Algorithms, Design & Analysis

Introduction

Dr. Mudassir Shabbir

Information Technology University

June 19, 2025

About Your Fellows

- Hi there! We are Khadijah Farooqi and Abdullah Salman.
- We are Associate Students at ITU.

Recap of the prev lecture

- Problem: Find the k^{th} smallest element in an array with distinct elements.
- Algorithm: Guess Select Algorithm
  - Input: Array A and k, where k is the position of the desired element.
  - Uniformly Randomly pick a number g GUESS as an element from A.
  - Partition A into:
    - L: Elements less than g.
    - R: Elements greater than g.
  - Compute the rank of g using L and R.
  - If $|L| = k - 1$, return g.
  - Else if $|L| > k - 1$, recursively select k^{th} element from L.
  - Otherwise, recursively select $(k - |L| - 1)^{th}$ element from R.

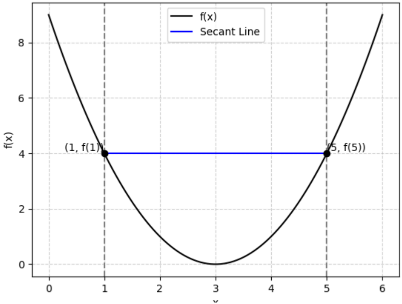
Recap of the prev lecture

- Problem: Find the k^{th} largest element in an array with distinct elements.
  - Method 1: Swap the Definition of L and R
    - L: Elements greater than g.
    - R: Elements smaller than g.
    - Now repeat the same guess algorithm stated above to find the k^{th} largest element.
  - Method 2: Transforming k
    - Let $k = n - k$ and use the selection process for this adjusted k.

Convex Function

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain is a convex set and for all x, y in its domain, and all $\lambda \in [0, 1]$, we have:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$



Relation to Digression

The term Digression refers to the deviation from the straight-line connection.
-> A convex function digresses downward from the secant line.
-> A linear function shows no digression.
-> A concave function digresses upward from the secant line.

Facts About Convex Functions

- The sum of two convex functions is also convex. If $g(x)$ and $f(x)$ are convex, then $r(x) = g(x) + f(x)$ is also convex.
- Geometrically, the line segment connecting $(x, f(x))$ to $(y, f(y))$ lies above the graph of f.
- If f is continuous, checking convexity for a fixed $\lambda \in (0, 1)$, such as $\lambda = \frac{1}{2}$, is sufficient.
- A function f is concave if $-f$ is convex.

Examples of Convex Functions

Examples of Univariate Convex Functions:

- e^{ax}
- $-\textsc{log}(x)$
- x^a (defined on $\mathbb{R}_{++}$), for $a \geq 1$ or $a \leq 0$
- $-x^a$ (defined on $\mathbb{R}_{++}$), for $0 \leq a \leq 1$
- $|x|^a$, for $a \geq 1$
- $x\,\textsc{log}(x)$ (defined on $\mathbb{R}_{++}$)

Reference: S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004. Available at: `http://stanford.edu/~boyd/cvxbook/`

Convex Theorem

Definition

A function $f(x)$ is convex if it satisfies the inequality:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all x, y in its domain and $\lambda \in [0,1]$.

Importance

The convex function theorem plays an important role in:

- Dynamic Programming
- Divide and Conquer
- Optimization Problems

Proof of Convex Theorem

Using the Convex Theorem:

$$f(\lambda a + (1{-}\lambda)b) \leq \lambda f(a) + (1{-}\lambda)f(b).$$

Let's Take:

$$\lambda = \frac{a}{a+b}, \quad (1-\lambda) = \frac{b}{a+b}$$

Why Choosing $\lambda = \frac{a}{a+b}$ and $(1-\lambda) = \frac{b}{a+b}$:

Its because λ must satisfy $\lambda \in [0,1]$ and $\lambda + (1{-}\lambda) = 1$,

so, taking $\lambda = \frac{a}{a+b}$ and $(1-\lambda) = \frac{b}{a+b}$ ensures that λ and $1{-}\lambda$ are positive and sum to 1.

$$\lambda + (1-\lambda) = \frac{a}{a+b} + \frac{b}{a+b} = \frac{a+b}{a+b} = 1$$

Proof of Convex Theorem

Substituting $\lambda = \frac{a}{a+b}$ and $(1-\lambda) = \frac{b}{a+b}$, we get:

$$f\left(\frac{a}{a+b}a + \frac{b}{a+b}b\right) \leq \frac{a}{a+b}f(a) + \frac{b}{a+b}f(b)$$

Simplifying the left-hand side:

$$f\left(\frac{a^2+b^2}{a+b}\right) \leq \frac{af(a)+bf(b)}{a+b}$$

Example: Quadratic Function

Function Choice

Let $f(x) = x^2$ (a convex function). Choose $a = 2$, $b = 4$.

Applying Convexity

$$\lambda = \frac{2}{2+4} = \frac{1}{3}, \quad 1 - \lambda = \frac{4}{6} = \frac{2}{3}$$

$$\lambda a + (1-\lambda)b = \frac{1}{3}(2) + \frac{2}{3}(4) = \frac{10}{3}$$

Checking convexity inequality:

$$f\left(\frac{10}{3}\right) \leq \frac{1}{3}f(2) + \frac{2}{3}f(4)$$

Example: Quadratic Function (Continued)

Simplification

Substituting $f(x) = x^2$:

$$\left(\frac{10}{3}\right)^2 \leq \frac{1}{3}(4) + \frac{2}{3}(16)$$

Simplifying,

$$\frac{100}{9} \leq 12$$

Conclusion

Since $\frac{100}{9} \approx 11.11 \leq 12$, the inequality holds, confirming the convexity of $f(x) = x^2$.

Mathematical Induction (Revision)

Climbing a Infinite Ladder:

Suppose that we have an infinite ladder, we want to know whether we can reach every step on this ladder. We know two things :

a) We can reach the first step of the ladder.
b) If we can reach a particular step of the ladder, then we can reach the next step.
Induction says that these two pieces of information enough to reach the required conclusion!

Here is how it works (informally!)

From (a) we can reach the first step. Then by applying (b) we can reach the second step. Applying (b) again, the third step. And so on. We can apply (b) any number of times to reach any particular step, no matter how high up.

After 100 uses of (b), we know that we can reach the 101st step.

Principal of Mathematical Induction (Revision)

To prove that $P(n)$ is true for all positive integers n, we complete these steps:

Basis Step: Show that $P(1)$ is true.

Inductive hypotheis: We asume that statement is true for some positive integer k

Inductive Step(proof): Based on the assumption, we prove that the statement must be true for $k + 1$

To complete the inductive step, assuming the inductive hypothesis that $P(k)$ holds for an arbitrary integer k, show that must $P(k + 1)$ be true.

Proof by Induction: Example (Revision)

Example: Prove the following using induction.
"The sum 1 + 2 + ... + n is equivalent to $\frac{n(n+1)}{2}$."

Base Case:
Does our statement hold for the first value of n, which is $n = 1$?
$P(1) = 1 = \frac{1(2)}{2} = 1$

Inductive Hypothesis:
We assume that the statement is true for some k. So, we assume the following is true.
$P(k) = \sum_{i=1}^{k} i = \frac{k(k+1)}{2}$

Reference: Professor's DS slides of F2023

Proof by Induction: Example (Revision)

> Example: Prove the following using induction.
> "The sum 1 + 2 + ... + n is equivalent to $\frac{n(n+1)}{2}$."

Inductive Hypothesis(Cont.):

Now, using this assumption, i.e., $P(k) = \sum_{i=1}^{k} i = \frac{k(k+1)}{2}$, we need to prove that the statement is also true for $k+1$.

