Efficient Algorithms

Ekkapot Charoenwanit

Software Systems Engineering

TGGS

KMUTNB

Lecture 7: Dynamic Programming

Fibonacci Numbers

Fibonacci numbers can be defined as a recurrence as follows:
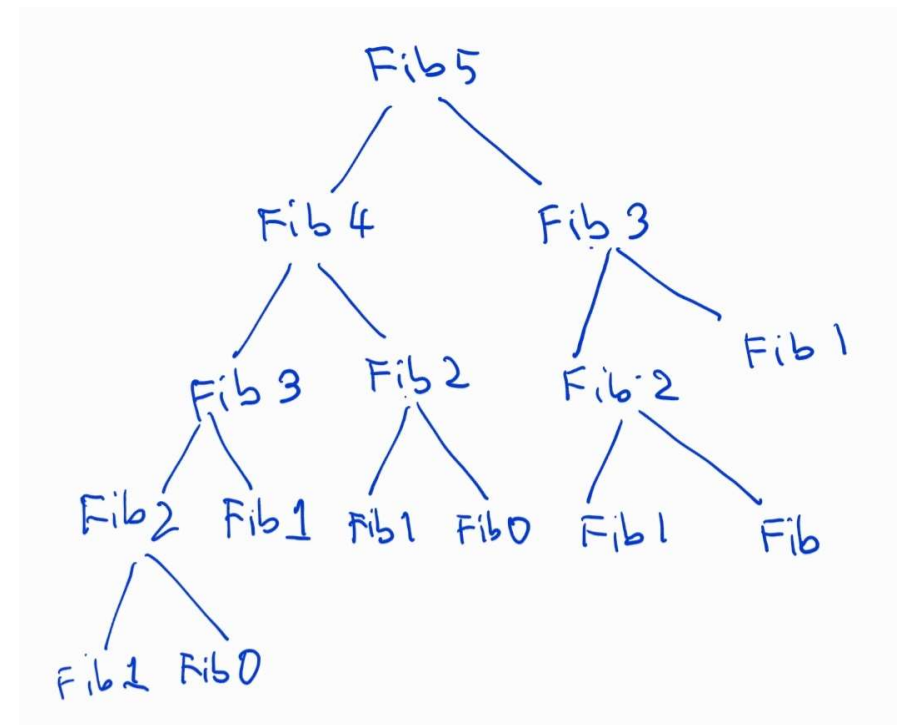
$$F_0 = 0$$
$$F_1 = 1$$
$$F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 2$$

Fibonacci Numbers: Top-Down Approach

Top-Down Approach

- implemented via recursion

```
1: procedure FIB(n)
2:     if n ≤ 1 then
3:         return n
4:     else
5:         FIB(n − 1) + FIB(n − 2)
```

Fibonacci Numbers: Top-Down Approach

Let $N(n)$ be the number of recursive calls $FIB(n)$ makes.

We can write $N(n)$ as

$$N(n) = N(n-1) + N(n-2) + 1 \text{ for } n \geq 2$$

$$N(0) = 1 \text{ and } N(1) = 1$$

Solving the recurrence, we have

$$N(n) = 2F(n+1) - 1$$

,where $F(n) = \frac{\phi^n - (1-\phi)^n}{\sqrt{5}}$,where $\phi = \frac{1+\sqrt{5}}{2} \approx 1.61803$

Fibonacci Numbers: Top-Down Approach

$$F(n) = \frac{\phi^n - (1 - \phi)^n}{\sqrt{5}}$$

The number of recursive calls is ***exponential*** in the value of n.

Therefore, the running time of $FIB(n)$ exhibits an ***exponential growth*** in n.

Space complexity is proportional to the depth of the recursion tree, i.e., $\Theta(n)$.

Fibonacci Numbers: Overlapping Subproblems

The recursion tree of $FIB(n)$ exhibits a property known as **overlapping subproblems**.

