

CS4102 Algorithms

Spring 2020 – Horton's Slides

Dynamic Programming, Part Deux

- Our mid-term is coming!
- Spring Break's also coming!
- You'll make it! 😊 Hang in there!!!

Midterm

- Wednesday, March 4 in class
 - SDAC: Please schedule with SDAC for Wednesday
 - Mostly in-class with a (required) take-home portion
 - Take-home “bonus” – If you do better on take-home than on its “starter” question on the in-class, you can earn back half the difference.
- Practice Midterm and Solutions on Collab
- Review Session on Panopto
- More office hours by me! See Piazza

Today's Keywords

- Dynamic Programming
- Longest Common Subsequence
- Seam Carving

CLRS Readings

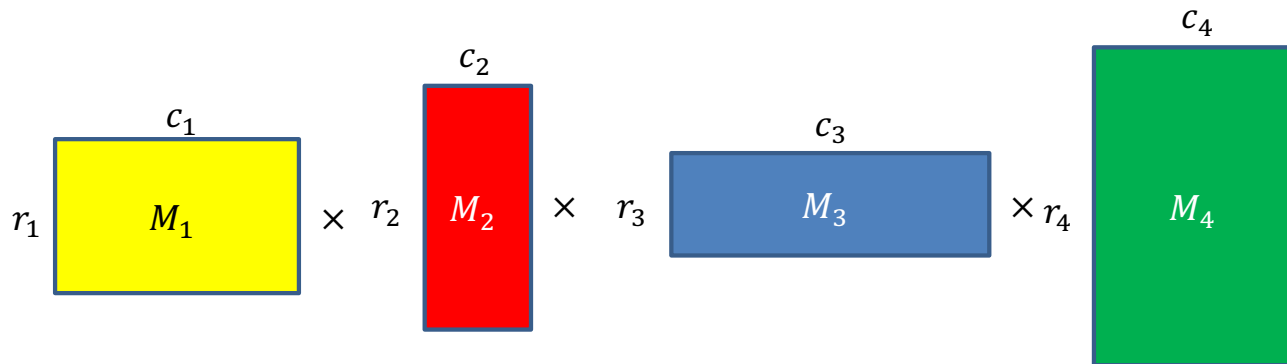
- Chapter 15
 - Section 15.1, Log/Rod cutting, optimal substructure property
 - Note: r_i in book is called Cut() or C[] in our slides. We use their example.
 - Section 15.3, More on elements of DP, including optimal substructure property
 - **Section 15.2, matrix-chain multiplication**
 - **Section 15.4, longest common subsequence**

Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

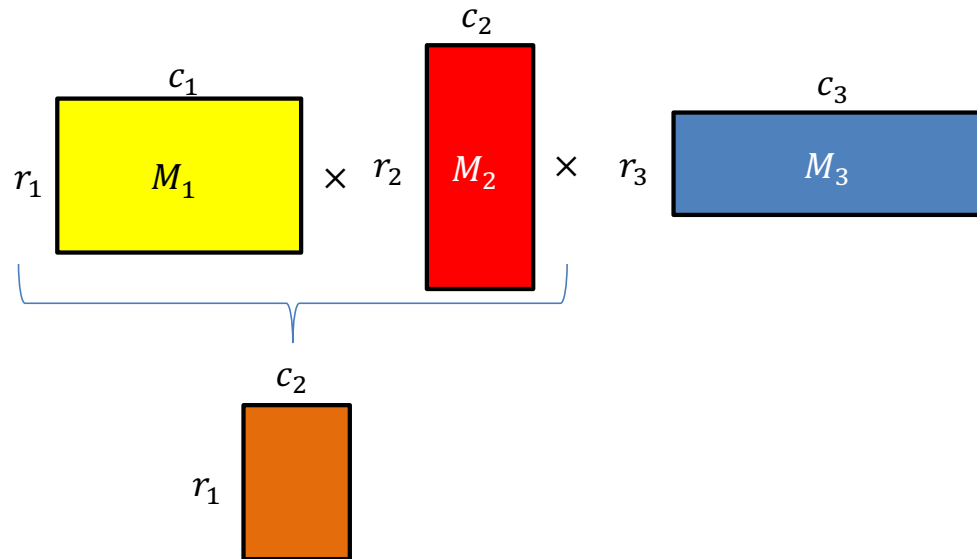
Matrix Chaining

- Given a sequence of Matrices (M_1, \dots, M_n) , what is the most efficient way to multiply them?



