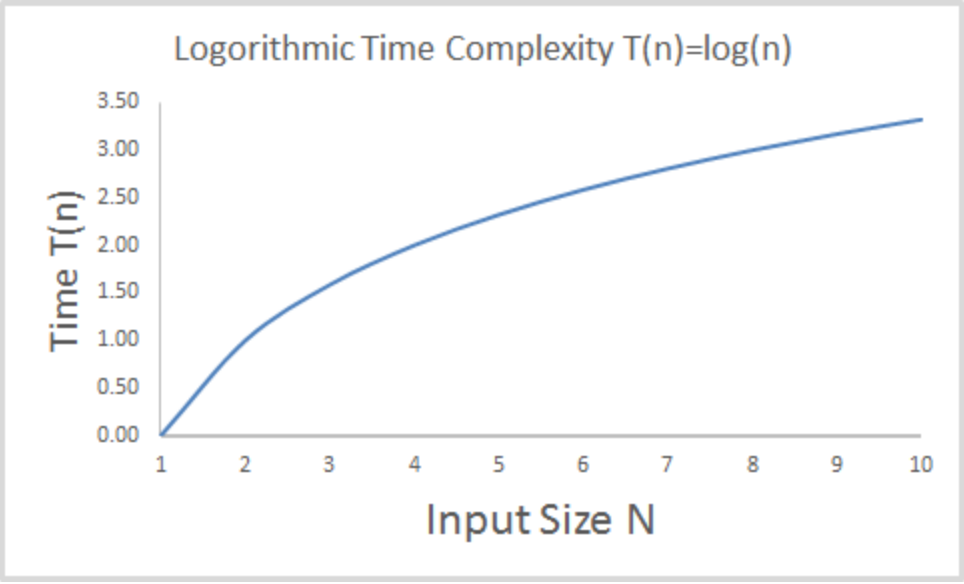
- When the size of input is N but the number of steps to execute the code is log(N), such a code is said to be executing in Logarithmic Time. This definition is quite vague but if we take an example, it will be quite clear.

What is Logarithmic Time Complexity?

- When the size of input is N but the number of steps to execute the code is log(N), such a code is said to be executing in Logarithmic Time. This definition is quite vague but if we take an example, it will be quite clear.
- Let's say we have a very large number which is a power of 2 i.e. we have 2^x. We want to find x. For eg: 64 = 2^6. So x is 6.

```
pow(n){
   x = 0
   while (n > 1){
      n = n/2
      x = x+1}
   return x }
```

T(n) = log(n)



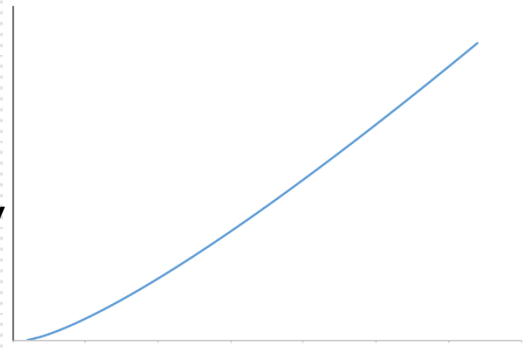
What is Log-Linear Time Complexity?

- When we call a Logarithmic Time Algorithm inside a loop, it would result into a Log-Linear Time Complexity program.
- For example: Let's say we have long sorted list of size N. And we have Q numbers, for each of those Q numbers we have to find the index of it in the given list.

```
for i in Qlist:
    print binary_search(x, search_list)  #This statement is
                                         #executed Q times
```

Analyzing above code, we know that the call to Binary Search function takes (log N) times. We are calling it Q times. Hence the overall time complexity is Q(log N).

Log Linear Time Complexity



What is Polynomial Time Complexity?

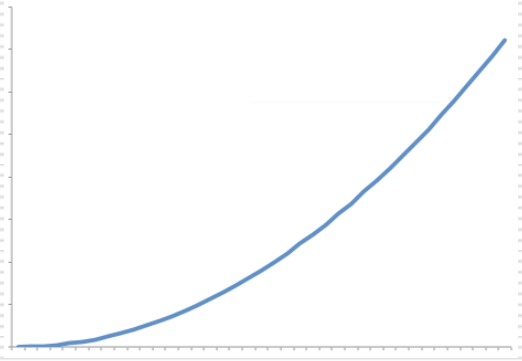
When the computation time increases as function of N raised to some power, N being the size of input. Such a code has Polynomial Time Complexity.

For example, let's say we have a list of size N and we have nested loops on that list.

```
for i in N:
    for j in N:
        # Some processing
```

In the above code, the processing part is executed N*N times i.e. N^2 times. Such a code has *(N^2)* time complexity.

Quadratic Time Complexity

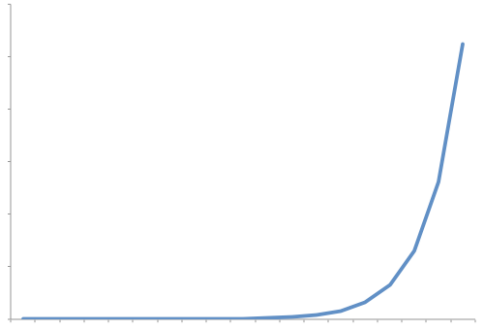


What is Exponential Time Complexity?

- When the computation time of a code increases as function of X^N, N being the size of input. Such a code has Polynomial Time Complexity.
- For example, following recursive code to find Nth fibonacci number has Time Complexity as (**2^N**)

```
def F(n):
    if n == 0:
        return 0
    elif n == 1:
        return 1
    else:
        return F(n-1) + F(n-2) # For every call to F, we make 2 more calls to F itself
```

Exponential Time Complexity



-We consider only leading term in the expression T(n), since lower-order terms are relatively insignificant for large n.

-We are moving from **Actual cost** to **Growth of Cost (or Rate of Growth).**

-We are interested to know what is the term that dominates so that if we arbitrarily keep on increasing *n* that is the term which primarily decide how the computing time will grow.

Order of Growth

Measuring the performance of an algorithm in relation with the input size n is called Order of growth or Rate of Growth

n	log2(n)	n	nlog2(n)	n^2	2^n	n^3
1	0.00	1	0.00	1	2	1
2	1.00	2	2.00	4	4	8
3	1.58	3	4.75	9	8	27
4	2.00	4	8.00	16	16	64
5	2.32	5	11.61	25	32	125
6	2.58	6	15.51	36	64	216
7	2.81	7	19.65	49	128	343
8	3.00	8	24.00	64	256	512
9	3.17	9	28.53	81	512	729
10	3.32	10	33.22	100	1024	1000

Order of growth for varying input size of n

Order of Growth

Measuring the performance of an algorithm in relation with the input size n is called order of growth . Some of the popular order which we will see is:-

- Order 1 : Constant.
- Order log(n) : Logarithmic
- Order (n) : linear
- Order nlog(n): log linear, occurs very often
- order (n ^ C) : polynomial
- order (C ^ n) : exponential

n	log2(n)	n	nlog2(n)	n^2	2^n	n^3
1	0.00	1	0.00	1	2	1
2	1.00	2	2.00	4	4	8
3	1.58	3	4.75	9	8	27
4	2.00	4	8.00	16	16	64
5	2.32	5	11.61	25	32	125
6	2.58	6	15.51	36	64	216
7	2.81	7	19.65	49	128	343
8	3.00	8	24.00	64	256	512
9	3.17	9	28.53	81	512	729
10	3.32	10	33.22	100	1024	1000

Order of growth for varying input size of n

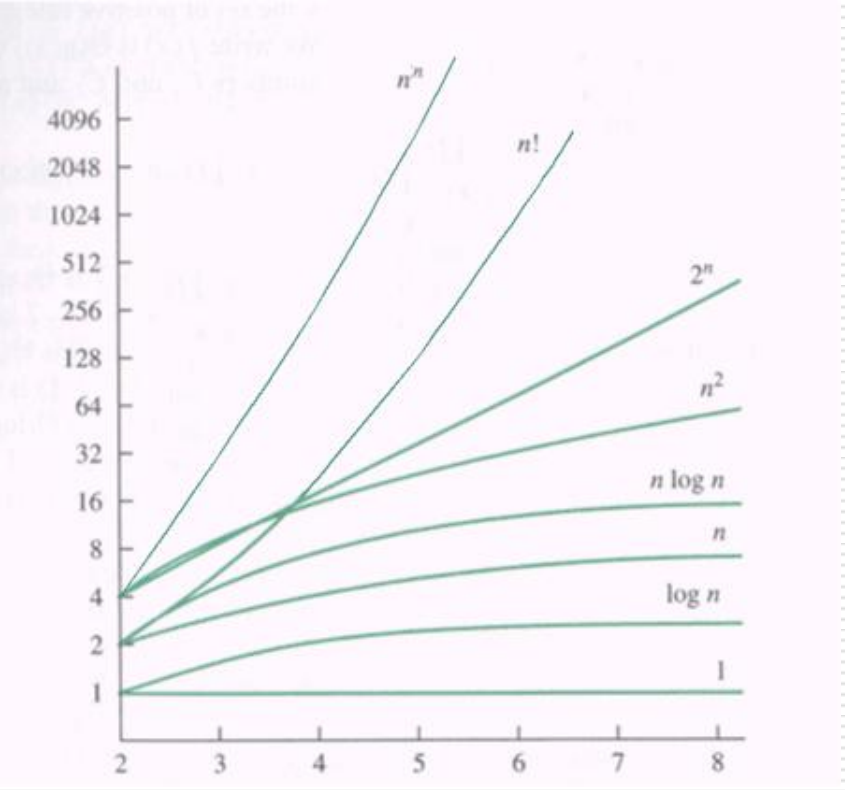
Order of Growth

Measuring the performance of an algorithm in relation with the input size n is called order of growth

Some of the popular order which we will see is:-

- Order 1 : Constant.
- Order log(n) : Lograthimic
- Order (n) : liner
- Order nlog(n): log liner, occurs very often
- order (n ^ C) : polynomial
- order (C ^ n) : exponential

n	log2(n)	n	nlog2(n)	n^2	2^n	n^3
1	0.00	1	0.00	1	2	1
2	1.00	2	2.00	4	4	8
3	1.58	3	4.75	9	8	27
4	2.00	4	8.00	16	16	64
5	2.32	5	11.61	25	32	125
6	2.58	6	15.51	36	64	216
7	2.81	7	19.65	49	128	343
8	3.00	8	24.00	64	256	512
9	3.17	9	28.53	81	512	729
10	3.32	10	33.22	100	1024	1000



Order of growth for varying input size of n

Quiz

- Which kind of growth best characterizes each of these functions?

	Constant	Linear	Polynomial	Exponential
$3n$	○	○	○	○
$3n^2$	○	○	○	○
2^n	○	○	○	○
$(3/2)^n$	○	○	○	○
1000	○	○	○	○
1	○	○	○	○
$(3/2)n$	○	○	○	○
$2n^3$	○	○	○	○

Answer

- Which kind of growth best characterizes each of these functions?