As you can see in the recursion tree for $FIB(5)$,

- $FIB(0)$ is called 3 times
- $FIB(1)$ is called 5 times
- $FIB(2)$ is called 3 times
- $FIB(3)$ is called 2 times
- $FIB(4)$ is called 1 time
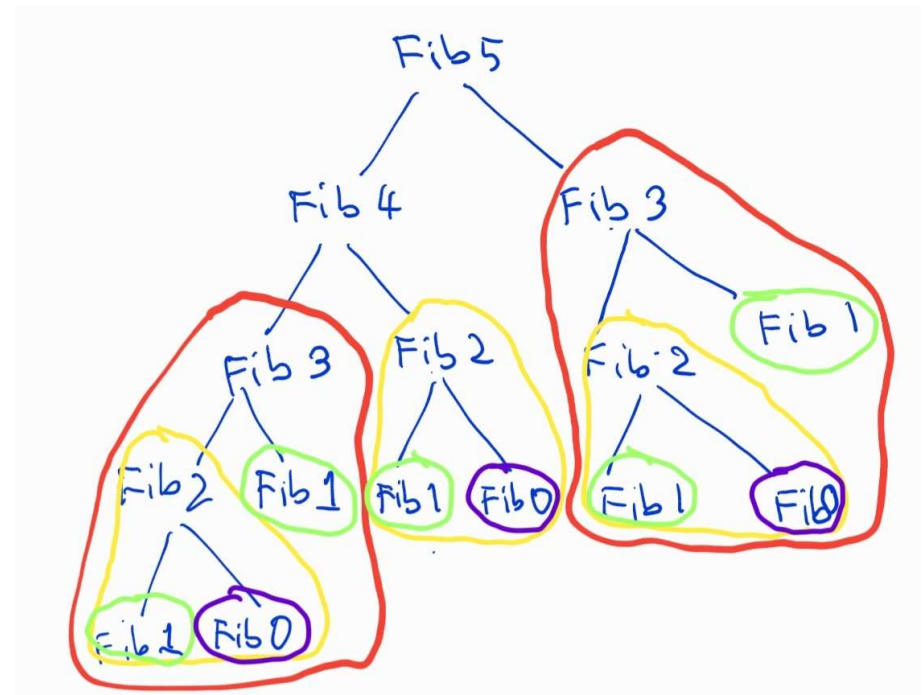
Fibonacci Numbers: Memoization

Memoization means "**to remember**" by storing values into a **table**.

All the entries of the table *F* are initially initialized to *0*.

- The number of different subproblems is $n + 1$.
- The number of recursive calls is $n + 1 = \Theta(n)$.
- Work done per call is $\Theta(1)$.

1: **procedure** FIB(n, $F[1...n]$)
2:      **if** $n \leq 1$ **then**
3:          **return** n
4:      **else**
5:          **if** F$(n) > 0$ **then**
6:              **return** $F[n]$
7:          **return** FIB($n - 1, F$) + FIB($n - 2, F$)

The running time = #recursive calls * work done per call = $\Theta(n) \cdot \Theta(1) = \Theta(n)$.

Fibonacci Numbers: Bottom-Up Approach

Bottom-Up Approach:

- Implemented via tabulation

- Smaller values -> Larger values

- Space complexity is $\Theta(n)$

- Time complexity is $\Theta(n)$

```
1: procedure FIB(n)
2:      F = NEW TABLE[0...n]
3:      F[0] = 0
4:      F[1] = 1
5:      for i = 2 → n do
6:          F[i] = F[i − 1] + F[i − 2]
7:      return F[n]
```

***Space complexity can be further optimized. (See PS 5.1.2)

Top-Down vs Bottom-Up

Top-Down Approach:

- Divides the original problem into smaller subproblems
- Solves the subproblems
- Combines the solutions to subproblems
- Is implemented via recursion
- Minimizes the number of subproblems via memoization

Bottom-Up Approach:

- Solves smaller subproblems first
- Generates solutions to larger subproblems from solutions to smaller ones
- Is implemented using loop iterations
- Solves each subproblem only once

Dynamic Programming

Dynamic programming (DP) is an algorithm **design paradigm** where a problem is solved recursively by dividing into smaller subproblems like divide-and-conquer.

However, DP exploits the following **two properties** of the problem:

- **Overlapping subproblems**
    - Subproblems solving the same problems are repeated many times
    - Solve such repeated subproblems only once and reuse their solutions via
        - Memoization (top-down)
        - Tabulation (bottom-up)
- **Optimal Substructure**
    - An optimal solution to a problem contains optimal solutions to smaller subproblems

DP is usually used to solve **optimization problems** as we will see shortly.
- **Minimization**
- **Maximization**

Matrix Chain Multiplication

Recalling matrix multiplication:

1) The product $C = AB$ of a $p \times q$ matrix A and a $q \times r$ matrix B is a $p \times r$ matrix C computed by

$$c_{ij} = \sum_{k=1}^{q} a_{ik} \cdot b_{kj}$$

for $1 \leq i \leq p$ and $1 \leq j \leq r$.