Show that : $P(k+1) = \frac{(k+1)(k+2)}{2}$

$P(k+1) = 1 + 2 \ldots + k - 1 + k + (k+1) = P(k) + (k+1) = \frac{(k)(k+1)}{2} + (k+1)$

$= \frac{k(k+1)+2(k+1)}{2}$

$= \frac{(k+1)(k+2)}{2}$ Q.E.D.

Reference: Professor's DS slides of F2023

Recursion

Recursion is a method to solve a problem where the solution depends on solutions to smaller subproblems of the same problem. Recursive functions (function calling itself) are used to solve problems based on Recursion. The main challenge with recursion is to find the time complexity of the Recursive function.

Methods of Solving Recursion

- Tree Method
- Substitution Method
- Master Theorem

Recursion

The Recursion-Tree Method $\rightarrow$ useful for guessing the asymptotic bound of a recurrence relation.

Recurrence - Recursion Tree Relationship

Recurrence - Recursion Tree Relationship

$T(1) = c$

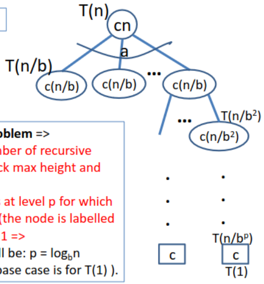
Problem size

The local TC at the node

$T(n) = a*T(n/b) + cn$

Number of subproblems =>
Number of children of a node in the recursion tree. =>
Affects the number of nodes per level. At level i there will be a^i nodes.
Affects the level TC.

Size of a subproblem =>
Affects the number of recursive calls (frame stack max height and tree height)
Recursion stops at level p for which the pb size is 1 (the node is labelled T(1)) => $n/b^p = 1$ =>
Last level, p, will be: $p = \log_b n$
(assuming the base case is for T(1)).



TC = time complexity

5

Reference: https://ranger.uta.edu/ alex/courses/3318/lectures/08recurrences.pdf

Recursion Tree Method

Recurrence Relation:

$$T(n) = 2T\left(\frac{n}{2}\right) + c$$

Base Case:

$$T(1) = c$$

Recursion Tree:

- At each level, the problem size halves (in this case).
- Number of nodes doubles at each level.
- Level p has 2^p nodes, each with problem size $\frac{n}{2^p}$.
- Recursion stops when problem size is 1:

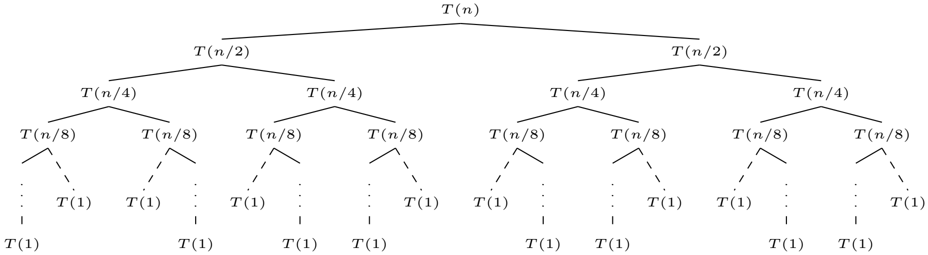
$$\frac{n}{2^p} = 1 \Rightarrow p = \lg n$$

Total Time Complexity:

$$\text{Total TC} = c(1 + 2 + 4 + \cdots + 2^p) = c \cdot (2^{p+1} - 1) = 2c \cdot 2^p = 2cn = \Theta(n)$$

Recursion Tree for $T(n) = 2T(n/2) + c$

Recursion Tree:



Tree Method - Complexity Analysis

Table Representation:

Level	Arg/pb size	TC of 1 node	Nodes per level	Level TC
0	n	c	1	c
1	$n/2$	c	2	$2c$
2	$n/4$	c	4	$4c$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$n/2^i$	c	2^i	$2^i c$