Order Matters!

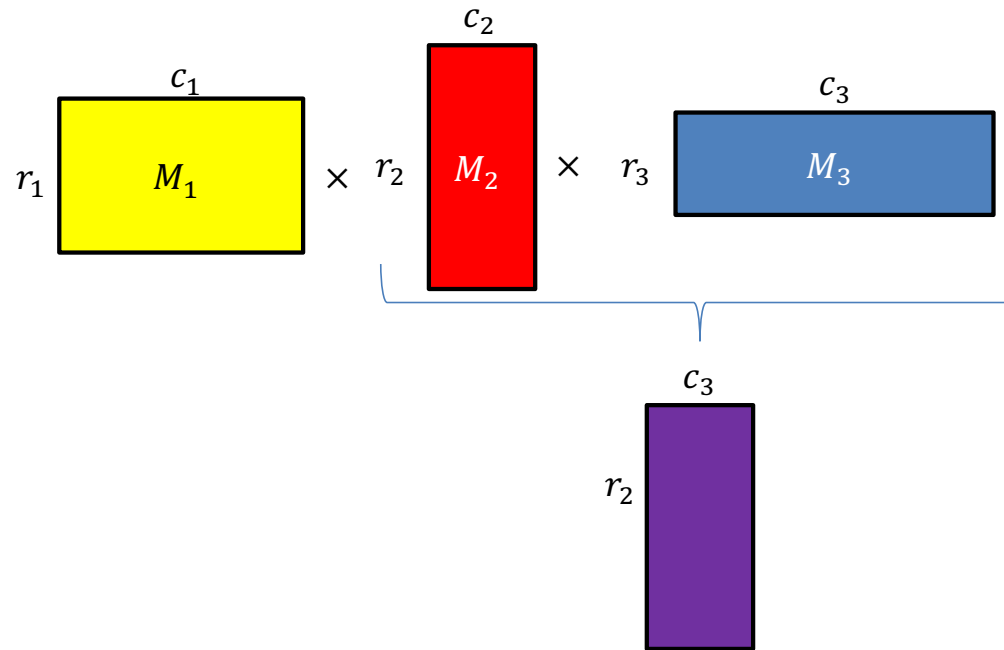
$$c_1 = r_2$$
$$c_2 = r_3$$



- $(M_1 \times M_2) \times M_3$
 - uses $(c_1 \cdot r_1 \cdot c_2) + c_2 \cdot r_1 \cdot c_3$ operations

Order Matters!

$$c_1 = r_2$$
$$c_2 = r_3$$



- $M_1 \times (M_2 \times M_3)$
 - uses $c_1 \cdot r_1 \cdot c_3 + (c_2 \cdot r_2 \cdot c_3)$ operations

Order Matters!