2) Matrix multiplication is **associative**, i.e., $A(BC) = (AB)C$ so different **parenthesizations** do not alter the value.

Matrix Chain Multiplication

$$C \quad = \quad A \quad \times \quad B$$

$$[p \times r] \qquad [p \times {\color{red}q}] \qquad [{\color{red}q} \times r]$$

To multiply A and B,

pqr multiplications are needed.

Matrix Chain Multiplication: Example

Example: Let A_1, A_2 and A_3 be matrices of the following dimensions 10×100, 100×5 and 5×50, respectively.

There are *2* different parenthesizations for a matrix chain of length *3*.

Case I: $A_1(A_2\ A_3)$
$A_2 A_3$ requires $100 \cdot 5 \cdot 50 = 25000$ multiplications, whose result is a matrix $A_{2,3}$ of dimension 100×50.
$A_1 A_{2,3}$ requires $10 \cdot 100 \cdot 50 = 50000$ multiplications, whose result is a matrix $A_{1,3}$ of dimension 10×50.
Therefore, the total number of multiplications is $25000 + 50000 = 75000$.

Case II: $(A_1 A_2)\ A_3$
$A_1 A_2$ requires $10 \cdot 100 \cdot 5 = 5000$ multiplications, whose result is a matrix $A_{1,2}$ of dimension 10×5.
$A_{1,2} A_3$ requires $10 \cdot 5 \cdot 50 = 2500$ multiplications, whose result is a matrix $A_{1,3}$ of dimension 10×50.
Therefore, the total number of multiplications is $5000 + 2500 = 7500$.

Matrix Chain Multiplication: Brute Force

How many different ways can we parenthesize a matrix chain of length n?

Let $C(n)$ denote the number of ways of parenthesizations of a matrix chain of length n.

Therefore, $C(n) = \begin{cases} 1, & n \leq 2 \\ \sum_{k=1}^{n-1} C(k)\, C(n-k), & n \geq 3 \end{cases}$

$C(n) = \Omega\left(\frac{4^n}{n^{1.5}}\right)$, which is related to the Catalan numbers.

Since the number of ways of placing parenthesis is ***exponential*** in the length of a matrix chain, a brute force approach is impractical.

Matrix Chain Multiplication: Notation

Given a matrix chain $A_1 A_2 \ldots A_n$ of length n,

A_1 has a dimension of $p_0 \times p_1$,

A_2 has a dimension of $p_1 \times p_2$,

$\ldots$

A_n has a dimension of $p_{n-1} \times p_n$.

$A_i \ldots A_j$ is a matrix of size $p_{i-1} \times p_j$.

Let $m(i, j)$ be the number of multiplications for $A_i \ldots A_j$.

We want to find $m(1, n)$.

Matrix Chain Multiplication: Optimal Substructure

Suppose we **split** a matrix chain $A_i \dots A_j$ at some position $i \leq k < j$.

$$A_i \dots A_j = (A_i \dots A_k)(A_{k+1} \dots A_j)$$

$A_i \dots A_k$ evaluates to a $p_{i-1} \times p_k$ matrix $A_{i,k}$ whose number of multiplications is $m(i,k)$.

$A_{k+1} \dots A_j$ evaluates to a $p_k \times p_j$ matrix $A_{(k+1),j}$ whose number of multiplications is $m(k+1,j)$.

Multiplying A_{ik} and $A_{(k+1),j}$ requires $p_{i-1}p_k p_j$ multiplications.

Therefore, the total number of multiplications is $m(i,k) + m(k+1,j) + p_{i-1}p_k p_j$.

Matrix Chain Multiplication: Optimal Substructure

If an optimal solution to $A_i \ldots A_j$ involves splitting into $A_i \ldots A_k$ and $A_{k+1} \ldots A_j$ at the final step, solutions to parenthesizations of $A_i \ldots A_k$ and $A_{k+1} \ldots A_j$ are also optimal.

Proof: We will use a ***Cut-and-Paste*** argument.

Suppose the solution $m(i,k)$ to $A_i \ldots A_k$ is ***not optimal***. We can then replace this solution to $A_i \ldots A_k$ with a ***better solution*** $m'(i,k) < m(i,k)$, hence a ***better solution*** to $A_i \ldots A_j$

$$m'(i,j) = m'(i,k) + m(k+1,j) + p_{i-1}p_k p_j < m(i,j)$$

,which contradicts the optimality of the solution to $A_i \ldots A_j$.