	Constant	Linear	Polynomial	Exponential
$3n$	○	●	○	○
$3n^2$	○	○	●	○
2^n	○	○	○	●
$(3/2)^n$	○	○	○	●
1000	●	○	○	○
1	●	○	○	○
$(3/2)n$	○	●	○	○
$2n^3$	○	○	●	○

Quiz

- Rank these functions according to their growth, from slowest growing (at the top) to fastest growing (at the bottom).

n^2
2^n
n
n^3
$(3/2)^n$
1

Answer

- Rank these functions according to their growth, from slowest growing (at the top) to fastest growing (at the bottom).

1

n

n^2

n^3

$(3/2)^n$

2^n

Quiz

- Rank these functions according to their growth, from slowest growing to fastest growing.

$6n^3$
$n \log_6 n$
$4n$
$8n^2$
$\log_2 n$
$n \log_2 n$
$\log_8 n$
64
8^{2n}

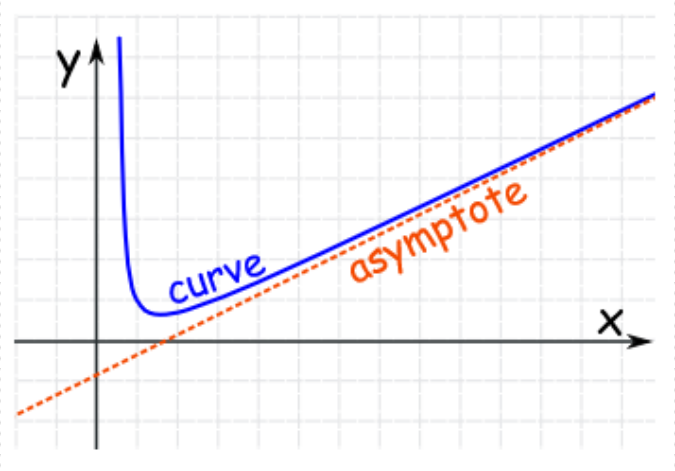
Answer

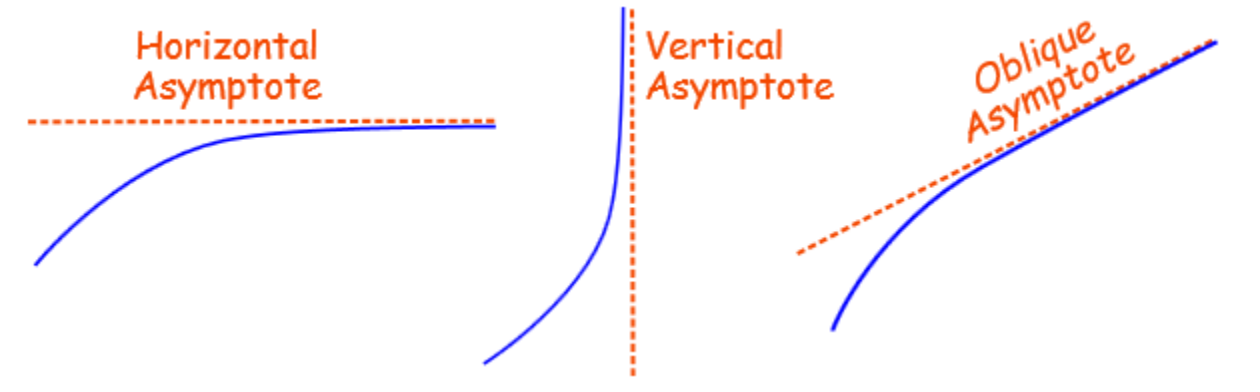
- Rank these functions according to their growth, from slowest growing to fastest growing.

64
$\log_8 n$
$\log_2 n$
$4n$
$n \log_6 n$
$n \log_2 n$
$8n^2$
$6n^3$
8^{2n}

Asymptote

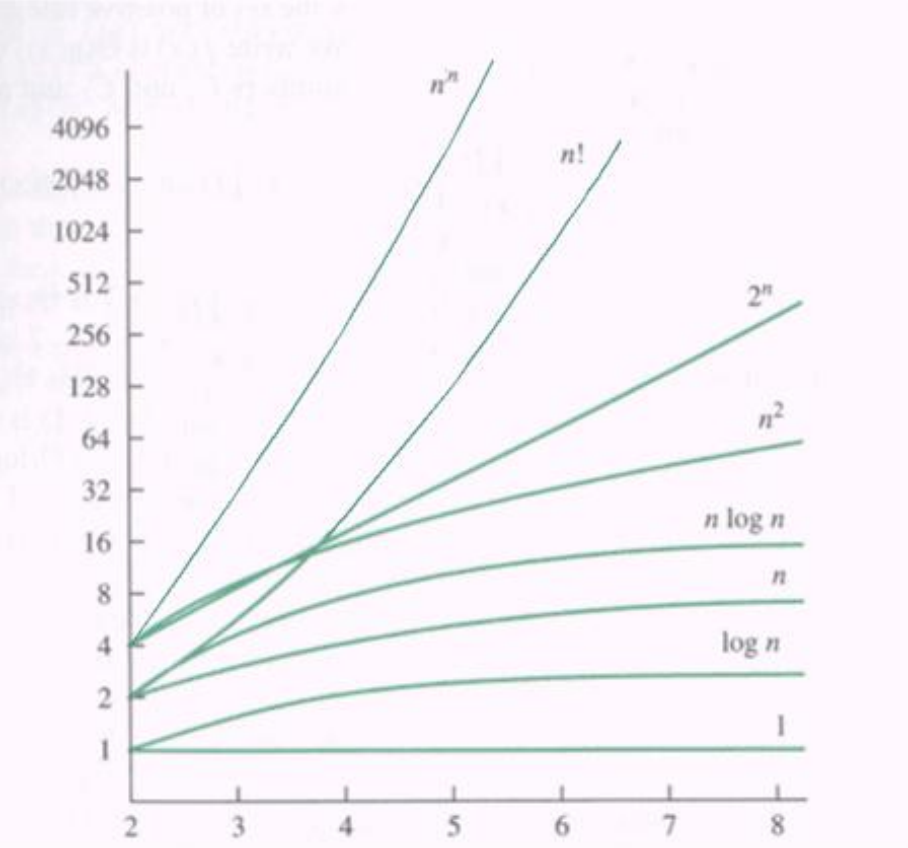
Asymptote: A straight line that continually approaches a given curve but does not meet it at any finite distance





Rate of Growth ordering

$$O(1) < O(\log n) < O(n) < O(n \log n) < O(n^2) < O(2^n) < O(n!)$$



Asymptotic notation

- **Asymptotic notation of an algorithm is a mathematical representation of its complexity**
- In asymptotic notation, when we want to represent the complexity of an algorithm, we use only the most significant terms in the complexity of that algorithm and ignore least significant terms in the complexity of that algorithm (Here complexity may be Space Complexity or Time Complexity).

For example, consider the following time complexities of two algorithms...

- **Algorithm 1 : $5n^2 + 2n + 1$**
- **Algorithm 2 : $10n^2 + 8n + 3$**

Generally, when we analyze an algorithm, we consider the time complexity for larger values of input data (i.e. '**n**' value). In above two time complexities, for larger value of '**n**' the term in algorithm 1 '**2n + 1**' has least significance than the term '**$5n^2$**', and the term in algorithm 2 '**8n + 3**' has least significance than the term '**$10n^2$**'.

Here for larger value of '**n**' the value of most significant terms (**$5n^2$** and **$10n^2$**) is very larger than the value of least significant terms (**2n + 1** and **8n + 3**). So for larger value of '**n**' we ignore the least significant terms to represent overall time required by an algorithm. In asymptotic notation, we use only the most significant terms to represent the time complexity of an algorithm.

By **dropping the less significant terms** and the **constant coefficients**, we can focus on the important part of an algorithm's running time—its **rate of growth**—without getting mired in details that complicate our understanding. When we drop the constant coefficients and the less significant terms, we use **asymptotic notation**.

We'll see THREE types of Asymptotic Notations:

Big - Oh (O) UPPER BOUNDING function

Big - Omega (Ω) LOWER BOUNDING function

Big - Theta (Θ) ORDER or TIGHT BOUNDING function

Analysis of Linear Search

```
Algorithm SequentialSearc(A[0..n-1,K)
i=0
While i<n and A[i]!=K do
        { i=i+1 }
If i<n
   return i
else
   return -1
```

Question:

If the Key element is in the first position of the Array then How many times the operation i=i+1 will be executed ?

If the Key element is in the last position of the Array then How many times the operation i=i+1 will be executed ?

Analysis of Linear Search

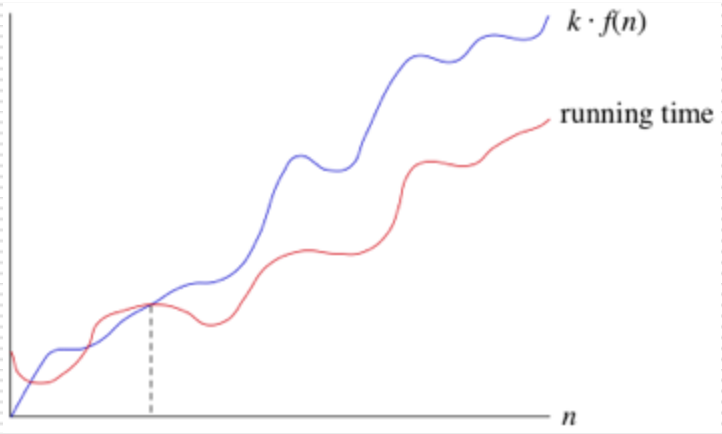
```
Algorithm SequentialSearc(A[0..n-1,K)
i=0
While i<n and A[i]!=K do
        { i=i+1 }
If i<n
   return i
else
   return -1
```

Answer:

Find, if the Key element is in the first position of the Array then
How many times the operation i=i+1 will be executed ?
Find, if the Key element is in the last position of the Array then
How many times the operation i=i+1 will be executed ?
Find the total lower bound and upper bound (Best and Worst case)
Running time ? $T_{lower}(n)=1$ $T_{upper}(n)=n$

