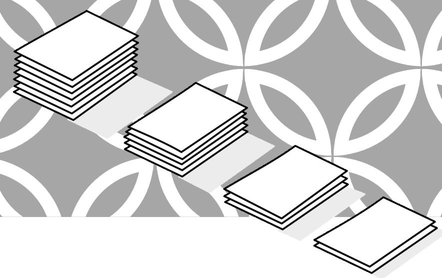


CSCI 3100

COMPARISON BASED SORTING AND ASYMPTOTIC ANALYSIS

REVIEW & OVERVIEW

Review from last class:

- *Search algorithms: min, min+max, second largest*
  - *How many comparisons?*
  - *Is this the best we can do?*
- *Insertion sort*

Analysis of Insertion Sort

- Best case
- Worst case
- Average case

Asymptotic Analysis

Insertion Sort

```
INSERTION-SORT(A)
for j = 2 to A.length
    key = A[j]
    // Insert A[j] into the sorted
          sequence A[1..j − 1].
    i = j − 1
    while i > 0 and A[i] > key
        A[i + 1] = A[i]
        i = i − 1
    A[i + 1] = key
```

INSERTION-SORT(A)	*cost*	*times*
for $j = 2$ **to** $A.length$	c_1	n
$key = A[j]$	c_2	$n-1$
// Insert $A[j]$ into the sorted sequence $A[1..j-1]$.	0	$n-1$
$i = j - 1$	c_4	$n-1$
while $i > 0$ and $A[i] > key$	c_5	$\sum_{j=2}^{n} t_j$
$A[i+1] = A[i]$	c_6	$\sum_{j=2}^{n} (t_j - 1)$
$i = i - 1$	c_7	$\sum_{j=2}^{n} (t_j - 1)$
$A[i+1] = key$	c_8	$n-1$

$$\sum_{j=2}^{n} j = \frac{n(n+1)}{2} - 1$$

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

$$T(n) = c_1 n + c_2(n-1) + c_4(n-1) + c_5 \sum_{j=2}^{n} t_j + c_6 \sum_{j=2}^{n} (t_j - 1) + c_7 \sum_{j=2}^{n} (t_j - 1) + c_8(n-1) .$$

BEST, WORST, AND AVERAGE CASES

What is the best case running time of this algorithm?

What is the worst case running time of this algorithm?

What is the average case running time of this algorithm?

ASYMPTOTIC ANALYSIS

rough measure that characterizes **how fast functions grow** with respect to **n**

Big O notation: asymptotic "less than" or "at most":

f(n)=O(g(n)) implies: f(n) "$\leq$" g(n)

Ω notation: asymptotic "greater than" or "at least":

f(n)= Ω (g(n)) implies: f(n) "$\geq$" g(n)

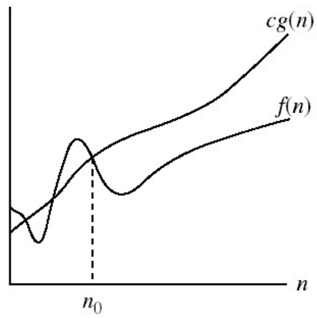
Θ notation: asymptotic "equality" or "exactly":

f(n)= Θ (g(n)) implies: f(n) "=" g(n)

6

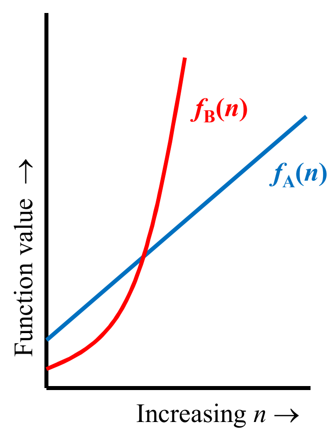
DEFINITION OF BIG O

$$O(g(n)) = \{f(n) : \text{there exist positive constants } c \text{ and } n_0 \text{ such that } 0 \le f(n) \le cg(n) \text{ for all } n \ge n_0\}.$$



$g(n)$ is an ***asymptotic upper bound*** for $f(n)$.

BIG-O NOTATION



We say

$f_A(n) = 7n+18$ is ***order n***, or **O (n)**
It is, at most, roughly *proportional* to *n.*

$f_B(n) = 3n^2+5n+4$ is ***order n²***, or **O(*n*²)**.

It is, at most, roughly proportional to n^2.

In general, any O(n^2) function is faster- growing than any O(n) function.

8

MORE EXAMPLES

$n^4 + 100n^2 + 10n + 50$ ➔ $O(n^4)$

$10n^3 + 2n^2$ ➔ $O(n^3)$

$n^3 - n^2$ ➔ $O(n^3)$

constants

10 is **O(1)**

1273 is **O(1)**

what is the rate of growth for ***INSERTION SORT*** tudied earlier (in Big O notation)?

