CS 577 - Basics of Algorithm Analysis

Marc Renault

Department of Computer Sciences
University of Wisconsin – Madison

Spring 2023
TopHat Section 001 Join Code: 020205
TopHat Section 002 Join Code: 394523

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Algorithm Analysis

Algorithm Evaluation

- Sound
- Complete
- Resource requirements:
  - Time
  - Space
  - Other...

Computational Tractability

Defining Efficiency

Definition 1[1]

An algorithm is efficient if, when implemented, it runs quickly on real input instances.

Issues:

- Not concrete enough for meaning algorithm comparison.
- What is "quickly"?
- What are "real input instances"?

[1] Algorithm Design, p. 30

Defining Efficiency

Definition 2[2]

An algorithm is efficient if it achieves qualitatively better *worst-case* performance, at an analytical level,than *brute force* search.

[2] Algorithm Design, p. 32

Quantifying an Algorithm's Performance

Brute-force

1. Enumerate all possible solutions.
2. Check all possible solutions and keep the best one.

Worst-case

Considering all possible inputs, what is worst possible performance of the algorithm?

- Absolute guarantee on performance.
- Only needs one data point.

DEFINING EFFICIENCY

Definition 2[2]

An algorithm is efficient if it achieves qualitatively better *worst-case* performance, at an analytical level,than *brute force* search.

Issues:

- Still too vague for a good measure.
- What exactly is "qualitative"?

[2] Algorithm Design, p. 32

Stable Marriage Problem (SMP) (1962)[1,2,3]

Problem Definition

Given a set of n men, M, and an opposite set of n women, W. Each person has a preference ranking of the opposite set. Compute a stable matching between M and W. A matching is stable if it is (i) perfect, and (ii) there are no pairs (m, w) and (m', w') in the matching where m prefers w' and w' prefers m.

- A.k.a Stable Matching Problem.
- There are more complicated variations of the model.
- Used in the real world (e.g. matching doctors to hospitals).
- Nobel Prize in Economics in 2012 (Shapley and Roth).

[1] Algorithm Design, Ch 1.

[2] Algorithms, Ch 4.5

[3] `http://mathsite.math.berkeley.edu/smp/smp.html`

Analysis of SMP

Algorithm: Gale-Shapley Algorithm (1962)

Initially all $m \in M$ and $w \in W$ are free
while *there is a man m who is free and hasn't proposed to every woman* **do**
    Choose such a man m
    Let w be the highest-ranked woman in m's preference list to whom m has not yet proposed
    if *w is free* **then**
        (m, w) become engaged
    else w is currently engaged to m'
        if *w prefers m' to m* **then**
            m remains free
        else w prefers m to m'
            (m, w) become engaged
            m' becomes free
        end
    end
end
return *the set S of engaged pairs*

TopHat 1

How many brute-force possibilities when there are n men and n women?

Defining Efficiency

Definition 3[3]

An algorithm is efficient if it has a polynomial running time with respect to the input size.

Polynomial: $f(n) = c_d \cdot n^d + c_{d-1} \cdot n^{d-1} + \cdots + c_1 \cdot n + c_0$, where d and c_i are constants.

Well defined notion:

- Natural follow-up: what is the most efficient algorithm possible?
- Not perfect: n^{100} is polynomial, but $n^{1+0.02(\log n)}$ is not.

[3] Algorithm Design, p. 32

Analysis of SMP

Algorithm: Gale-Shapley Algorithm (1962)

Initially all $m \in M$ and $w \in W$ are free
while *there is a man m who is free and hasn't proposed to every woman* **do**
    Choose such a man m
    Let w be the highest-ranked woman in m's preference list to whom m has not yet proposed
    if *w is free* **then**
        (m, w) become engaged
    else w is currently engaged to m'
        if *w prefers m' to m* **then**
            m remains free
        else w prefers m to m'
            (m, w) become engaged
            m' becomes free
        end
    end
end
return *the set S of engaged pairs*

TopHat 2

In an implementation of SMP, what would be the input size when there are n men and n women?