Summarizing Big-O, Big-Omega, Big-theta

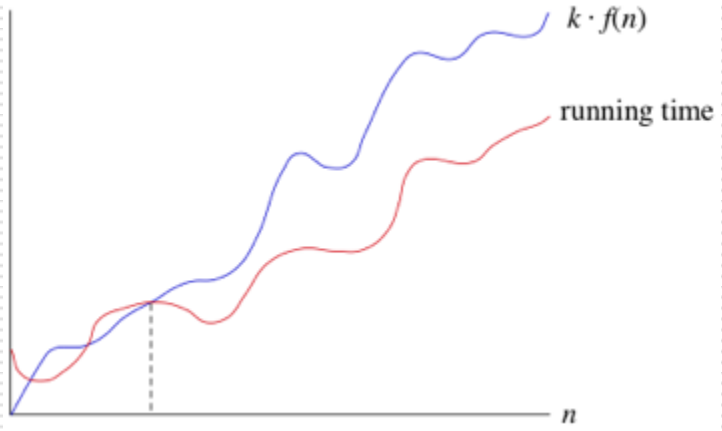
Big-O Upper Bound



no

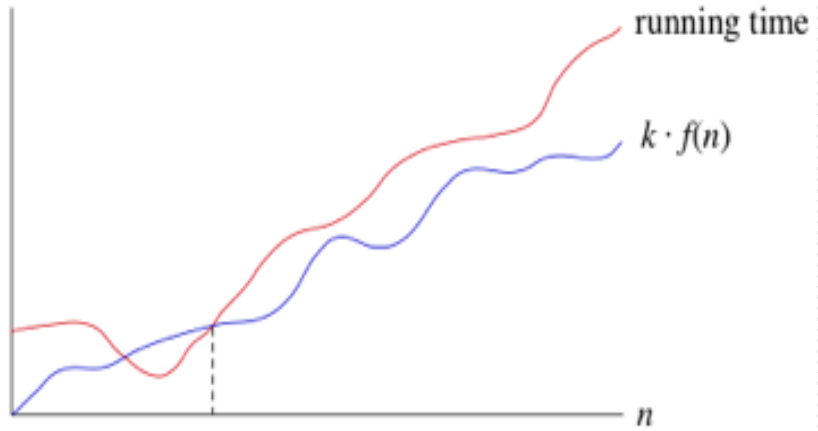
Summarizing Big-O, Big-Omega, Big-theta

Big-O Upper Bound



no

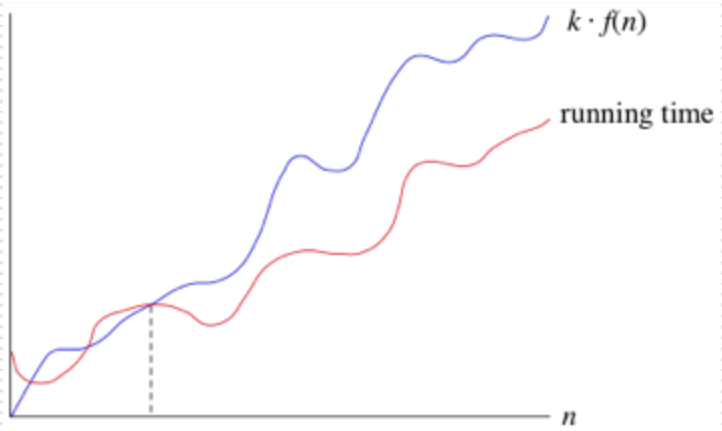
Big-Omega Lower Bound



no

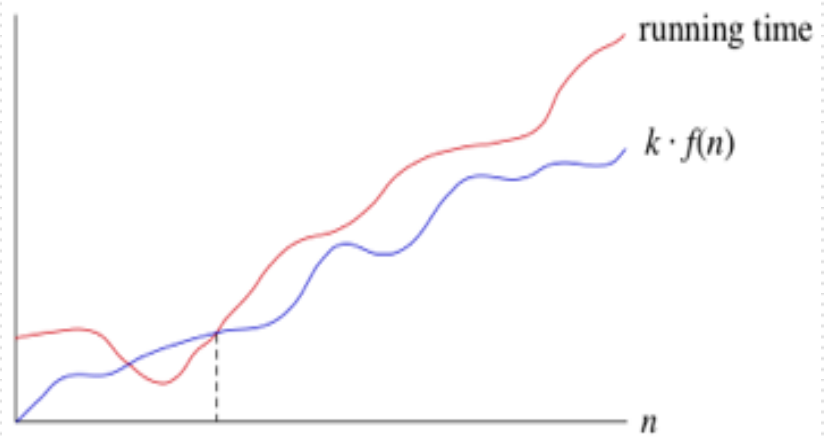
Summarizing Big-O, Big-Omega, Big-theta

Big-O Upper Bound



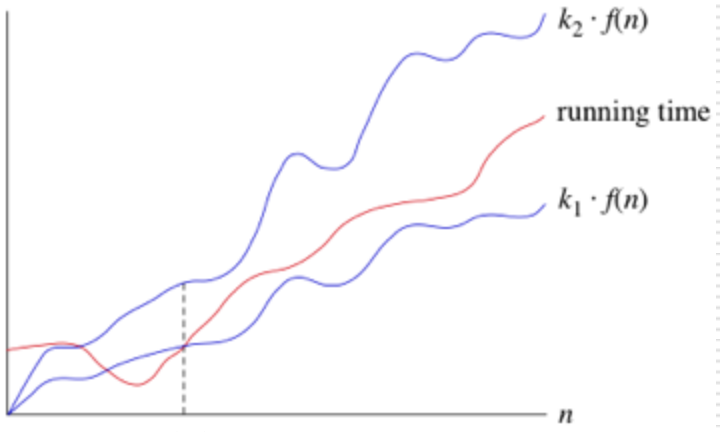
no

Big-Omega Lower Bound



no

Big-theta
Tight bound or Order bound



no

Question

Go through the following pseudocode

```
containsZero(arr, n){ #assume normal array of length n
  for i=1 to n {
      if arr[i] == 0  return true
    }
  return false
```

What's the lower bound or best case? Well, if the array we give it has 0 as the first value, it will take what time ?

What's the worst case? If the array doesn't contain 0, it will take what time

Question

Go through the following pseudocode

```
containsZero(arr, n){ #assume normal array of length n
  for i=1 to n {
      if arr[i] == 0  return true
    }
  return false
```

What's the lower bound or best case? Well, if the array we give it has 0 as the first value, it will take what time ?

- Constant time: Ω (1)

What's the worst case? If the array doesn't contain 0, it will take what time

- It will iterate through the whole array: O(n)

Question

Go through the following pseudocode

```
printNums(arr,n){
for i=1 to n {
  print(arr[i]);
}
```

Can you think of a best case and worst case??

Question

Go through the following pseudocode

```
printNums(arr,n){
for i=1 to n {
  print(arr[i]);
}
```

Can you think of a best case and worst case??

We can't! No matter what array we give it, we have to iterate through every value in the array. So the function will take AT LEAST n time ($\Omega(n)$), but we also know it won't take any longer than n time ($O(n)$). What does this mean? Our function will take exactly n time i.e., **$\Theta(n)$**

Asymptotic notation

To compare and rank orders of growth or rate of growth of the algorithms, Computer Scientists use three notations:

- **Big - Oh (O) UPPER BOUNDING function**
- **Big - Omega (Ω) LOWER BOUNDING function**
- **Big - Theta (Θ) ORDER BOUNDING function**

Big - Oh (O) UPPER BOUNDING function: Informal Intr.

- Let us consider t(n) and g(n) are non-negative functions (or expressions) which take non-negative arguments.
- **O(g(n)) is set of all functions with a smaller or same order of growth g(n)** (to within a constant multiple, as n goes to infinity).

Ex: $n \in O(n^2)$ for all n>=1

Informal Introduction

- Let us consider t(n) and g(n) are non-negative functions (or expressions) which take non-negative arguments.
- **O(g(n)) is set of all functions with a smaller or same order of growth g(n)** (to within a constant multiple, as goes to infinity).

Ex: $n \in O(n^2)$ for all n>=1

$n \in O(n^2)$ $\forall$ n>=1	
$n <= n^2$	
n	n^2
1	1
2	4
3	9

Informal Introduction

- Let us consider t(n) and g(n) are non-negative functions (or expressions) which take non-negative arguments.
- **O(g(n)) is set of all functions with a smaller or same order of growth g(n)** (to within a constant multiple, as goes to infinity).
- Ex: $100n+5 \in O(n^2)$ for all n>=1000

Big-Oh Upper Bounding:Informal Intr.

- Let us consider t(n) and g(n) are non-negative functions (or expressions) which take non-negative arguments.
- **O(g(n)) is set of all functions with a smaller or same order of growth g(n)** (to within a constant multiple, as goes to infinity).
- Ex: $100n+5 \in O(n^2)$ for all n>=1000

$100n+5 \in O(n^2)$ $\forall$ n>=10^3		
$100n+5 <= n^2$		
n	$100n+5$	n^2
10^2	10^4+5	10^4
10^3	**10^5+5**	**10^6**
10^4	10^6+5	10^8

Question

- Check whether the statement

$\frac{1}{2}n(n-1) \in O(n^2)$ is true

Answer

- Check whether the statement

$\frac{1}{2}n(n-1) \in O(n^2)$ is true

True for all n >=1

n	(1/2)n(n-1)	n^2
1	0	1
2	1	4
3	3	9
4	6	16
5	10	25

Question

Check which of the following statement is true

- $n^3 \in O(n^2)$
- $n^3 \notin O(n^2)$

Answer

Check which of the following statement is true

- $n^3 \in O(n^2)$ False
- $n^3 \notin O(n^2)$ True

Question

Check whether the following statement is true

- $n^4 + n + 1 \notin O(n^2)$

Answer

Check whether the following statement is true

- $n^4 + n + 1 \notin O(n^2)$

True

Formal Definition

Big - Oh (O) UPPER BOUNDING function

Big - Oh (O) UPPER BOUNDING function

- Big - Oh notation is used to define the **upper bound** of an algorithm in terms of Time Complexity. That means Big - Oh notation always indicates the maximum time required by an algorithm for all input values. That means Big - Oh notation describes the worst case of an algorithm time complexity.

Formal Definition

- A function f(n) is said to be in O(g(n)) , denoted

f(n) ∈ O(g(n)) (or f(n) = O(g(n))),

if f(n) is bounded above by some constant multiple of g(n) for all large n, i.e., if there exist some positive constant c and some nonnegative integer n_0 such that

$$f(n) <= cg(n) \text{ for all } n>=n_0$$

(f(n) is less than or equal to cg(n) for all values of n greater than or equal to n_0)

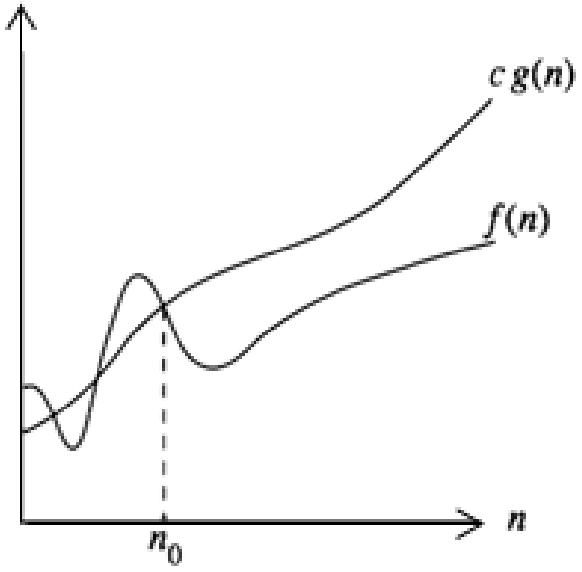
Big - Oh (O) UPPER BOUNDING function

- Big - Oh notation is used to define the **upper bound** of an algorithm in terms of Time Complexity. That means Big - Oh notation always indicates the maximum time required by an algorithm for all input values. That means Big - Oh notation describes the worst case of an algorithm time complexity.