Total Time = $aN^2 + bN + c$

If a, b, and c are constants then ***Total Time = O(?)***

NO UNIQUENESS

There is no unique set of values for n_0 and c in proving the asymptotic bounds

Prove that $100n + 5 = O(n^2)$

(i) $100n + 5 \leq 100n + n = 101n \leq 101n^2$ for all $n \geq 5$

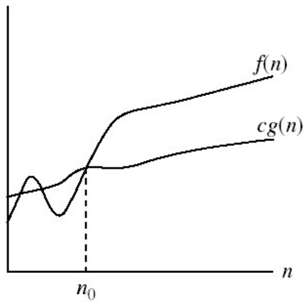
You may pick $n_0 = 5$ and $c = 101$ to complete the proof.

(ii) $100n + 5 \leq 100n + 5n = 105n \leq 105n^2$ for all $n \geq 1$

You may pick $n_0 = 1$ and $c = 105$ to complete the proof.

DEFINITION OF Ω

$$\Omega(g(n)) = \{f(n) : \text{there exist positive constants } c \text{ and } n_0 \text{ such that } 0 \le cg(n) \le f(n) \text{ for all } n \ge n_0\} .$$



$g(n)$ is an ***asymptotic lower bound*** for $f(n)$.

$\Omega(g(n))$ is the set of functions with larger or same order of growth as $g(n)$

EXAMPLES

- **$5n^2 = \Omega(n)$**

  $\exists$ c, n_0 such that: $0 \le cn \le 5n^2 \Rightarrow cn \le 5n^2 \Rightarrow c = 1$ and $n > n_0=1$
- **$100n + 5 \ne \Omega(n^2)$**

  $\exists$ c, n_0 such that: $0 \le cn^2 \le 100n + 5$

  since $100n + 5 \le 100n + 5n \quad \forall\, n \ge 1$

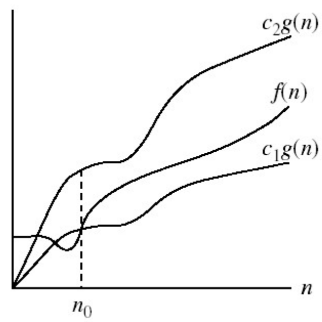
  $cn^2 \le 105n \Rightarrow n(cn - 105) \le 0$

  Since n is positive $\Rightarrow (cn - 105) \le 0 \Rightarrow n \le 105/c$

  $\Rightarrow$ contradiction: n cannot be smaller than a constant
- **$n = \Omega(2n)$, $n^3 = \Omega(n^2)$, $n = \Omega(\log n)$**

DEFINITION OF Θ

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



$\Theta(g(n))$ is the set of functions with the same order of growth as $g(n)$

$g(n)$ is an ***asymptotically tight bound*** for $f(n)$.

13

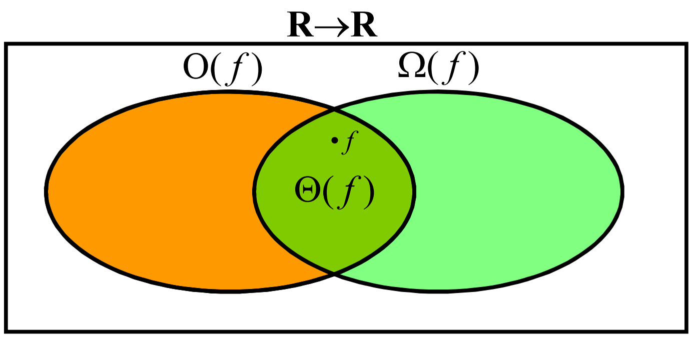
EXAMPLES

- $n^2/2 - n/2 = \Theta(n^2)$
  - $\frac{1}{2} n^2 - \frac{1}{2} n \le \frac{1}{2} n^2 \quad \forall n \ge 0 \quad \Rightarrow \quad c_2 = \frac{1}{2}$
  - $\frac{1}{4} n^2 \le \frac{1}{2} n^2 - \frac{1}{2} n \quad \forall n \ge 2 \quad \Rightarrow \quad c_1 = \frac{1}{4}$
- $n \ne \Theta(n^2)$: $c_1 n^2 \le n \le c_2 n^2$ $\quad \Rightarrow$ only holds for: $n \le 1/c_1$
- $6n^3 \ne \Theta(n^2)$: $c_1 n^2 \le 6n^3 \le c_2 n^2$

  $\Rightarrow$ only holds for: $n \le c_2 / 6$
- $n \ne \Theta(\log n)$: $c_1 \log n \le n \le c_2 \log n$

  $\Rightarrow c_2 \ge n/\log n, \forall n \ge n_0$ – impossible

RELATIONS BETWEEN DIFFERENT SETS



15

The Growth of Combinations of Functions

- 1
- $\log n$
- n
- $n \log n$
- n^2
- 2^n
- $n!$