$$c_1 = r_2$$

$$c_2 = r_3$$

- $(M_1 \times M_2) \times M_3$

– uses $(c_1 \cdot r_1 \cdot c_2) + c_2 \cdot r_1 \cdot c_3$ operations

– $(10 \cdot 7 \cdot 20) + 20 \cdot 7 \cdot 8 = 2520$

- $M_1 \times (M_2 \times M_3)$

– uses $c_1 \cdot r_1 \cdot c_3 + (c_2 \cdot r_2 \cdot c_3)$ operations

– $10 \cdot 7 \cdot 8 + (20 \cdot 10 \cdot 8) = 2160$

$$M_1 = 7 \times 10$$

$$M_2 = 10 \times 20$$

$$M_3 = 20 \times 8$$

$$c_1 = 10$$

$$c_2 = 20$$

$$c_3 = 8$$

$$r_1 = 7$$

$$r_2 = 10$$

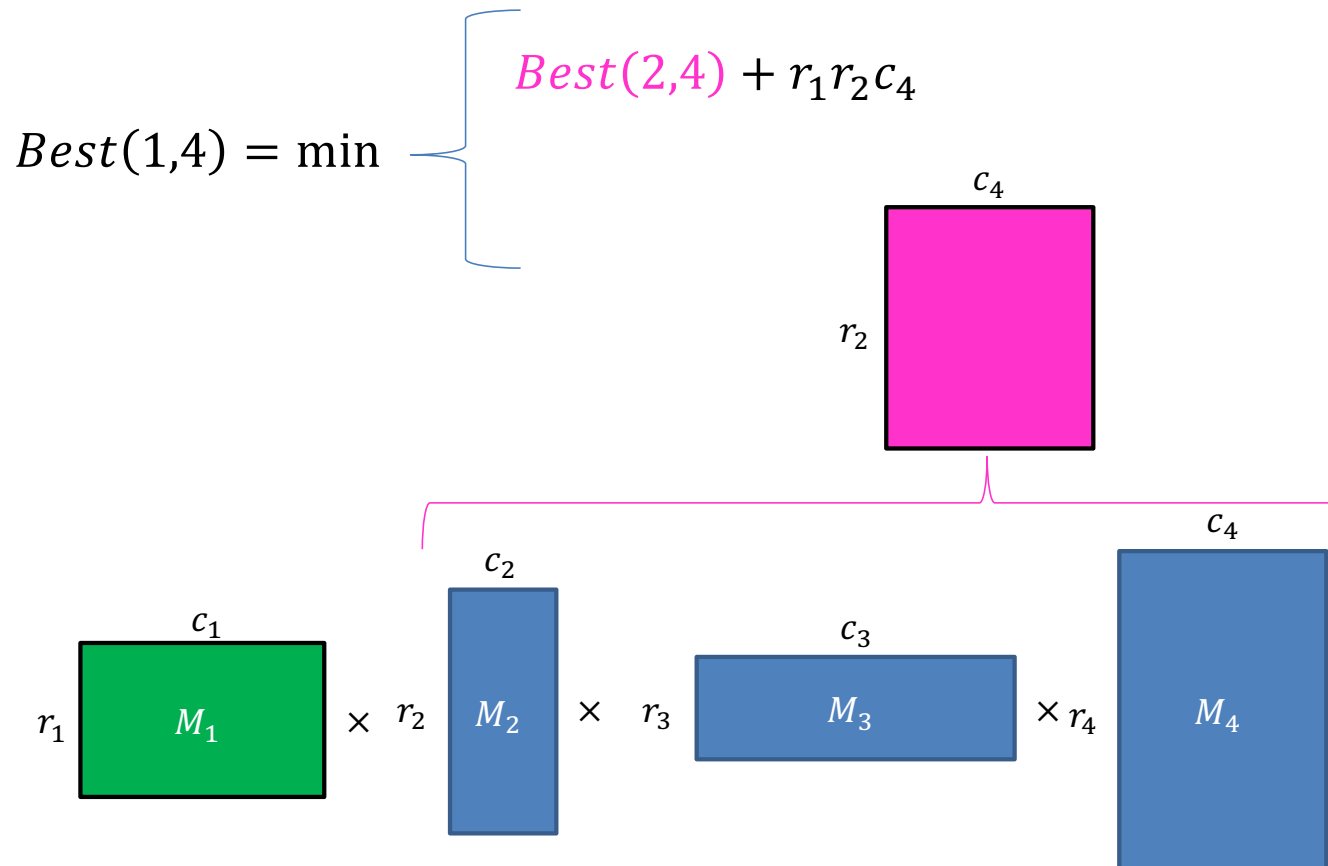
$$r_3 = 20$$

Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

1. Identify the Recursive Structure of the Problem

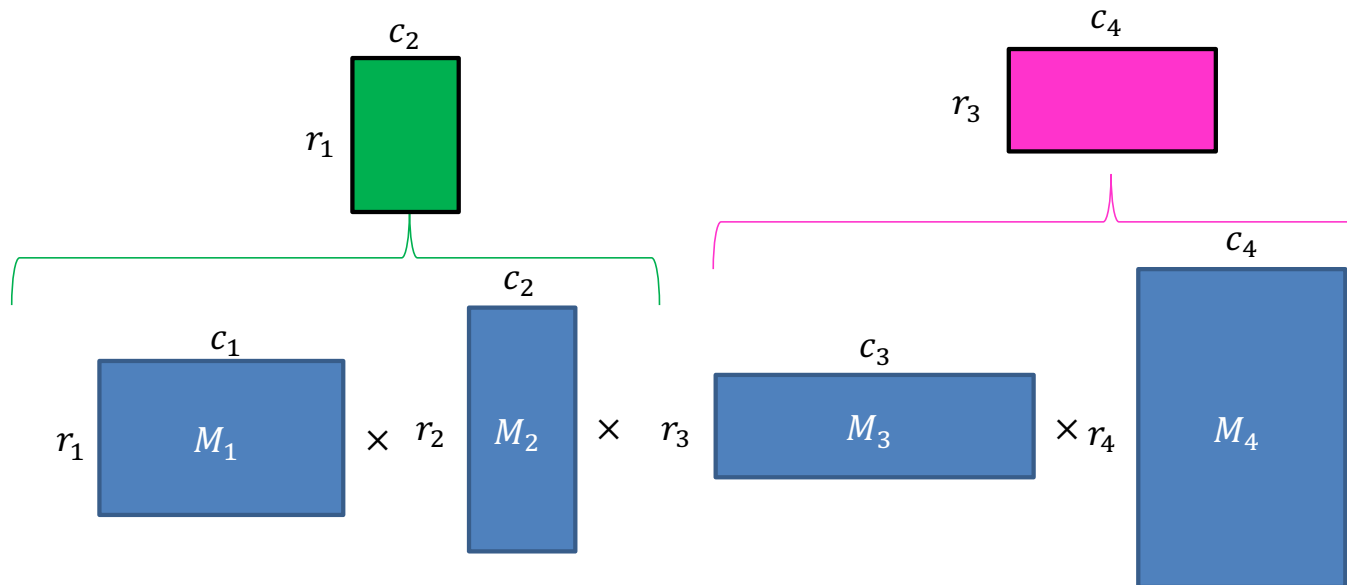
$Best(1, n)$ = cheapest way to multiply together M_1 through M_n