A symmetric argument is applied to optimality of $A_{k+1} \ldots A_j$. ∎

Matrix Chain Multiplication: Optimal Substructure

Having proved **optimal substructure** of the Matrix Chain Multiplication problem,

the next question is "**where** do we split ?, i.e., what is the position k ?".

Answer: Try them all !!!

Matrix Chain Multiplication: Recursive Formulation

Recursive Formulation: Let $M(i,j)$ be the minimum number of multiplications for a matrix chain $A_i \dots A_j$.

$$M(i,j) = \begin{cases} 1, & i = j \\ \min_{i \leq k < j} \{M(i,k) + M(k+1,j)\} + p_{i-1}p_k p_j & i < j \end{cases}$$

We do not know k so we try all the $j - i$ possible values and find k that provides the ***best*** (***smallest***) solution.

Matrix Chain Multiplication: Top-Down

```
1: procedure MCM(i,j,p[0...n])
2:     if i == j then
3:         return 0
4:     minMCM = ∞
5:     for k = i → j − 1 do
6:         minMCM = min(minMCM,
7:             MCM(i, k, p) + MCM(k + 1, j, p) + p[i] * p[k] * p[j])
8:     return minMCM
```

The running time is exponential in the chain length n.

Matrix Chain Multiplication: Overlapping Subproblems
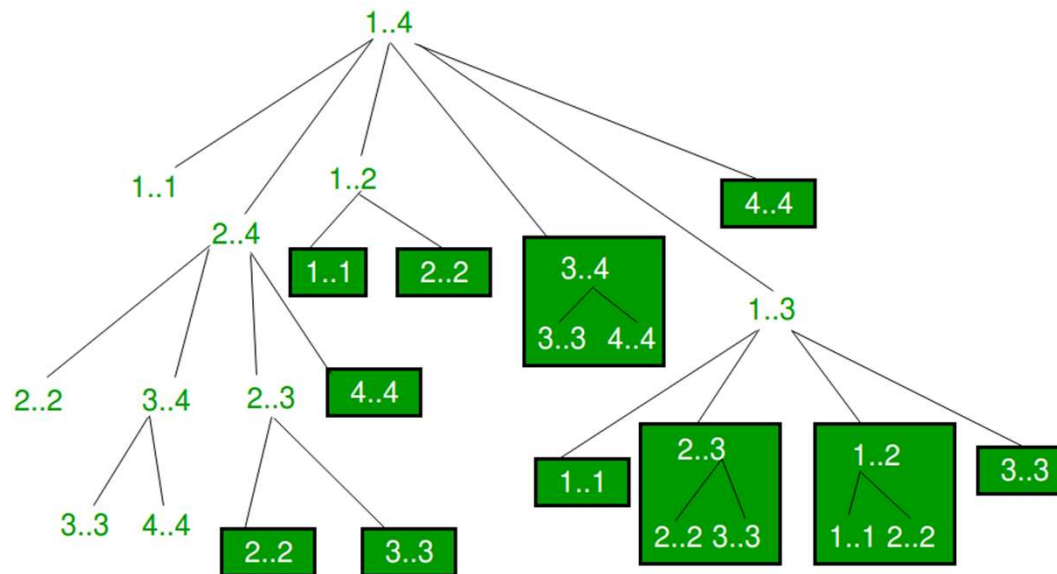
Matrix Chain Multiplication: Overlapping Subproblems

Additionally, the number of **distinct** subproblems is relatively **small** (i.e. **polynomial** in problem size).

The number of distinct subproblems is

$$\Theta(n^2), \text{ which is polynomial in the problem size } n.$$

Matrix Chain Multiplication: Memoization

```
1: procedure MCM(i,j,p[0...n], M[1...n][1...n])
2:     if i == j then
3:         return 0
4:     if M[i][j] > 0 then
5:         return M[i][j]
6:     M[i][j] = ∞
7:     for k = i → j − 1 do
8:         M[i][j] = min(M[i][j],
9:             MCM(i, k, p, M) + MCM(k + 1, j, p, M) + p[i] ∗ p[k] ∗ p[j])
10:    return M[i][j]
```

Work done per recursive calls is $\Theta(n)$.

#Recursive Calls is $\Theta(n^2)$.

Total work done is $\Theta(n^3)$.