Formal Definition

- A function f(n) is said to be in O(g(n)) , denoted **f(n) ∈ O(g(n)) (or f(n) = O(g(n)))**, if f(n) is bounded above by some constant multiple of g(n) for all large n, i.e., if there exist some positive constant c and some nonnegative integer n_0 such that

$f(n) <= cg(n)$ for all $n>=n_0$



n is Size of program's input.

f(n) Any real world function. Example: - Running time of a machine.

g(n) Another function that we want to use as an upper-bound. Not a real world function but preferably simple.

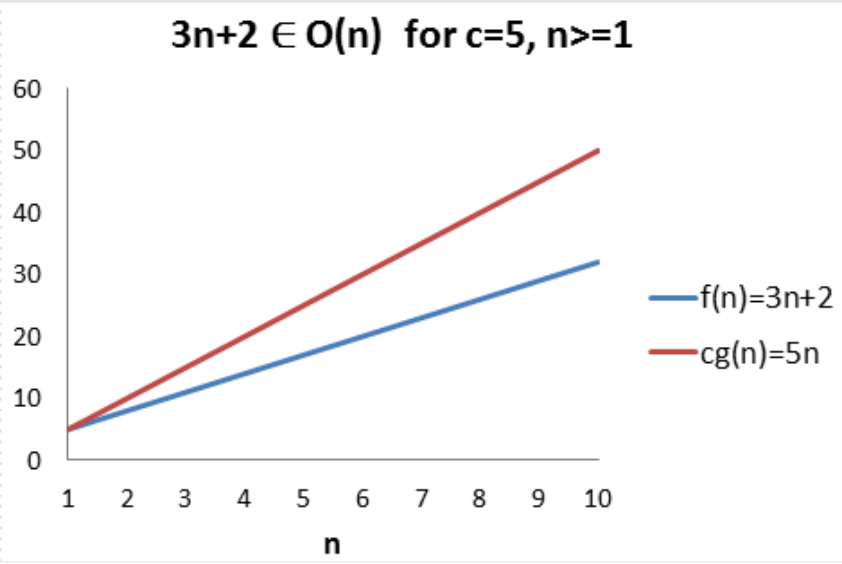
Example

- Consider the following f(n) and g(n)...
  f(n) = 3n + 2
  g(n) = n
  If we want to represent **f(n)** as **O(g(n))** then it must satisfy **f(n) <= Cg(n)** for all values of **C > 0** and **n_0>= 1**

Example

- Consider the following t(n) and g(n)...
  f(n) = 3n + 2
  g(n) = n
  If we want to represent **f(n)** as **O(g(n))** then it must satisfy **f(n) <= Cg(n)** for all values of **C > 0** and **n_0>= 1**
- f(n) <= C g(n) ⇒ 3n + 2 <= Cn
  Above condition is always TRUE for all values of **C = 5** and **n >= 1**.

		g(n)=n
n	**f(n)=3n+2**	**cg(n)=5n**
1	5	5
2	8	10
3	11	15
4	14	20
5	17	25
6	20	30
7	23	35
8	26	40
9	29	45
10	32	50



By using Big - Oh notation we can represent the time complexity as follows...
3n + 2 ∈ O(n) or 3n + 2 = O(n)

Example

Consider the following t(n) and g(n)...

$t(n) = 6*2^n + n^2$

$g(n) = 2^n$

Represent $6*2^n + n^2 \in O(2^n)$

C= ?? , n_0= ??

i.e., $6*2^n + n^2 <= C(2^n)$ for all $n >= n_0$

Example

Consider the following t(n) and g(n)...

$t(n) = 6*2^n + n^2$

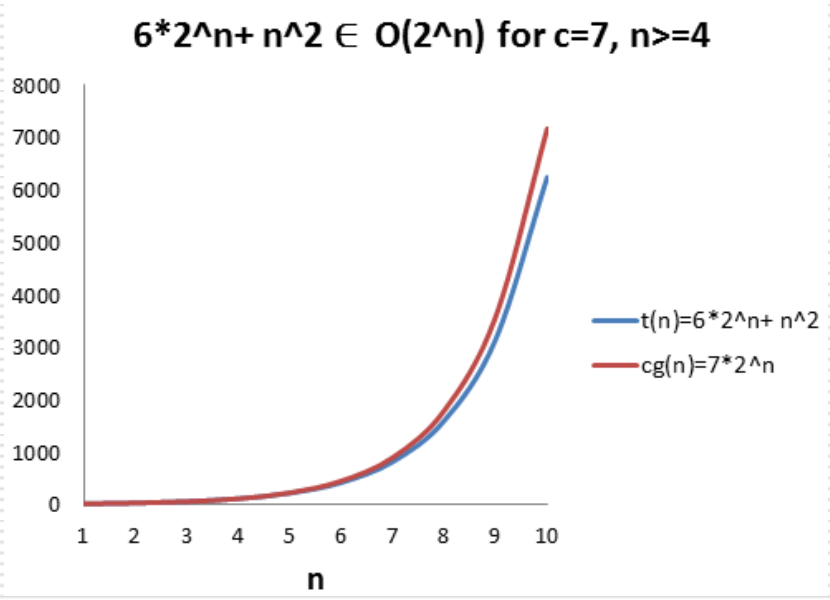
$g(n) = 2^n$

Represent $6*2^n + n^2 \in O(2^n)$

$C=7$, $n_0=4$

i.e., $6*2^n + n^2 <= 7(2^n)$ for all $n >= 4$

		g(n)=2^n
n	t(n)=6*2^n+ n^2	cg(n)=7*2^n
1	13	14
2	28	28
3	57	56
4	112	112
5	217	224
6	420	448
7	817	896
8	1600	1792
9	3153	3584
10	6244	7168



Question

Find the values for **c** and **n** to show that the assertion $3n^3 + 2n^2 \in O(n^3)$ is true

Question

Find the values for **c** and **n** to show that the assertion $3n^3 + 2n^2 \in O(n^3)$ is true

Answer:

- $C=5, n_0=1$
- $3n^3 + 2n^2 <= 5(n^3)$ for all $n >=1$

Question

- Is the following assertion True

$3^n \notin O(2^n)$

Question

- Is the following assertion True

$3^n \notin O(2^n)$

Answer: Yes

Question

Prove that $n^3 + n^2 + n \in O(n^3)$

Question

Prove that $n^3 + n^2 + n \in O(n^3)$

- ***Sol.*** For C=3, and n_0=1,

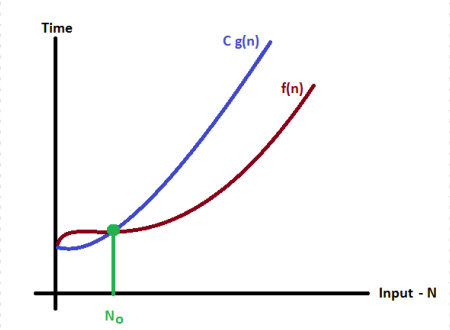
Big - Oh Notation (O)

- Big - Oh notation is used to define the **upper bound** of an algorithm in terms of Time Complexity.
  That means Big - Oh notation always indicates the maximum time required by an algorithm for all input values. That means Big - Oh notation describes the worst case of an algorithm time complexity.
  Big - Oh Notation can be defined as follows...

Consider function f(n) the time complexity of an algorithm and g(n) is the most significant term. If f(n) <= C g(n) for all n >= n_0, C > 0 and n_0 >= 1. Then we can represent f(n) as O(g(n)).

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

- Consider the following graph drawn for the values of f(n) and C g(n) for input (n) value on X-Axis and time required is on Y-Axis



- In above graph after a particular input value n_0, always C g(n) is greater than f(n) which indicates the algorithm's upper bound.

Big-O

- Big-O, commonly written as **O**, is an Asymptotic Notation for the worst case, or ceiling of growth for a given function. It provides us with an ***asymptotic upper bound*** for the growth rate of runtime of an algorithm. Say f(n) is your algorithm runtime, and g(n) is an arbitrary time complexity you are trying to relate to your algorithm. f(n) is O(g(n)), if for some real constants c (c > 0) and n_0, f(n) <= c g(n) for every input size n (n > n_0).

Example 1

- f(n) **=** 3log n **+** 100 g(n) **=** log n
- Is f(n) =O(g(n))? Is 3 log n + 100=O(log n)? Let's look to the definition of Big-O.
- **3**log n **+** 100 **<=** c ***** log n
- Is there some pair of constants c, n_0 that satisfies this for all n > n_0?
- **3**log n **+** 100 **<=** 150 ***** log n, n **>** 2 (undefined at n **=** 1)
- Yes! The definition of Big-O has been met therefore f(n) is O(g(n)).

Big-O

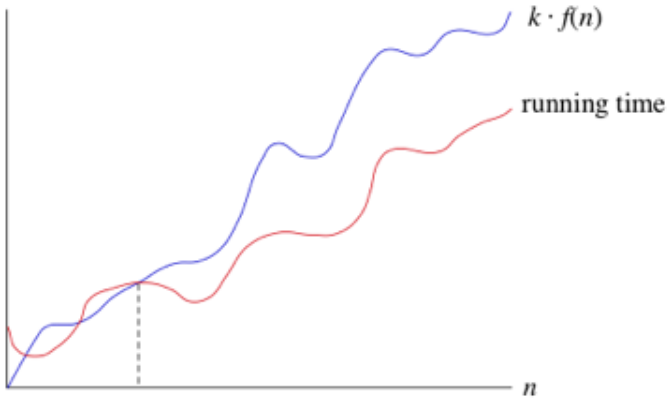
- Big-O, commonly written as **O**, is an Asymptotic Notation for the worst case, or ceiling of growth for a given function. It provides us with an ***asymptotic upper bound*** for the growth rate of runtime of an algorithm. Say f(n) is your algorithm runtime, and g(n) is an arbitrary time complexity you are trying to relate to your algorithm. f(n) is O(g(n)), if for some real constants c (c > 0) and n_0, f(n) <= c g(n) for every input size n (n > n_0).

Example 2

- f(n) **=** 3*****n**^**2 g(n) **=** n
- Is f(n) O(g(n))? Is 3 * n^2 O(n)? Let's look at the definition of Big-O.
- **3 *** n**^**2 **<=** c ***** n
- Is there some pair of constants c, n_0 that satisfies this for all n > $_0$? No, there isn't. f(n) is NOT O(g(n)).

Summarizing Big-Oh

- It would be convenient to have a form of asymptotic notation that means "the running time grows at most this much, but it could grow more slowly." We use "big-O" notation for just such occasions.
- If a running time is O(f(n)), then for large enough n, the running time is at most $k\,f(n)$ for some constant k. Here's how to think of a running time that is O(f(n)):



- We say that the running time is "big-O of f(n)" or just "O of f(n)." We use big-O notation for **asymptotic upper bounds**, since it bounds the growth of the running time from above for large enough input sizes.