1. Identify the Recursive Structure of the Problem

$Best(1, n)$ = cheapest way to multiply together M_1 through M_n

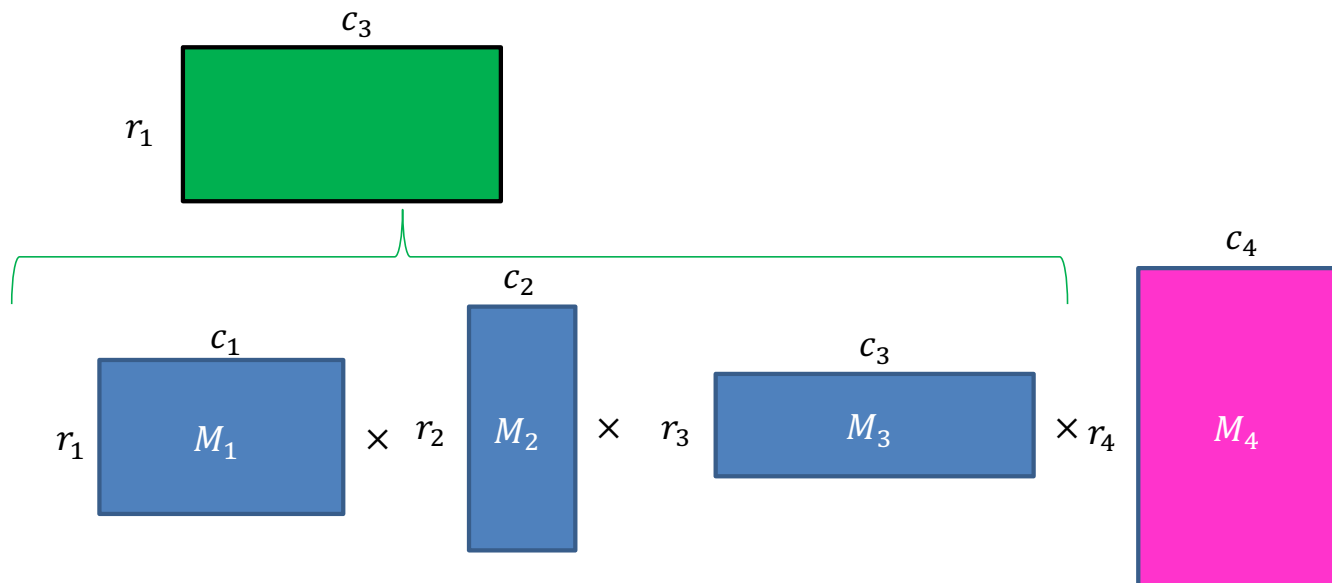
$$Best(1,4) = \min \left\{ \begin{array}{l} Best(2,4) + r_1 r_2 c_4 \\ Best(1,2) + Best(3,4) + r_1 r_3 c_4 \end{array} \right.$$



1. Identify the Recursive Structure of the Problem

$Best(1, n)$ = cheapest way to multiply together M_1 through M_n

$$Best(1,4) = \min \begin{cases} Best(2,4) + r_1 r_2 c_4 \\ Best(1,2) + Best(3,4) + r_1 r_3 c_4 \\ Best(1,3) + r_1 r_4 c_4 \end{cases}$$



1. Identify the Recursive Structure of the Problem

- In general:

$Best(i, j)$ = cheapest way to multiply together M_i through M_j

$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k + 1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

$$Best(1, n) = \min \left\{ \begin{array}{l} Best(2, n) + r_1 r_2 c_n \\ Best(1, 2) + Best(3, n) + r_1 r_3 c_n \\ Best(1, 3) + Best(4, n) + r_1 r_4 c_n \\ Best(1, 4) + Best(5, n) + r_1 r_5 c_n \\ \dots \\ Best(1, n - 1) + r_1 r_n c_n \end{array} \right.$$

Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

2. Save Subsolutions in Memory

- In general:

$Best(i, j)$ = cheapest way to multiply together M_i through M_j

$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

Save to M[n]

Read from M[n]
if present

$$Best(1, n) = \min$$

$$Best(2, n) + r_1 r_2 c_n$$

$$Best(1, 2) + Best(3, n) + r_1 r_3 c_n$$

$$Best(1, 3) + Best(4, n) + r_1 r_4 c_n$$

$$Best(1, 4) + Best(5, n) + r_1 r_5 c_n$$

...

$$Best(1, n-1) + r_1 r_n c_n$$

Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

3. Select a good order for solving subproblems

- In general:

$Best(i, j)$ = cheapest way to multiply together M_i through M_j

$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

Save to M[n]