Matrix Chain Multiplication: Bottom-Up

```
 1: procedure MCM(p[0...n])
 2:     M = NEW TABLE[0...n][0...n]
 3:     P = NEW TABLE[1...n − 1][2...n]
 4:     for i = 1 → n do
 5:         M[i][i] = 0
 6:     for l = 2 → n do
 7:         for i = 1 → n − l + 1 do
 8:             j = i − l + 1
 9:             M[i][j] = ∞
10:             for k = i → j − 1 do
11:                 q = M[i, k] + M[k + 1, j] + p[i] * p[k] * p[j]
12:                 if q < M[i][j] then
13:                     M[i][j] = q
14:                     S[i][j] = k
15:     return (M, S)
```

Matrix Chain Multiplication: Solution Reconstruction

```
1: procedure PRINTOPTIMALPAREN(i, j, S)
2:     if i == j then
3:         PRINT A_i
4:     else
5:         PRINT (
6:         PRINTOPTIMALPAREN(i, S[i][j], S)
7:         PRINTOPTIMALPAREN(S[i][j] + 1, j, S)
8:         PRINT )
```

Longest Common Subsequence

Definition: The ***Longest Common Subsequence*** (***LCS***) problem is as follows. Given two strings X of length m and Y of length n, our goal is to determine the longest common subsequence, that is, the longest sequence of characters that do not necessarily appear contiguously in the two strings.
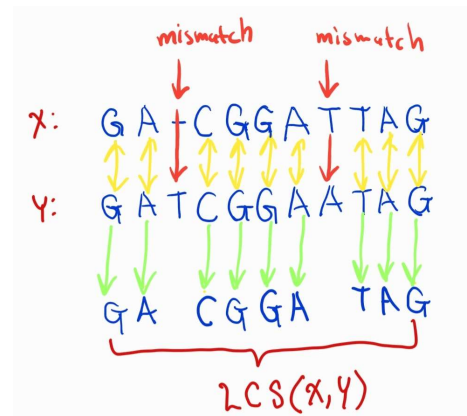
LCS finds its application in DNA sequence alignment
- used to compares similarity between two DNA sequences
- Longest Common Subsequence -> ***Best Alignment***

Longest Common Subsequence: Example

Consider the following DNA fragments *X* and *Y*:

X : GACGGATTAG and *Y* : GATCGGAATAG

Therefore, the longest common subsequence is GACGGATAG.

Longest Common Subsequence: Notation

Notation:

Given two strings X of length m and Y of length n,

$$X = < x_1, x_2, x_3, \ldots, x_m >$$
$$Y = < y_1, y_2, y_3, \ldots, y_n >$$
$$X_i = < x_1, x_2, x_3, \ldots, x_i >$$
$$Y_j = < y_1, y_2, y_3, \ldots, y_j >$$

$LCS(X_i, Y_j)$: longest common subsequence of X_i and Y_j

$LCS(X, Y) = LCS(X_m, Y_n)$

$LCS(i, j)$: ***length of*** $LCS(X_i, Y_j)$

Longest Common Subsequence: Optimal Substructure

Theorem (Optimal Substructure):

Let $X = < x_1, x_2, \ldots, x_m >$ and $Y = < y_1, y_2, \ldots, y_n >$ be sequences.
Let $Z = < z_1, z_2, \ldots, z_k >$ be any LCS of X and Y.

1) If $x_m = y_m$, then $z_k = x_m = y_m$ and Z_{k-1} is an LCS of X_{m-1} and Y_{n-1}.

2) If $x_m \neq y_m$, then $z_k \neq x_m$ implies Z is an LCS of X_{m-1} and Y_n.

3) If $x_m \neq y_m$, then $z_k \neq y_n$ implies Z is an LCS of X_m and Y_{n-1}.

Longest Common Subsequence: Optimal Substructure

CASE I:

If $x_m = y_m$, then $z_k = x_m = y_m$ and Z_{k-1} is an LCS of X_{m-1} and Y_{n-1}.
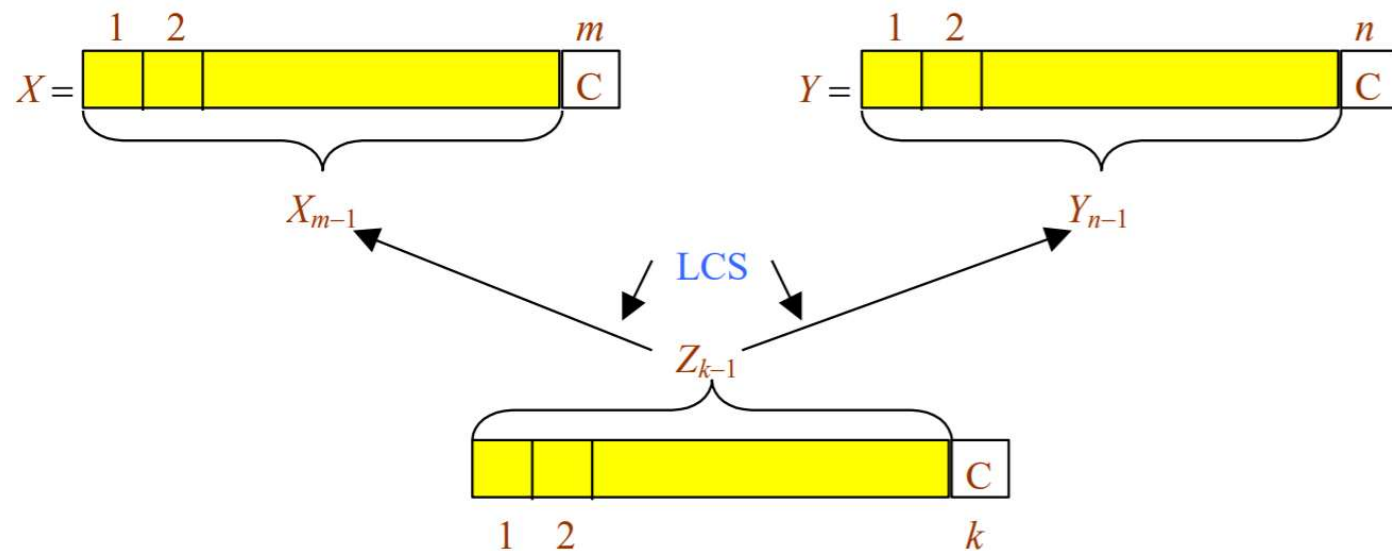
Longest Common Subsequence: Optimal Substructure

Proof: Assume $z_k \neq x_m = y_m$.

We can append $x_m = y_m$ to Z to obtain a subsequence of length $k + 1$, which contradicts optimality of Z.

Therefore, $z_k = x_m = y_m$.

Hence, the prefix Z_{k-1} is a common subsequence (**CS**) of length $k - 1$.

We must show that Z_{k-1} is, in fact, a LCS of X_{m-1} and Y_{n-1}.
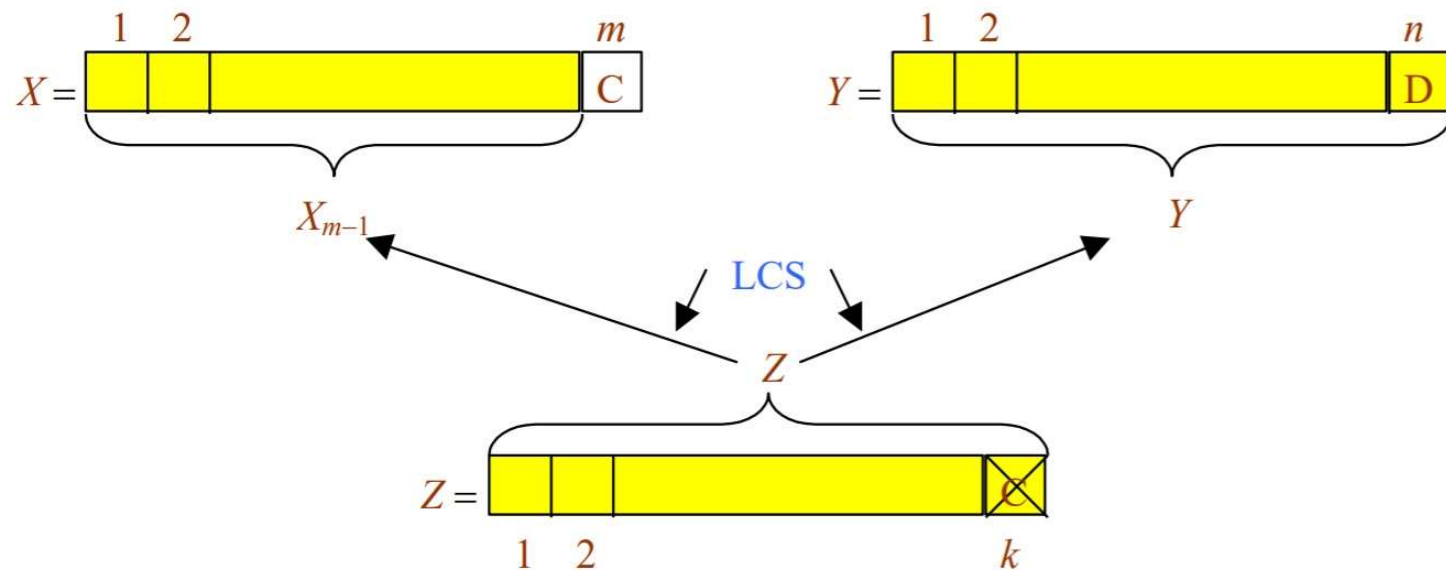
We will prove using a ***cut-and-paste argument*** as follows:

Assume there exists a CS W of X_{m-1} and Y_{n-1} with $|W| = k$.
Appending $x_m = y_n$ to W will produce a CS of length $k + 1$, contradicting optimality of Z whose length is k. ∎

Longest Common Subsequence: Optimal Substructure
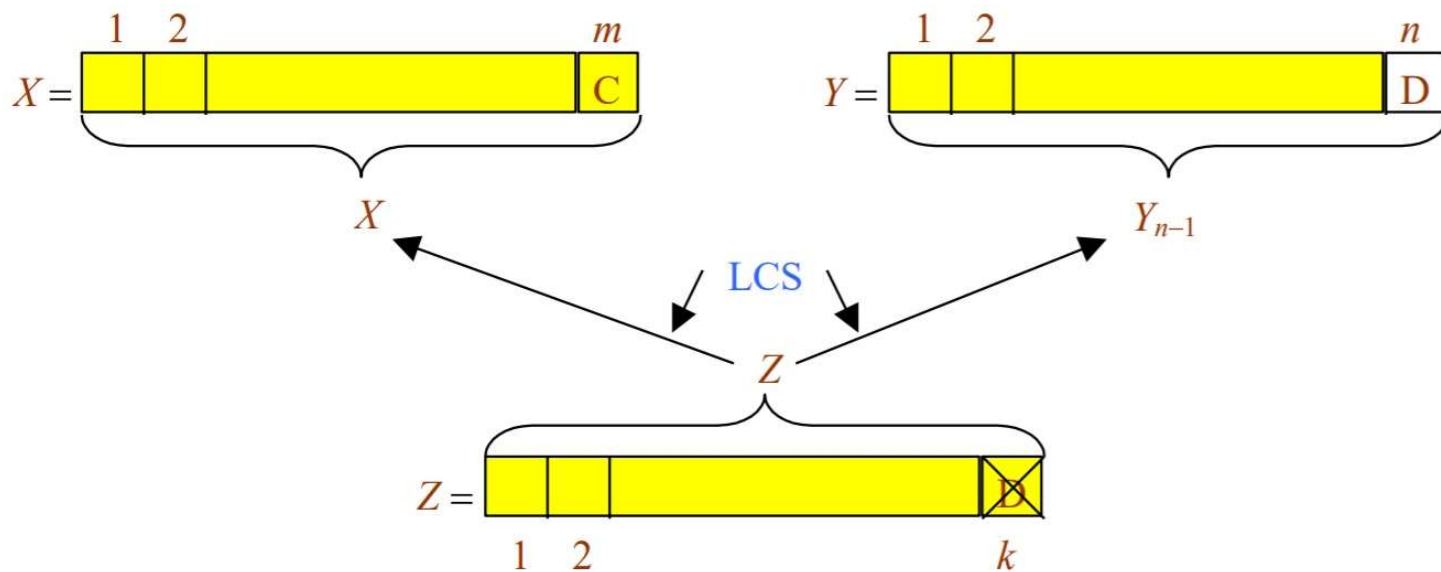
CASE II:

If $x_m \neq y_m$, then $z_k \neq x_m$ implies Z is an LCS of X_{m-1} and Y_n.

Longest Common Subsequence: Optimal Substructure

CASE III:

If $x_m \neq y_m$, then $z_k \neq y_n$ implies Z is an LCS of X_m and Y_{n-1}.

Longest Common Subsequence: Optimal Substructure

Proof: If $z_k \neq x_m$ then Z is a CS of X_{m-1} and Y_n.

We have to show that Z is, in fact, an LCS of X_{m-1} and Y_n.

Assume that there exists a CS W of X_{m-1} and Y_n with $|W| > k$.