Stopping Condition:

$$\frac{n}{2^p} = 1 \Rightarrow p = \log_2 n$$

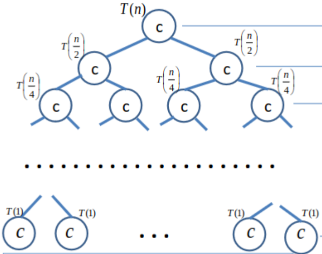
Total Tree Complexity:

$$\text{Tree TC} = c(1 + 2 + 2^2 + \cdots + 2^p) = c \cdot \frac{2^{p+1} - 1}{2 - 1} = 2c \cdot 2^p = 2cn = \Theta(n)$$

Tree Method - Reference Image

Recursion Tree for: T(n) = 2T(n/2)+c

Base case: T(1) = c



Level	Arg/ pb size	TC of 1 node	Nodes per level	Level TC
0	n	c	1	c
1	n/2	c	2	2c
2	n/4	c	4	4c
…				
i	$n/2^i$	c	2^i	$2^i c$
…				
p=lgn	1 ($=n/2^p$)	c	2^p ($=n$)	$2^k c$

Stop at level p, when the subtree is T(1).
=> The problem size is 1, but the general formula for the problem size, at level p is:
$n/2^p$=> $n/2^p = 1$ => $p = \lg n$

Tree TC $= c(1+2+2^2+2^3+\ldots+2^i+\ldots+2^p) = c2^{p+1}/(2-1)$
$= 2c2^p = 2cn = \Theta(n)$

Reference: https://ranger.uta.edu/ alex/courses/3318/lectures/08recurrences.pdf

Recursion

Substitution - The substitution method is a technique used to solve recurrence relations by guessing the form of the solution and then verifying it using mathematical induction.

Substitution

KEY CONCEPTS

- Recursive Relation
- Base case
- Inductive Hypothesis

Substitution

Recursive Relation->An equation that recursively defines a sequence of values.

Base Case->The condition under which the recursion stops.

Inductive Hypothesis->An assumption that the solution holds for a smaller input.

Steps in Substitution Method

- Guess the Form of the Solution: Make an educated guess about the form of the solution based on the recurrence relation.
- Verify the Guess: Use mathematical induction to prove that the guessed solution is correct.
- Solve for Constants: Determine the constants in the solution to match the initial conditions.

Substitution - Example

- Recurrence Relation: $T(n) = 2T(n/2) + n$
- Guess: $T(n) = O(n \log n)$
- Verification: Use induction to prove that $T(n) \leq cn \log n$ for some constant c.

Substitution - Mathematical Induction

Base Case->Show that the solution holds for the smallest value of n.

For $n = 1$,

$$T(1) = 1 \leq c \cdot 1 \cdot \log 1$$

Inductive Step->Assume the solution holds for n/2 and prove it holds for n.

Substitution- Common Mistakes

- Incorrect Guesses: Choosing a form that does not satisfy the recurrence relation.
- Insufficient Verification: Not fully proving the solution with induction.
- Ignoring Base Cases: Failing to verify the base case in the induction process.

Substitution- Application

Merge Sort: Analyzing the time complexity using the substitution method.

- Recurrence:

$$T(n) = 2T(n/2) + n$$

- Solution:

$$T(n) = O(n \log n)$$

Quick Sort: Understanding the worst-case and average-case scenarios.

- Recurrence:

$$T(n) = T(k) + T(n - k - 1) + n$$

- Solution:
  O(nlogn) on average and in the best case, but $O(n^2)$ in the worst case.