Read from M[n]
if present

$$Best(1, n) = \min$$

$$Best(2, n) + r_1 r_2 c_n$$

$$Best(1, 2) + Best(3, n) + r_1 r_3 c_n$$

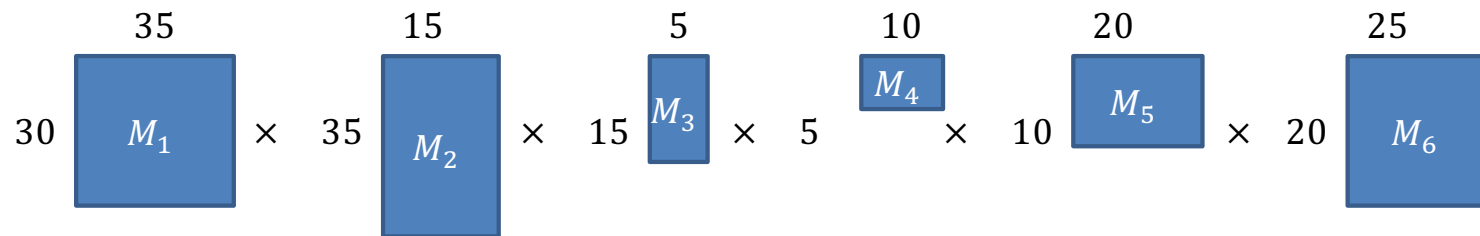
$$Best(1, 3) + Best(4, n) + r_1 r_4 c_n$$

$$Best(1, 4) + Best(5, n) + r_1 r_5 c_n$$

...

$$Best(1, n-1) + r_1 r_n c_n$$

3. Select a good order for solving subproblems

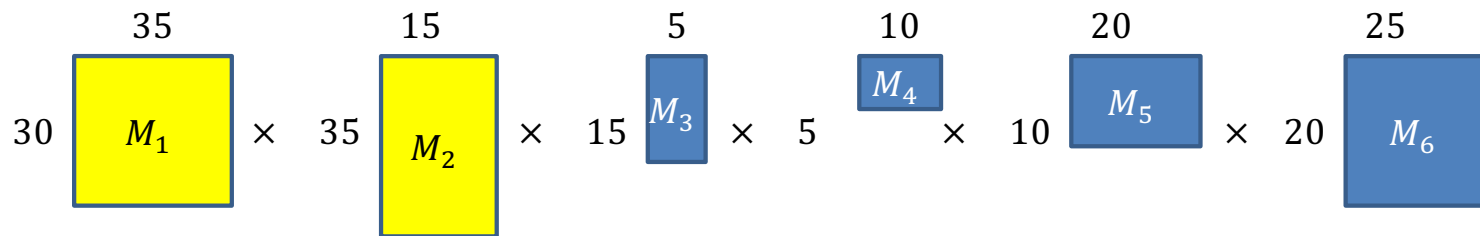


$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	
$i = 1$	0						1
2		0					2
3			0				3
4				0			4
5					0		5
6						0	6

3. Select a good order for solving subproblems



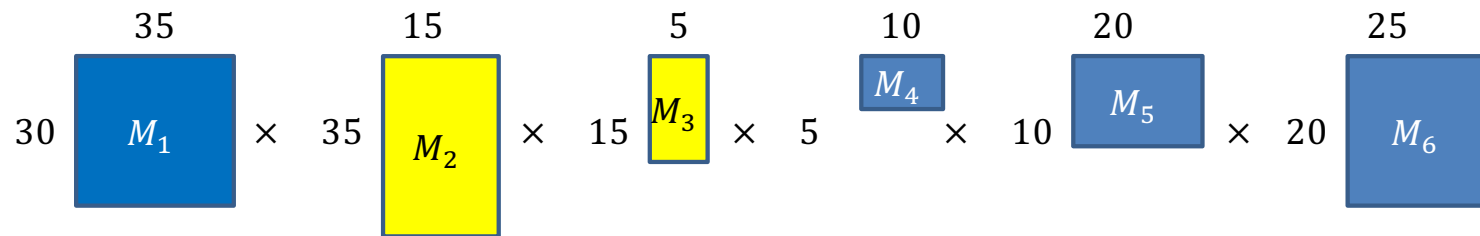
$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	$i =$
1	0	15750					
2		0					
3			0				
4				0			
5					0		
6						0	