Then, W would also be a CS of X_m and Y_n, hence contradicting optimality of Z whose length is k.

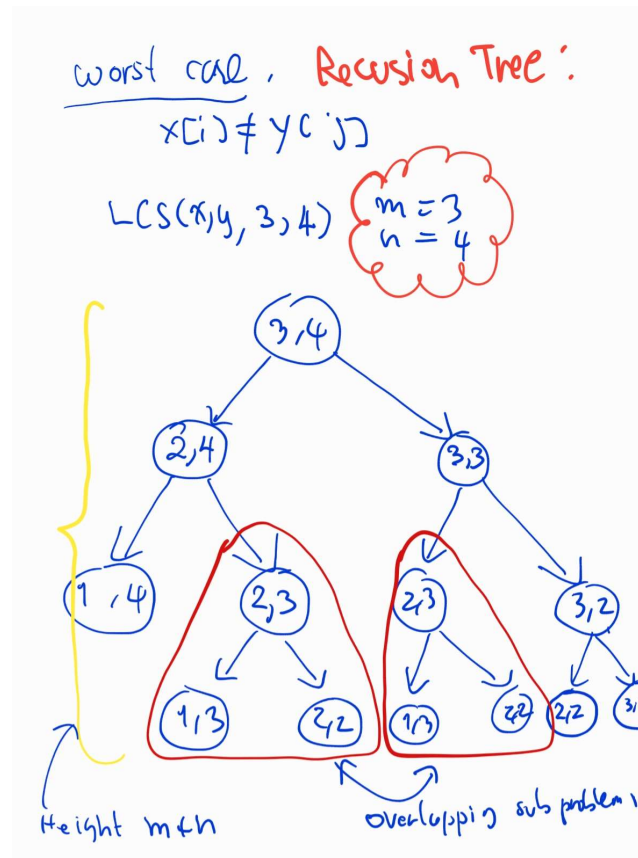Therefore, Z is a LCS of X_{m-1} and Y_n. ■

Proof for **Case III** is symmetric to the proof of **Case II**.

Longest Common Subsequence: Recursive Formulation

Recursive Formulation:

$$c[i,j] = \begin{cases} 0 & if\ i = 0\ or\ j = 0 \\ c\,[i-1, j-1] + 1 & if\ i,j > 0\ and\ x_i = y_j \\ \max\{c[i, j-1], c[i-1, j]\} & if\ i,j > 0\ and\ x_i \neq y_j \end{cases}$$

Longest Common Subsequence: Overlapping Subproblems

Longest Common Subsequence: Overlapping Subproblems

Additionally, the number of **distinct** subproblems is relatively **small** (i.e. **polynomial** in problem size).

The number of distinct subproblems is

$$\Theta(mn),$$ where m and n are the lengths of X and Y, respectively .

Therefore, we can use **memoization** or **tabulation** to solve the LCS problem.

Longest Common Subsequence: Bottom-Up

LCS-LENGTH(X, Y)

```
1   m = X.length
2   n = Y.length
3   let b[1..m, 1..n] and c[0..m, 0..n] be new tables
4   for i = 1 to m
5       c[i, 0] = 0
6   for j = 0 to n
7       c[0, j] = 0
8   for i = 1 to m
9       for j = 1 to n
10          if xᵢ == yⱼ
11              c[i, j] = c[i − 1, j − 1] + 1
12              b[i, j] = "↖"
13          elseif c[i − 1, j] ≥ c[i, j − 1]
14              c[i, j] = c[i − 1, j]
15              b[i, j] = "↑"
16          else c[i, j] = c[i, j − 1]
17              b[i, j] = "←"
18  return c and b
```

Longest Common Subsequence: Solution Reconstruction

PRINT-LCS(b, X, i, j)

1   **if** $i == 0$ or $j == 0$
2   **return**
3   **if** $b[i, j] == $ "↖"
4   PRINT-LCS$(b, X, i - 1, j - 1)$
5   print x_i
6   **elseif** $b[i, j] == $ "↑"
7   PRINT-LCS$(b, X, i - 1, j)$
8   **else** PRINT-LCS$(b, X, i, j - 1)$

Summary

We have covered the topic of **Dynamic Programming** using

- Fibonacci numbers
- Matrix Chain Multiplication
- Longest Common Sequence

as examples.

Central to DP, are **optimal substructure** and **overlapping subproblem** properties.

We will cover **Greedy Algorithms** next week.