Recurrence Solving - Example no. 1

Solving $T(n) = C + T(9n/10)$

Using Tree Method:

Expanding the recurrence:

$$\begin{aligned}
T(n) &= C + T(9n/10) \\
T(9n/10) &= C + T((9/10)^2 n) \\
T(n) &= C + \big(C + T((9/10)^2 n)\big) \\
&= 2C + T((9/10)^2 n) \\
T((9/10)^2 n) &= C + T((9/10)^3 n) \\
T(n) &= 3C + T((9/10)^3 n)
\end{aligned}$$

Generalizing:

$$T(n) = iC + T((9/10)^i n)$$

Summing the Contributions Each level contributes: $C(9/10)^i n$.

$$T(n) = \sum_{i=0}^{\infty} C(9/10)^i n$$

This is a geometric series with sum:

$$T(n) = Cn \sum_{i=0}^{\infty} (9/10)^i$$

Using the formula for an infinite geometric series:

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}, \quad \text{where } r = 9/10,$$

we get:

$$T(n) = Cn \frac{1}{1 - 9/10} = Cn \cdot 10.$$

Final Result: $T(n) = 10Cn$.

Using Substitution Method:

We now have a guess: $T(n) \leq C \cdot 10n$.
Defining the Constant:

$$T(n) \leq Can, \quad \text{where } a \text{ is a constant }.$$

Base Case: For $n \leq 5$, we assume:

$$T(n) \leq Ca \cdot 5.$$

Inductive Hypothesis: Assuming it is true for all $k < n$, we prove it for n:

$$Cn\left(1 + \frac{9}{10}a\right) \leq Cna.$$

Recurrence Solving - Example no. 1

Solving $T(n) = C + T(9n/10)$

Proof:

$$T(n) \leq \text{Can}$$

$$T(n) = C.n + T(\frac{9n}{10})$$

By inductive Hypothesis:

$$= \text{C.n} + \frac{\cdot 9}{10} n \cdot Ca$$

$$T(n) = \text{C.n}\left(1 + \frac{9}{10}a\right)$$

$$\text{C.n}\left(1 + \frac{9}{10}a\right) \leq \text{C.n.}a$$

$$1 + \frac{9}{10}a \leq a$$

$$1 \leq a - \frac{9}{10}a$$

$$1 \leq \frac{a}{10} \rightarrow 10 \leq a$$

Q/A

That's why we assume 10 as 'a' in first in substitution in the class

Ammar asked prof how did we replace $T(n) = cn + T\left(\frac{9}{10}n\right)$ to $cn + \left(\frac{9}{10}\right) nca$

Professor: We did this through Inductive Hypothesis

Hamza asked prof about the sign in $T(n) = cn \cdot (1 + \frac{9}{10}nca)$ that it should be less than or equal to ($\leq$).

Hamza was right, and professor changed it to $T(n) \leq$ Cn $+ \frac{9}{10}n \cdot ca$

If we change number of chunks what will happen??

Recurrence would change

Recurrence Solving - Example no. 2

Solving $T(n) = 2T(n/2) + Cn$

Guessing a Solution: Assume $T(n) = aCn$.

Base Case: Same as before.

Inductive Hypothesis: Assume true for all $k < n$.

$$\begin{aligned} T(n) &= 2T(n/2) + Cn \\ &= 2\left(aC\frac{n}{2}\right) + Cn \\ &= aCn + Cn \\ &= Cn(a+1). \end{aligned}$$

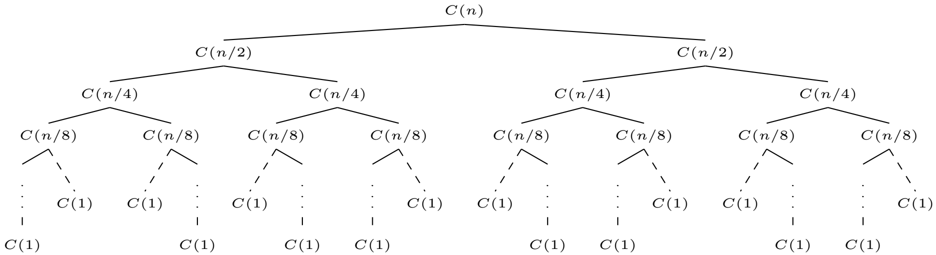
Final Complexity: $O(n)$.