$$Best(1, 2) = \min \left\{ Best(1, 1) + Best(2, 2) + r_1 r_2 c_2 \right.$$

3. Select a good order for solving subproblems



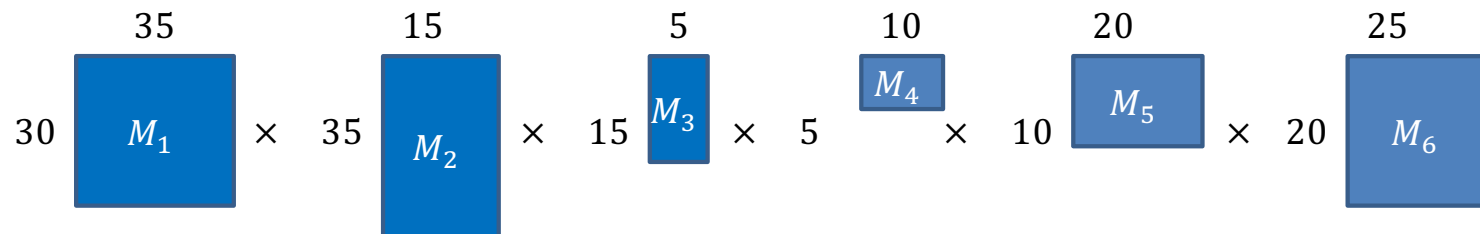
$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	$i =$
1	0	15750					
2		0	2625				
3			0				
4				0			
5					0		
6						0	

$$Best(2,3) = \min \left\{ Best(2,2) + Best(3,3) + r_2 r_3 c_3 \right.$$

3. Select a good order for solving subproblems

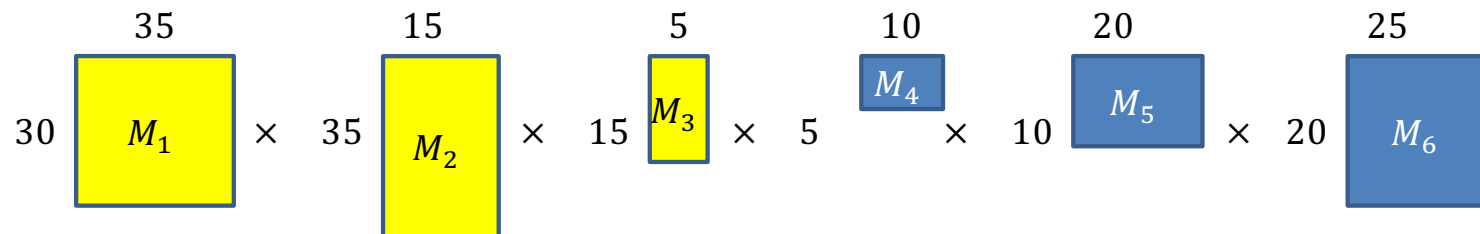


$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	$i =$
1	0	15750					1
2		0	2625				2
3			0	750			3
4				0	1000		4
5					0	5000	5
6						0	6

3. Select a good order for solving subproblems



$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

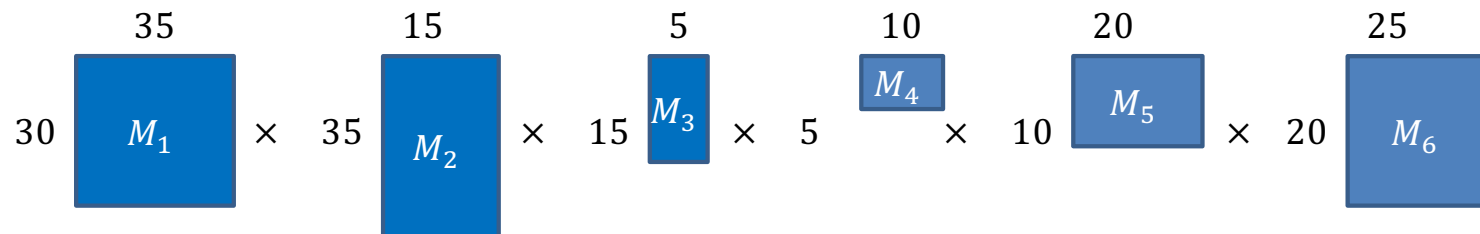
$$r_1 r_2 c_3 = 30 \cdot 35 \cdot 5 = 5250$$