Formal Definition

Big - Omega (Ω) LOWER BOUNDING function

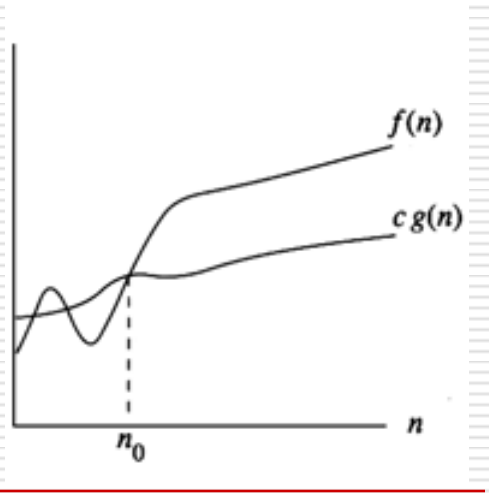
Big - Omega (Ω) LOWER BOUNDING function

- Big - Omega notation is used to define the **lower bound** of an algorithm in terms of Time Complexity. That means Big - Omega notation always indicates the **minimum time required** by an algorithm for all input values. That means Big - Omega notation describes the best case of an algorithm time complexity.

Formal Definition

- A function f(n) is said to be in **Ω**(g(n)) , denoted f(n) ∈ **Ω**(g(n)), if f(n) is bounded below by some positive constant multiple of g(n) for all large n, i.e., if there exist some positive constant c and some nonnegative integer n_0 such that

$$f(n) >= cg(n) \text{ for all } n>=n_0$$



Big - Omega (Ω)

Big - Omega notation is used to define the **lower bound** of an algorithm in terms of Time Complexity.

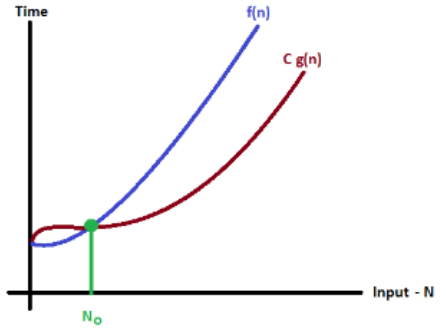
That means Big - Omega notation always indicates the minimum time required by an algorithm for all input values. That means Big - Omega notation describes the best case of an algorithm time complexity.

Big - Omega Notation can be defined as follows...

Consider function f(n) the time complexity of an algorithm and g(n) is the most significant term. If f(n) >= C x g(n) for all n >= n_0, C > 0 and n_0 >= 1. Then we can represent f(n) as Ω(g(n)).

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

Consider the following graph drawn for the values of f(n) and C g(n) for input (n) value on X-Axis and time required is on Y-Axis



In above graph after a particular input value n_0, always C x g(n) is less than f(n) which indicates the algorithm's lower bound.

Example

- Consider the following f(n) and g(n)...
  f(n) = 3n + 2
  g(n) = n
  If we want to represent **f(n)** as **Ω(g(n))** then it must satisfy **f(n) >= C g(n)** for all values of **C > 0** and **n_0>= 1**

Example

- Consider the following f(n) and g(n)...
  f(n) = 3n + 2
  g(n) = n
  If we want to represent **f(n)** as **Ω(g(n))** then it must satisfy **f(n) >= C g(n)** for all values of **C > 0** and **n_0>= 1**

- f(n) >= C g(n)
  ⇒3n + 2 >= C n
  Above condition is always TRUE for all values of
  C = 1 and **n >= 1**.

- By using Big - Omega notation we can represent the time complexity as follows...
  3n + 2 = Ω(n)

Example

- Consider the following f(n) and g(n)...
  f(n) = $n^3 + 4n^2$
  g(n) = n^2

If we want to represent **f(n)** as **Ω(g(n))** then it must satisfy

f(n) >= C g(n) for all values of **C > 0** and **n_0>= 1**

Example

- Consider the following f(n) and g(n)...
  f(n) = $n^3 + 4n^2$
  g(n) = n^2

If we want to represent **f(n)** as **Ω(g(n))** then it must satisfy **f(n) >= C g(n)** for all values of **C > 0** and $\mathbf{n_0 >= 1}$

- f(n) >= C g(n)
  $\Rightarrow n^3 + 4n^2 >= C\ n^2$
  Above condition is always TRUE for all values of
  C = 1 and **n >= 1**.

- By using Big - Omega notation we can represent the time complexity as follows...
  $n^3 + 4n^2 = \mathbf{\Omega}(n^2)$

Big-Oh and Big-Omega

- Big-Oh
- Think of it this way. Suppose you have 10 rupees in your pocket. You go up to your friend and say, "I have an amount of money in my pocket, and I guarantee that **it's no more** than one thousand rupees." Your statement is absolutely true, though not terribly precise.

- Big-Omega
- For example, if you really do have a one thousand rupees in your pocket, you can truthfully say "I have an amount of money in my pocket, and it's **at least** 10 rupees." That is correct, but certainly not very precise.

Formal Definition

Big - Theta (Θ) ORDER BOUNDING function

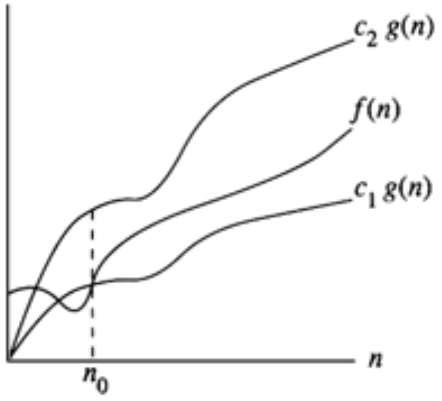
Big - Theta (Θ) ORDER BOUNDING function

- Big - Theta notation is used to define the **average bound** of an algorithm in terms of Time Complexity. That means Big - Theta notation always indicates the average time required by an algorithm for all input values. That means Big - Theta notation describes the average case of an algorithm time complexity.

Formal Definition

- A function f(n) is said to be in **Θ**(g(n)) , denoted f(n) ∈ **Θ**(g(n)), if f(n) is bounded above and below by some positive constant multiple of g(n) for all large n, i.e., if there exist some positive constant c_1 and c_2 and some nonnegative integer n_0 such that

$$c_1 g(n) <= f(n) <= c_2 g(n) \text{ for all } n >= n_0$$



Example

- Consider the following f(n) and g(n)...
  f(n) = 3n + 2
  g(n) = n
  If we want to represent **f(n)** as **Θ(g(n))** then it must satisfy **C_1 g(n) <= f(n) <= C_2 g(n)** for all values of **C_1, C_2 > 0** and **n_0>= 1**

Example

- Consider the following f(n) and g(n)...
  f(n) = 3n + 2
  g(n) = n
  If we want to represent **f(n)** as **Θ(g(n))** then it must satisfy **C_2 g(n) <= f(n) <= C_1 g(n)** for all values of **C_1, C_2 > 0** and **n_0>= 1**
- C_1 g(n) <= f(n) <= C_2 g(n) ⇒
  C_1 n <= 3n + 2 <= C_2 n

  Above condition is always TRUE for all values of **C_1 = 1, C_2 = 5** and **n >= 1**.
  By using Big - Theta notation we can represent the time complexity as follows...
  3n + 2 = Θ(n)

Example

- Consider the following f(n) and g(n)...

  $f(n) = 10n^3+5$

  $g(n) = n^3$

If we want to represent **f(n)** as **Θ(g(n))** then it must satisfy **C_2 g(n) <= f(n) <= C_1 g(n)** for all values of **C_1, C_2 > 0** and **n_0>= 1**

Example

- Consider the following f(n) and g(n)...
  $f(n) = 10n^3+5$
  $g(n) = n^3$

If we want to represent **f(n)** as **Θ(g(n))** then it must satisfy **C_1 g(n) <= f(n) <= C_2 g(n)** for all values of **C_1, C_2 > 0** and **n_0>= 1**

- C_1 g(n) <= f(n) <= C_2 g(n) ⇒
  C_1 **n^3** <= **$10n^3+5$** <= C_2 **n^3**

  Above condition is always TRUE for all values of
  C_1 = 10, C_2 = 11 and **n >= 2**.

By using Big - Theta notation we can represent the time complexity as follows...
$10n^3+5 = Θ(n^3)$

Summarizing:
Formal Definitions of Asymptotic Notations

Big - Oh (O) UPPER BOUNDING function

Big - Omega (Ω) LOWER BOUNDING function

Big - Theta (Θ) ORDER BOUNDING function

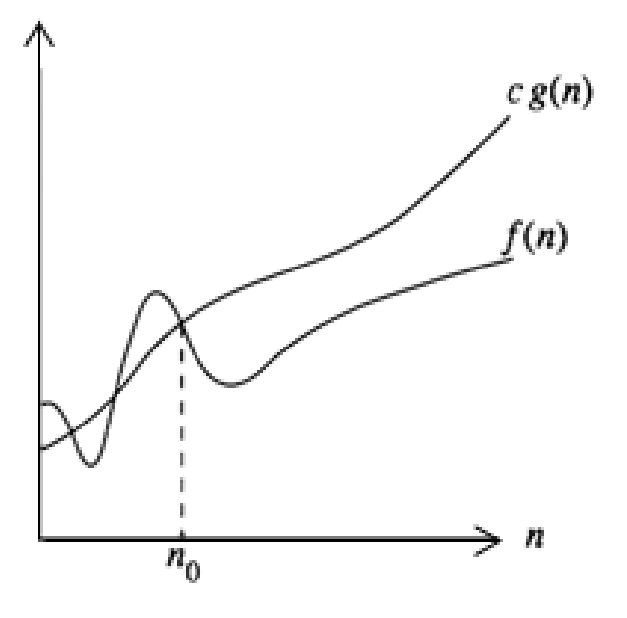
Big - Oh (O) UPPER BOUNDING function

- Big - Oh notation is used to define the **upper bound** of an algorithm in terms of Time Complexity. That means Big - Oh notation always indicates the maximum time required by an algorithm for all input values. That means Big - Oh notation describes the worst case of an algorithm time complexity.

Formal Definition

- A function f(n) is said to be in O(g(n)) , denoted **f(n) ∈ O(g(n)) (or f(n) = O(g(n)))**, if f(n) is bounded above by some constant multiple of g(n) for all large n, i.e., if there exist some positive constant c and some nonnegative integer n_0 such that

$$f(n) <= cg(n) \text{ for all } n>=n_0$$



n is Size of program's input.

t(n) Any real world function. Example: - Running time of a machine.

g(n) Another function that we want to use as an upper-bound. Not a real world function but preferably simple.

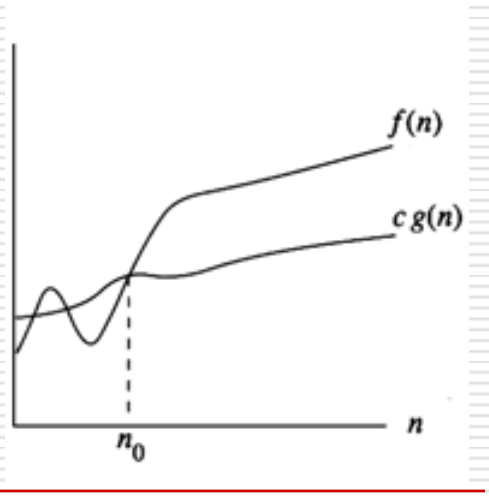
Big - Omega (Ω) LOWER BOUNDING function

- Big - Omega notation is used to define the **lower bound** of an algorithm in terms of Time Complexity. That means Big - Omega notation always indicates the **minimum time required** by an algorithm for all input values. That means Big - Omega notation describes the best case of an algorithm time complexity.