Analysis of SMP

Algorithm: Gale-Shapley Algorithm (1962)

Initially all $m \in M$ and $w \in W$ are free
while *there is a man m who is free and hasn't proposed to every woman* **do**
  Choose such a man m
  Let w be the highest-ranked woman in m's preference list to whom m has not yet proposed
  if *w is free* **then**
    (m, w) become engaged
  else w is currently engaged to m'
    if *w prefers m' to m* **then**
      m remains free
    else w prefers m to m'
      (m, w) become engaged
      m' becomes free
    end
  end
end
return *the set S of engaged pairs*

TopHat 3

In the Gale-Shapely algorithm, what is the maximum number of iterations when there are n men and n women?

Quantifying an Algorithm's Performance

Brute-force

1. Enumerate all possible solutions.
2. Check all possible solutions and keep the best one.

Worst-case

Considering all possible inputs, what is worst possible performance of the algorithm?

- Absolute guarantee on performance.
- Only needs one data point.

Quantifying an Algorithm's Performance

Worst-case

Considering all possible inputs, what is worst possible performance of the algorithm?

- Absolute guarantee on performance.
- Only needs one data point.

Average-case

Given a distribution over the possible inputs, what is the expected performance of the algorithm?

- Without mention of distribution, uniform is assumed.
- Analysis typically more complicated.

Quantifying an Algorithm's Performance

Worst-case

Considering all possible inputs, what is worst possible performance of the algorithm?

- Absolute guarantee on performance.
- Only needs one data point.

Best-case

Considering all possible inputs, what is best possible performance of the algorithm?

- Tends to be meaningless
- Could used when choosing between 2 otherwise equivalent algorithms.

Insertion Sort Analysis

(Introduction to Algorithms, p.26)

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	$n-1$

Overall:

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

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

Insertion Sort Analysis

(Introduction to Algorithms, p.26)

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	$n-1$

Overall:

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

$$= an^2 + bn - d$$

Asymptotic Order of Growth

Asymptotic Order of Growth

Bounding $f(n)$ as n grows

- Bound $f(n)$ from above.
- Bound $f(n)$ from below.

Bachmann–Landau notation (Asymptotic notation)

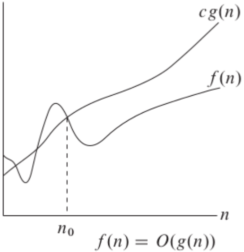
- Big-Oh: O $(\leq)$
- Big-Omega: Ω $(\geq)$
- Big-Theta: Θ (equivalent)
- Little-oh: o $(<<)$
- Little-omega: ω $(>>)$

Big-Oh

Asymptotic upper bound

Formal Definition[1]

$$O(g(n)) = \{f(n) : \exists c, n_0 > 0 \mid 0 \le f(n) \le cg(n)\ \forall n \ge n_0\}$$



Insertion sort:

$$T(n) = an^2 + bn - d \in O(n^2)$$

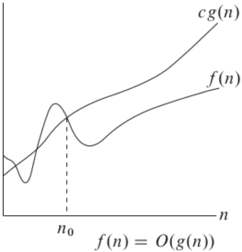
[1] Introduction to Algorithms, Ch 3.1

Big-Oh

Asymptotic upper bound

Formal Definition [1]

$$O(g(n)) = \{f(n) : \exists c, n_0 > 0 \mid 0 \leq f(n) \leq cg(n)\ \forall n \geq n_0\}$$



Insertion sort:

$$T(n) = an^2 + bn - d = O(n^2)$$

Often used, but technically an abuse of notation

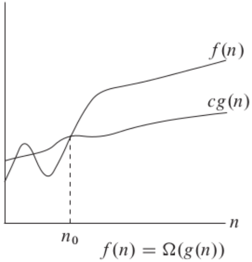
[1] Introduction to Algorithms, Ch 3.1

Big-Omega

Asymptotic lower bound

Formal Definition[1]

$$\Omega(g(n)) = \{f(n) : \exists c, n_0 > 0 \mid 0 \leq cg(n) \leq f(n)\ \forall n \geq n_0\}$$