Professor asked can anyone find problem in it Hamza: Our guess was a constant and Professor in beginning of this solution said that he knows this solution is wrong, so basically our random guess was wrong. Now we are making another guess through Tree method.

Recurrence Solving - Example no. 2

Solving $T(n) = 2T(n/2) + Cn$

Recursion Tree:



Recurrence Solving - Example no. 2

Solving $T(n) = 2T(n/2) + Cn$

Recursion Tree Analysis:
The recursion tree for $T(n) = 2T(n/2) + Cn$ shows the following:

- At level 0: We have $T(n)$, which contributes Cn.
- At level 1: We have 2 nodes, each contributing $C \cdot \frac{n}{2}$, so the total contribution at level 1 is Cn.
- At level 2: We have 4 nodes, each contributing $C \cdot \frac{n}{4}$, so the total contribution at level 2 is Cn.
- At level k: We have 2^k nodes, each contributing $C \cdot \frac{n}{2^k}$, so the total contribution at level k is Cn.

Height of the tree: The height of the tree is $\log_2 n$, because we divide n by 2 at each level.
Total Work: The total work is the sum of the contributions at each level:

$$T(n) = Cn \cdot \log_2 n$$

Hence, the overall time complexity is:

$$T(n) = O(n \log n)$$

Home Work

Prove this guess from Substitution method is your homework

- Remember not every correct guess can be proven through Substitution Method

Recursion Solving - Example No. 3

Initial Guesses for the Recursion

Recursion: $T(n) = 2T(n/5) + 5$
Professor asked the class to make random guesses:

- Hamza's Guess: $T(n) \leq 5 \log n$
- Jalal's Guess: $T(n) \leq C \cdot n \log n$

Results:

- Hamza's guess was incorrect.
- Jalal's guess was also incorrect (Result was ≤ 10, not 5).

Recursion Solving - Example No. 3

New Guess: $T(n) \leq an$

Inductive Hypothesis:

$$T(k) \leq a \cdot k \quad \text{for all} \quad k < n$$

Substitute in the Recurrence:

$$T(n) \leq 2 \cdot \left(\frac{an}{2}\right) + 5$$

$$T(n) \leq an + 5$$

Issue: The extra $+5$ is not what we need.

- We can change the guess multiple times until we find the desired result.

Recursion Solving - Example No. 3

Saifullah's Suggested Guess: $T(n) \leq (a-5)n$

Inductive Hypothesis:

$$T(k) \leq (a-5) \cdot k \quad \text{for all} \quad k < n$$

Substitute in the Recurrence:

$$T(n) \leq 2 \cdot \left(\frac{(a-5)n}{2} \right) + 5$$

$$T(n) \leq (a-5)n + 5$$

$$T(n) = an - 5n + 5$$

$$a = 5 - \frac{5}{n}$$

Issue: This approach still didn't work.

Recursion Solving - Example No. 3

Final Correct Guess: $T(n) \leq an - 5$

Inductive Hypothesis:

$$T(k) \leq (a-1) \cdot k \quad \text{for all} \quad k < n$$

Substitute in the Recurrence:

$$T(n) \leq 2 \cdot \left(\frac{(a-1)n}{2} \right) + 5$$

$$T(n) = an - n + 5$$

$$T(n) \leq an - 5$$

Conclusion: This is the correct guess.

- Not every correct guess can be proven by substitution method.
- We need to make the hypothesis strong to make induction work.

Homework: Analyzing Recurrences

1. Recurrence: $T(n) = 2T(n/2) + x$
   - Where x can be one of the following:
     - $x = n \log n$
     - $x = n^2$
     - $x = \log n$

2. Recurrence: $T(n) = 3T(n/2) + Cn$