Formal Definition

- A function f(n) is said to be in **Ω**(g(n)) , denoted f(n) ∈ **Ω**(g(n)), if f(n) is bounded below by some positive constant multiple of g(n) for all large n, i.e., if there exist some positive constant c and some nonnegative integer n_0 such that

$$f(n) >= cg(n) \text{ for all } n>=n_0$$



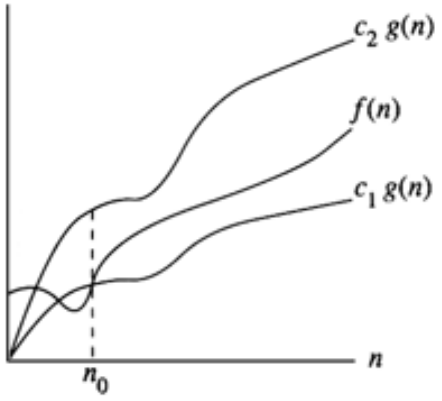
Big - Theta (Θ) ORDER BOUNDING function

- Big - Theta notation is used to define the **average bound** of an algorithm in terms of Time Complexity. That means Big - Theta notation always indicates the average time required by an algorithm for all input values. That means Big - Theta notation describes the average case of an algorithm time complexity.

Formal Definition

- A function f(n) is said to be in **Θ**(g(n)) , denoted f(n) ∈ **Θ**(g(n)), if f(n) is bounded above and below by some positive constant multiple of g(n) for all large n, i.e., if there exist some positive constant c_1 and c_2 and some nonnegative integer n_0 such that

$$c_1 g(n) <= f(n) <= c_2 g(n) \text{ for all } n >= n_0$$



Question

Let us consider the problem of finding the minimum of an array x[1..n]. The input size is n and the corresponding algorithm can be described as follows:

```
minimum(x[1..n])
1:      m ← x[1]
2:      for i ← 2, n do
3:          if m > x[i] then
4:                  m ← x[i]
5:          endif
6:      endfor
7: return m
```

In the given algorithm above, how many times the fourth operation (m=x[i]) will be executed if the given array (x) is having elements in **ascending order** ?

Answer

Let us consider the problem of finding the minimum of an array x[1..n]. The input size is n and the corresponding algorithm can be described as follows:

```
minimum(x[1..n])
1:      m ← x[1]
2:      for i ← 2, n do
3:          if m > x[i] then
4:                  m ← x[i]
5:          endif
6:      endfor
7: return m
```

It will not get executed

Best Case (Lower Bound)

T(n) = 0

In the given algorithm above, how many times the fourth operation (m=x[i]) will be executed if the given array (x) is having elements in **ascending order** ?

Question

Let us consider the problem of finding the minimum of an array x[1..n]. The input size is n and the corresponding algorithm can be described as follows:

```
minimum(x[1..n])
1:      m ← x[1]
2:      for i ← 2, n do
3:          if m > x[i] then
4:                  m ← x[i]
5:          endif
6:      endfor
7: return m
```

In the given algorithm above, how many times the fourth operation (m=x[i]) will be executed if the given array (x) is having elements in **Descending order** ?

Answer

Let us consider the problem of finding the minimum of an array x[1..n]. The input size is n and the corresponding algorithm can be described as follows:

```
minimum(x[1..n])
1:    m ← x[1]
2:    for i ← 2, n do
3:        if m > x[i] then
4:            m ← x[i]
5:        endif
6:    endfor
7: return m
```

It will get executed (n-1) times

Worst Case (Upper bound)

T(n) = (n-1)

In the given algorithm above, how many times the fourth operation (m=x[i]) will be executed if the given array (x) is having elements in **Descending order** ?

Best and Worst case

```
minimum(x[1..n])
1:    m ← x[1]
2:    for i ← 2, n do
3:        if m > x[i] then
4:            m ← x[i]
5:        endif
6:    endfor
7: return m
```

Operation	Cost	Repetitions
1	1	1
2	$2n$	1
3	1	$n-1$
4	1	$\tau(n)$

Unlike in previous examples we cannot find a general expression for the running time. This happens because the running time of the fourth operation does not depend only on the input size but also on the properties of the array (mainly on the position where the minimal value appears for the first time).

If the minimum is on the first position then the assignment 4 is not at all executed, τ (n) = 0. This is the **best case** which could appear.

If, on the other hand, the array is decreasingly sorted the assignment 4 is executed at each iteration. Thus τ(n) = n − 1. This is the **worst case**. Taking into consideration the **best** and the **worst case** we can establish a **lower** and an **upper bound** for the running time:

3n ≤ T(n) ≤ 4n−1. It is easy to see that both bounds depend linearly on the input size.

Example

```
int a;
a = 5
a++;
```

Simple statements

Fragment 1

O(1)

Example

```
int a;
a = 5
a++;
```

Simple statements

Fragment 1

O(1)

```
for(i=0;i<n;i++)
{
  // simple statements
}
```

Single loop

Fragment 2

O(n)

Example

```
int a;
a = 5
a++;
```

Simple statements

Fragment 1

$O(1)$

```
for(i=0;i<n;i++)
{
  // simple statements
}
```

Single loop

Fragment 2

$O(n)$

```
for(i=0;i<n;i++
{
  for(j=0;j<n;j+
  {
    // simple state
  }
}
```

Nested loop

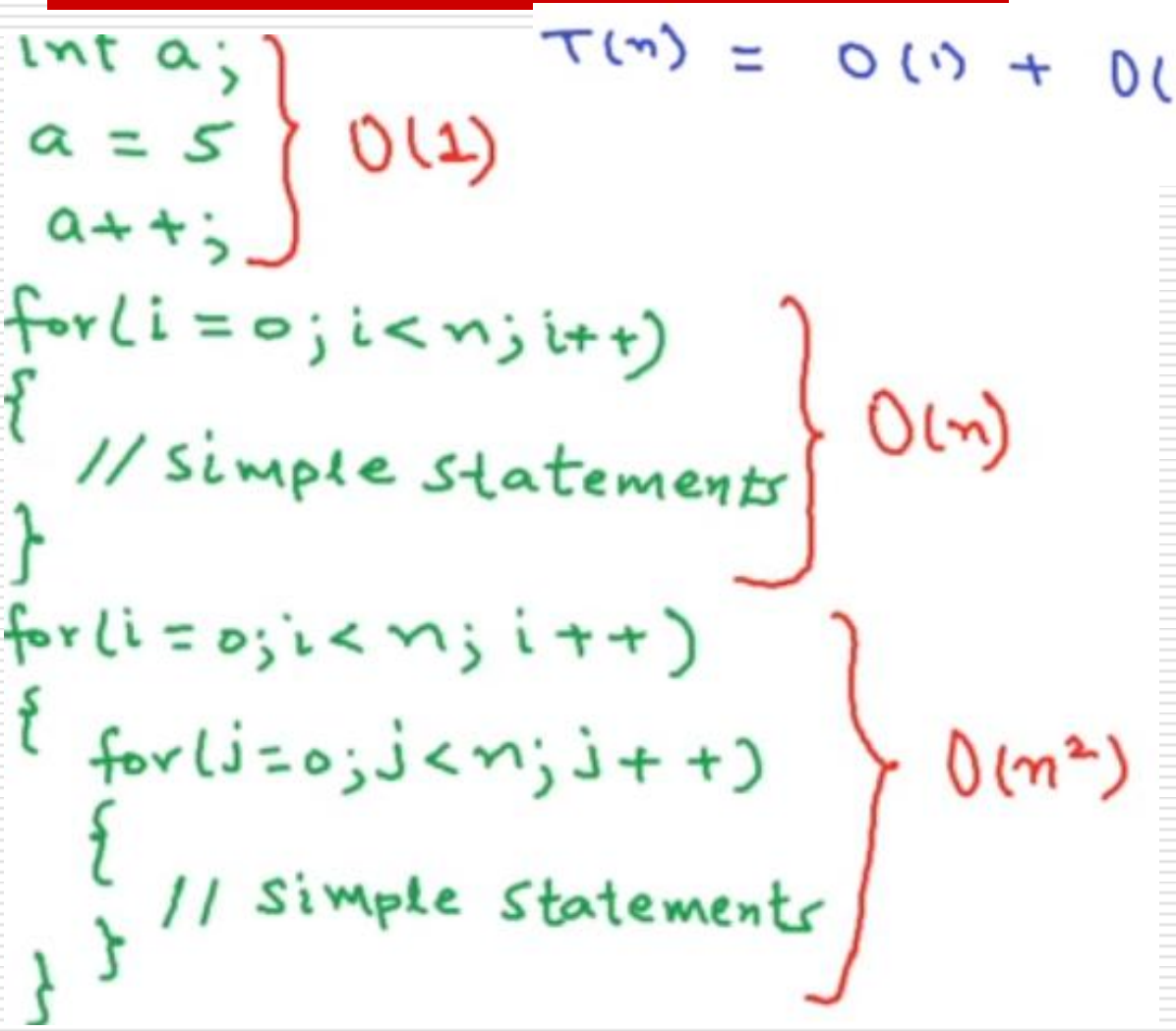
Fragment 3

$O(n^2)$

Question: What is the running time of this code ?

```
int a;
a = 5
a++;
for(i=0;i<n;i++)
{
  // simple statements
}
for(i=0;i<n;i++)
{
  for(j=0;j<n;j++)
  {
    // simple statements
  }
}
```