Insertion sort:

$$T(n) = an^2 + bn - d \in \Omega(n^2)$$

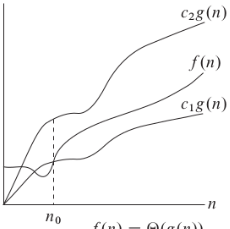
[1] Introduction to Algorithms, Ch 3.1

Big-Theta

Asymptotic tight bound

Formal Definition[1]

$$\Theta(g(n)) = \{f(n) : \exists c_1, c_2, n_0 > 0 \mid$$
$$0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n) \ \forall n \geq n_0\}$$



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

Insertion sort:

$$T(n) = an^2 + bn - d \in \Theta(n^2)$$

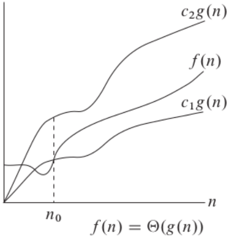
[1] Introduction to Algorithms, Ch 3.1

Big-Theta

Asymptotic tight bound

Formal Definition[1]

$$\Theta(g(n)) = \{f(n) : \exists c_1, c_2, n_0 > 0 \mid 0 \le c_1 g(n) \le f(n) \le c_2 g(n)\ \forall n \ge n_0\}$$



Key Property

For any two functions $f(n)$ and $g(n)$, we have $f(n) = \Theta(g(n))$ if and only it $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

[1] Introduction to Algorithms, Ch 3.1

Little-oh

Formal Definition[1]

$$o(g(n)) = \{f(n) : \forall c > 0 \exists n_0 > 0 \mid 0 \leq f(n) < cg(n)\ \forall n \geq n_0\}$$

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$$

Insertion sort:

$$\begin{aligned} T(n) = an^2 + bn - d &\in o(n^3) \\ &\in O(n^3) \\ &\in O(n^2) \\ &\notin o(n^2) \end{aligned}$$

[1] Introduction to Algorithms, Ch 3.1

Little-omega

Formal Definition[1]

$$\omega(g(n)) = \{f(n) : \forall c > 0 \exists n_0 > 0 \mid 0 \leq cg(n) < f(n)\ \forall n \geq n_0\}$$

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \infty$$

Insertion sort:

$$\begin{aligned} T(n) = an^2 + bn - d &\in \omega(n) \\ &\in \Omega(n) \\ &\in \Omega(n^2) \\ &\notin \omega(n^2) \end{aligned}$$

[1] Introduction to Algorithms, Ch 3.1

Useful Asymptotic Properties

Polynomial Bound

For $c_d > 0, f(n) = c_d \cdot n^d + c_{d-1} \cdot n^{d-1} + \cdots + c_1 \cdot n + c_0 = O(n^d)$

Logarithms

- $\log_b n = \frac{\log_a n}{\log_a b} = \Theta(\log n)$
- $(\log n)^a = o(n^b)$ for any $a, b > 0$

Exponential

- For every $r > 1$ and every $d > 0$, $n^d = o(r^n)$
- $r^n = o(s^n)$ for $r < s$

Analysis of SMP

Algorithm: Gale-Shapley Algorithm (1962)

Initially all $m \in M$ and $w \in W$ are free
while *there is a man m who is free and hasn't proposed to every woman* **do**
    Choose such a man m
    Let w be the highest-ranked woman in m's preference list to whom m has not yet proposed
    if *w is free* **then**
        (m, w) become engaged
    else w is currently engaged to m'
        if *w prefers m' to m* **then**
            m remains free
        else w prefers m to m'
            (m, w) become engaged
            m' becomes free
        end
    end
end
return *the set S of engaged pairs*

Exercise

How would you implement this algorithm so that it has a running time of $O(n^2)$?

Common Runtimes