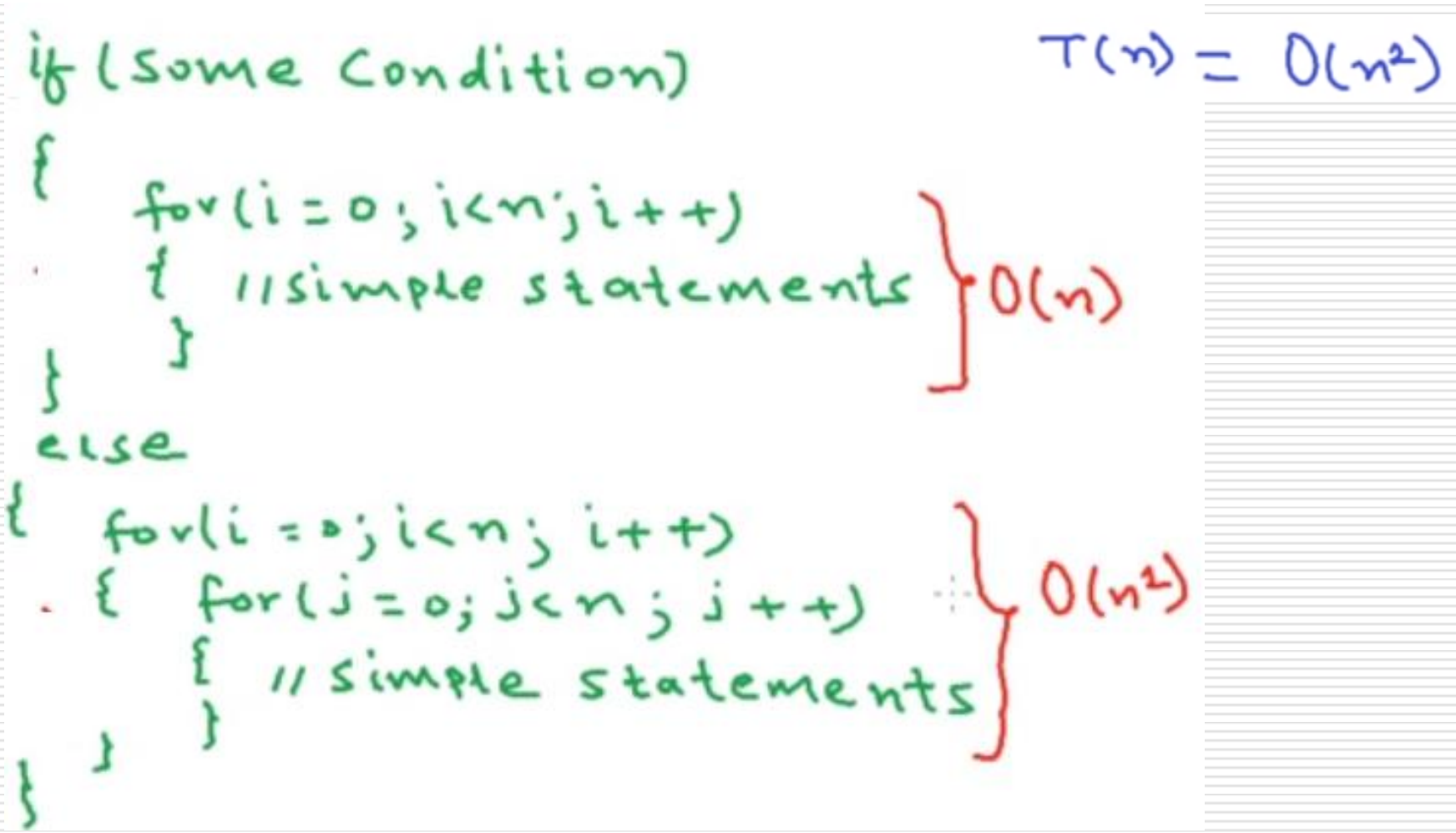
Answer


int a;
a = 5
a++;
O(1)
for(i=0; i<n; i++)
{
// simple statements
}
O(n)
for(i=0; i<n; i++)
{
for(j=0; j<n; j++)
{
// simple statements
}
}
O(n²)
T(n) = O(1) + O(n) + O(n²) = O(n²)

Question: What is the running time of this code ?

```
if (some condition)
{
    for(i = 0; i<n; i++)
    { //simple statements
    }
}
else
{   for(i = 0; i<n; i++)
    {   for(j = 0; j<n; j++)
        { // simple statements
        }
    }
}
```

Answer


if (some Condition)
{
for(i=0;i<n;i++)
{ //simple statements
}
}
else
{ for(i=0;i<n;i++)
{ for(j=0;j<n;j++)
{ //simple statements
}
}
}
O(n)
O(n²)
T(n) = O(n²)

Question

Use the informal definitions of O, Θ, and Ω to determine whether the following assertions are true or false.

a. $n(n+1)/2 \in O(n^3)$

b. $n(n+1)/2 \in O(n^2)$

c. $n(n+1)/2 \in \Theta(n^3)$

d. $n(n+1)/2 \in \Omega(n)$

Question

Use the informal definitions of O, Θ, and Ω to determine whether the following assertions are true or false.

a. $n(n+1)/2 \in O(n^3)$

b. $n(n+1)/2 \in O(n^2)$

c. $n(n+1)/2 \in \Theta(n^3)$

d. $n(n+1)/2 \in \Omega(n)$

Answer

$n(n+1)/2 \approx n^2/2$ is quadratic. Therefore

a. $n(n+1)/2 \in O(n^3)$ is true.

b. $n(n+1)/2 \in O(n^2)$ is true.

c. $n(n+1)/2 \in \Theta(n^3)$ is false.

d. $n(n+1)/2 \in \Omega(n)$ is true.

Question

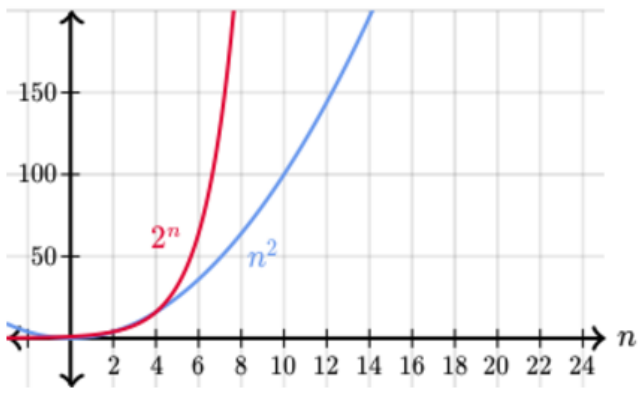
- For the functions, n^k and $c^n c$, what is the asymptotic relationship between these functions? Assume that k >= 1 and c > 1 are constants.
- n^k is $O(c^n)$
- n^k *is* $\Omega(c^n)$
- n^k is $\Theta(c^n)$

Answer

For the functions, n^k and $c^n c$, what is the asymptotic relationship between these functions? Assume that k >= 1 and c > 1 are constants.

- ***n^k* is $O(c^n)$**
- *n^k is $\Omega(c^n)$*
- n^k is $\Theta(c^n)$

n^k is a polynomial function, and c^n is a exponential function. We know that polynomials always grow more slowly than exponentials. We could prove that with a graph, but we have to make sure we look at it for big values of n, because the early behavior could be misleading. Here's a graph that compares the two functions (with k=2 and c=2), where we can clearly see the difference in growth:



Question

For the functions, 8^n and 4^n, what is the asymptotic relationship between these functions?

A. 8^n is $O(4^n)$

B. 8^n is $\Omega(4^n)$

C. 8^n is $\Theta(4^n)$

Question

For the functions, 8^n and 4^n, what is the asymptotic relationship between these functions?

A	8^n is $O(4^n)$	✗	8^n is $O(4^n)$
B	8^n is $\Omega(4^n)$	✓	8^n is $\Omega(4^n)$
C	8^n is $\Theta(4^n)$	✗	8^n is $\Theta(4^n)$

Question

Consider the following three claims

1. $(n + k)^m = \Theta(n^m)$, where k and m are constants
2. $2^{n+1} = O(2^n)$
3. $2^{2n+1} = O(2^n)$

Which of these claims are correct ?

(A) 1 and 2 **(B)** 1 and 3
(C) 2 and 3 **(D)** 1, 2, and 3

Question

Consider the following three claims

1. $(n + k)^m = \Theta(n^m)$, where k and m are constants
2. $2^{n+1} = O(2^n)$
3. $2^{2n+1} = O(2^n)$

Which of these claims are correct ?

(A) 1 and 2 **(B)** 1 and 3
(C) 2 and 3 **(D)** 1, 2, and 3

Answer: **(A)**

Explanation: $(n + k)^m$ and $\Theta(n^m)$ are asymptotically same as theta notation can always be written by taking the leading order term in a polynomial expression.

2^{n+1} and $O(2^n)$ are also asymptotically same as 2^{n+1} can be written as $2 * 2^n$ and constant multiplication/addition doesn't matter.

2^{2n+1} and $O(2^n)$ are not same as constant is in power.

Question

What is the asymptotic relationship between the functions $n^3 \lg n$ and $3n \log_8 n$?

A $n^3 \lg n$ is $O(3n \log_8 n)$	⊖ $n^3 \lg n$ is $O(3n \log_8 n)$
B $n^3 \lg n$ is $\Omega(3n \log_8 n)$	✓ $n^3 \lg n$ is $\Omega(3n \log_8 n)$
C $n^3 \lg n$ is $\Theta(3n \log_8 n)$	⊖ $n^3 \lg n$ is $\Theta(3n \log_8 n)$

Question

For the functions, $\lg n^{\lg 17}$ vs. $\lg 17^{\lg n}$, what is the asymptotic relationship between these functions?

Choose all answers that apply:

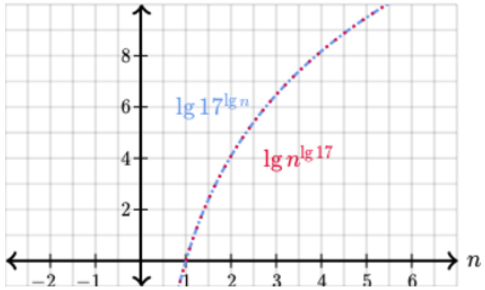
- (A) $\lg n^{\lg 17}$ is $O(\lg 17^{\lg n})$
- (B) $\lg n^{\lg 17}$ is $\Omega(\lg 17^{\lg n})$
- (C) $\lg n^{\lg 17}$ is $\Theta(\lg 17^{\lg n})$

Answer

In conclusion, all of the statements are true.

- $\lg n^{\lg 17}$ is $O(\lg 17^{\lg n})$
- $\lg n^{\lg 17}$ is $\Omega(\lg 17^{\lg n})$
- $\lg n^{\lg 17}$ is $\Theta(\lg 17^{\lg n})$

Both $\lg n^{\lg 17}$ vs. $\lg 17^{\lg n}$ are functions with logarithmic growth, and the same base. They differ in what they take the logarithm of: $n^{\lg 17}$ versus $17^{\lg n}$. Here's a graph of the two functions:



Notice something? It's the same graph! They're actually exactly equivalent functions, because of a particular property of logarithms:

$$\lg a^b = b \lg a$$

Let's re-write both of the original functions using that property:

original	becomes
$\lg\left(n^{(\lg 17)}\right)$	$\lg(17) \cdot \lg n$
$\lg\left(17^{\lg n}\right)$	$\lg(n) \cdot \lg(17)$

Useful Property Involving the Asymptotic Notations

THEOREM If $t_1(n) \in O(g_1(n))$ and $t_2(n) \in O(g_2(n))$, then

$$t_1(n) + t_2(n) \in O(\max\{g_1(n), g_2(n)\}).$$

PROOF The proof extends to orders of growth the following simple fact about four arbitrary real numbers a_1, b_1, a_2, b_2: if $a_1 \le b_1$ and $a_2 \le b_2$, then $a_1 + a_2 \le 2\max\{b_1, b_2\}$.

Since $t_1(n) \in O(g_1(n))$, there exist some positive constant c_1 and some non-negative integer n_1 such that

$$t_1(n) \le c_1 g_1(n) \quad \text{for all } n \ge n_1.$$

Similarly, since $t_2(n) \in O(g_2(n))$,

$$t_2(n) \le c_2 g_2(n) \quad \text{for all } n \ge n_2.$$

Let us denote $c_3 = \max\{c_1, c_2\}$ and consider $n \ge \max\{n_1, n_2\}$ so that we can use both inequalities. Adding them yields the following:

$$\begin{aligned} t_1(n) + t_2(n) &\le c_1 g_1(n) + c_2 g_2(n) \\ &\le c_3 g_1(n) + c_3 g_2(n) = c_3[g_1(n) + g_2(n)] \\ &\le c_3 2\max\{g_1(n), g_2(n)\}. \end{aligned}$$

Hence, $t_1(n) + t_2(n) \in O(\max\{g_1(n), g_2(n)\})$, with the constants c and n_0 required by the O definition being $2c_3 = 2\max\{c_1, c_2\}$ and $\max\{n_1, n_2\}$, respectively. ■

Useful Property Involving the Asymptotic Notations (Contd…)

So what does this property imply for an algorithm that comprises two consecutively executed parts? It implies that the algorithm's overall efficiency is determined by the part with a higher order of growth, i.e., its least efficient part:

$$\left.\begin{array}{|l} t_1(n) \in O(g_1(n)) \\ \hline t_2(n) \in O(g_2(n)) \end{array}\right\} \quad t_1(n) + t_2(n) \in O(\max\{g_1(n), g_2(n)\}).$$

For example, we can check whether an array has equal elements by the following two-part algorithm: first, sort the array by applying some known sorting algorithm; second, scan the sorted array to check its consecutive elements for equality. If, for example, a sorting algorithm used in the first part makes no more than $\frac{1}{2}n(n-1)$ comparisons (and hence is in $O(n^2)$) while the second part makes no more than $n-1$ comparisons (and hence is in $O(n)$), the efficiency of the entire algorithm will be in $O(\max\{n^2, n\}) = O(n^2)$.

Using Limits for Comparing Orders of Growth

- Though the formal definitions of *O*, *Ω* and *Θ* are indispensable for proving their abstract properties, they are rarely used for comparing the orders of growth of two specific functions. A much more convenient method for doing so is based on computing the limit of the ratio of two functions in question. Three principal cases may arise:

$$\lim_{n\to\infty} \frac{t(n)}{g(n)} = \begin{cases} 0 & \text{implies that } t(n) \text{ has a smaller order of growth than } g(n), \\ c & \text{implies that } t(n) \text{ has the same order of growth as } g(n), \\ \infty & \text{implies that } t(n) \text{ has a larger order of growth than } g(n).^3 \end{cases}$$

Note that the first two cases mean that $t(n) \in O(g(n))$, the last two mean that $t(n) \in \Omega(g(n))$, and the second case means that $t(n) \in \Theta(g(n))$.

Example

EXAMPLE 1 Compare the orders of growth of $\frac{1}{2}n(n-1)$ and n^2. (This is one of the examples we used at the beginning of this section to illustrate the definitions.)

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

Since the limit is equal to a positive constant, the functions have the same order of growth or, symbolically, $\frac{1}{2}n(n-1) \in \Theta(n^2)$. ■

Problem

For each of the following functions, indicate the class $\Theta(g(n))$ the function belongs to. (Use the simplest $g(n)$ possible in your answers.) Prove your assertions.

a. $(n^2+1)^{10}$

b. $\sqrt{10n^2+7n+3}$

a. Informally, $(n^2+1)^{10} \approx (n^2)^{10} = n^{20} \in \Theta(n^{20})$ Formally,

$$\lim_{n\to\infty} \frac{(n^2+1)^{10}}{n^{20}} = \lim_{n\to\infty} \frac{(n^2+1)^{10}}{(n^2)^{10}} = \lim_{n\to\infty} \left(\frac{n^2+1}{n^2}\right)^{10} == \lim_{n\to\infty} \left(1+\frac{1}{n^2}\right)^{10} = 1.$$

Hence $(n^2+1)^{10} \in \Theta(n^{20})$.

b. Informally, $\sqrt{10n^2+7n+3} \approx \sqrt{10n^2} = \sqrt{10}n \in \Theta(n)$. Formally,

$$\lim_{n\to\infty} \frac{\sqrt{10n^2+7n+3}}{n} = \lim_{n\to\infty} \sqrt{\frac{10n^2+7n+3}{n^2}} = \lim_{n\to\infty} \sqrt{10+\frac{7}{n}+\frac{3}{n^2}} = \sqrt{10}.$$

Hence $\sqrt{10n^2+7n+3} \in \Theta(n)$.

Thanks for Listening

Quiz

- Match each function with an equivalent function, in terms of their Θ. Only match a function if f(n) = Θ($g(n)$).

$f(n)$	$g(n)$
$n + 30$	$n^2 + 3n$
$n^2 + 2n - 10$	$\log_2 2x$
$n^3 * 3n$	$3n - 1$
$\log_2 x$	n^4

$f(x)$	Simplified	$g(x)$	Simplified
$n^2 + 2n - 10$	n^2	$n^2 + 3n$	n^2
$n^3 * 3n$	n^4	n^4	n^4
$n + 30$	n	$3n - 1$	n
$\log_2 x$	$\log_2 x$	$\log_2 2x$	$\log_2 x$

Given N distinct integers, how many triples sum to exactly zero?

30 -40 -20 -10 40 0 10 5

	a[i]	a[j]	a[k]	sum
1	30	-40	10	0
2	30	-20	-10	0
3	-40	40	0	0
4	-10	0	10	0

Brute-force algorithm

```
int count(int[] a)
{
   int N = a.length;
   int count = 0;
   for (int i = 0; i < N; i++)
      for (int j = i+1; j < N; j++)
         for (int k = j+1; k < N; k++)
            if (a[i] + a[j] + a[k] == 0)
               count++;
   return count;
}
```

```
int count = 0;
for (int i = 0; i < N; i++)
   if (a[i] == 0)
      count++;
```

operation	frequency
variable declaration	2
assignment statement	2
less than compare	N + 1
equal to compare	N
array access	N
increment	N to 2 N

Program Performance

- Program performance is the amount of computer memory and time needed to run a program.

- How is it determined?
  1. Analytically
     - performance analysis
  2. Experimentally
     - performance measurement

Criteria for Measurement

- Space
  - amount of memory program occupies
  - usually measured in bytes, KB or MB
- Time
  - execution time
  - usually measured by the number of executions

Components of Program Space

- Program space = Instruction space + data space + stack space

- The **instruction space** is dependent on several factors.
  - the compiler that generates the machine code
  - the compiler options that were set at compilation time
  - the target computer

Components of Program Space

- Data space
  - very much dependent on the computer architecture and compiler
  - The magnitude of the data that a program works with is another factor

char	1	float	4
short	2	double	8
int	2	long double	10
long	4	pointer	2

Unit: bytes

Components of Program Space

- Environment Stack Space
  - Every time a function is called, the following data are saved on the stack.
    1. the return address
    2. the values of all local variables and value formal parameters
    3. the binding of all reference and const reference parameters
  - What is the impact of recursive function calls on the environment stack space?

```
int mean(int a[], size_t n)
{
int sum = 0; // 1 step * 1
for (int i = 0; i < n; i++) // 1 step * (N+1)
sum += a[i]; // 1 step * N
return sum; // 1 step * 1
}
```

- Add up the steps: **1 + (N+1) + N + 1**
- Reduce: **2N + 3**

Quiz: Running time of binary search

32 teams qualified for the 2014 World Cup. If the names of the teams were arranged in sorted order (an array), how many items in the array would binary search have to examine to find the location of a particular team in the array, in the worst case?

A. At most, 1

B. At most, 16

C. At most, 6

D. At most, 32

Quiz: Running time of binary search

What is lg(32), the base-2 logarithm of 32?

A. 5

B. 32

C. 16

D. 1

Quiz: Running time of binary search

You have an array containing the prime numbers from 2 to 311 in sorted order: [2, 3, 5, 7, 11, 13, ..., 307, 311]. There are 64 items in the array. About how many items of the array would binary search have to examine before concluding that 52 is not in the array, and therefore not prime?

A. 32

B. 11

C. 64

D. 128

E. 22

F. 1

G. 7

- Suppose you have the following sorted list [3, 5, 6, 8, 11, 12, 14, 15, 17, 18] and are using the recursive binary search algorithm. Which group of numbers correctly shows the sequence of comparisons used to find the key 8.
- (A) 11, 5, 6, 8
  (B) **12, 6, 11, 8**
  (C) 3, 5, 6, 8
  (D) 18, 12, 6, 8

- Suppose you have the following sorted list [3, 5, 6, 8, 11, 12, 14, 15, 17, 18] and are using the recursive binary search algorithm. Which group of numbers correctly shows the sequence of comparisons used to search for the key 16?(A) 11, 14, 17
  (B) 18, 17, 15
  (C) 14, 17, 15
  (D) **12, 17, 15**

Pseudocode for (Euclid's Algorithm), **gcd(m, n)**

ALGORITHM Euclid(m, n)

// Computes gcd(m, n) by Euclid's algorithm

// Input: Two nonnegative, not-both-zero integers m and n

//Output: Greatest common divisor of m and n

while n ≠ 0 **do**

    r = m mod n

    m = n

    n = r

return m

gcd(m, n) = gcd(n, m mod n)

gcd(24, 60) = gcd(60, 24 mod 60

What is the minimum number of divisions among all inputs 1 ≤ m, n ≤ 10 ?

What happens if m < n ?

Second try: Consecutive Integer Checking, **gcd(m, n)**

- **Step 1:** Assign the value of min{m, n} to q.
- **Step 2:** Divide m by q. If the remainder is 0, go to Step 3; otherwise, go to Step 4.
- **Step 3:** Divide n by q. If the remainder is 0, return the value of q as the answer and stop; otherwise, proceed to Step 4.
- **Step 4:** Decrease the value of q by 1. Go to Step 2.

Does it work when m or n is 0 ?

Which one is faster, Euclid's or this one ?

Third try: Middle-school, **gcd(m, n)**

- **Step 1**: Find prime factors of m.
- **Step 2**: Find prime factors of n.
- **Step 3**: Identify all common prime factors of m and n (if p is a prime factor occurring p_m and p_n times in m and n, it should be repeated $\min\{p_m,p_n\}$ times)
- **Step 4**: Compute product of all common factors and return product as the answer.

Is it legitimate algorithm?

Test Your Knowledge

- Find Time Complexity of the following Algorithm

```
product(a[1..m, 1..n], b[1..n, 1..p])
1:    for i = 1,m do
2:        for j = 1,p do
3:            c[i, j] ← 0
4:            for k = 1,n do
5:                c[i, j] ← c[i, j] + a[i, k] * b[k, j]
6:            endfor
7:        endfor
8:    endfor
9: return c[1..m, 1..p]
```

- Thus the running time can be computed by using the following costs table:

Operation	Cost	Repetitions
1	$2(m+1)$	1
2	$2(p+1)$	m
3	1	$m \cdot p$
4	$2(n+1)$	$m \cdot p$
5	2	$m \cdot p \cdot n$

- The overall cost will be: T(m, n, p) = 4mnp + 5mp + 4m + 2.