$$r_1 r_3 c_3 = 30 \cdot 15 \cdot 5 = 2250$$

$$Best(1,3) = \min \left\{ \begin{array}{l} 0 \\ Best(1,1) + Best(2,3) + r_1 r_2 c_3 \\ Best(1,2) + Best(3,3) + r_1 r_3 c_3 \\ 15750 \end{array} \right.$$

	$j = 1$	2	3	4	5	6	$= i$
1	0	15750	7875				1
2		0	2625				2
3			0	750			3
4				0	1000		4
5					0	5000	5
6						0	6

3. Select a good order for solving subproblems



$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

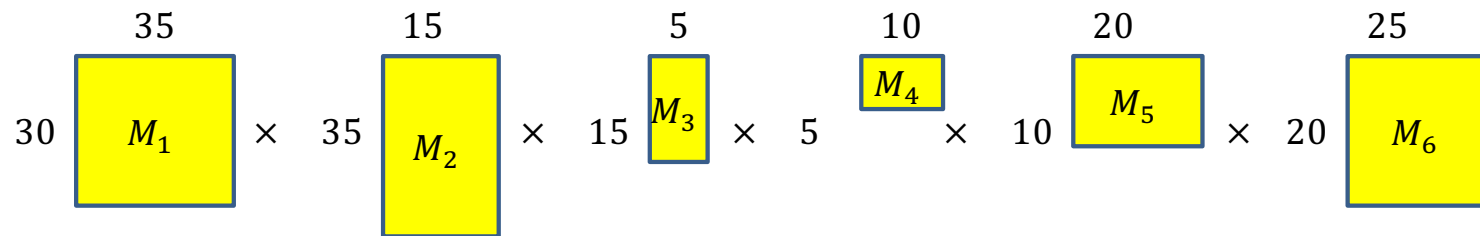
$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	$i =$
1	0	15750	7875				1
2		0	2625				2
3			0	750			3
4				0	1000		4
5					0	5000	5
6						0	6

To find $Best(i, j)$: Need all preceding terms of row i and column j

Conclusion: solve in order of diagonal

Matrix Chaining



$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	$i =$
	0	15750	7875	9375	11875	15125	1
		0	2625	4375	7125	10500	2
			0	750	2500	5375	3
				0	1000	3500	4
					0	5000	5
						0	6

$Best(1,6) = \min$

- $Best(1,1) + Best(2,6) + r_1 r_2 c_6$
- $Best(1,2) + Best(3,6) + r_1 r_3 c_6$
- $Best(1,3) + Best(4,6) + r_1 r_4 c_6$
- $Best(1,4) + Best(5,6) + r_1 r_5 c_6$
- $Best(1,5) + Best(6,6) + r_1 r_6 c_6$

Run Time

1. Initialize $Best[i, i]$ to be all 0s $\Theta(n^2)$ cells in the Array
2. Starting at the main diagonal, working to the upper-right, fill in each cell using:

1. $Best[i, i] = 0$

$\Theta(n)$ options for each cell

Each "call" to Best() is a $O(1)$ memory lookup

2. $Best[i, j] = \min_{k=i}^{j-1} (Best(i, k) + Best(k + 1, j) + r_i r_{k+1} c_j)$

$\Theta(n^3)$ overall run time

Backtrack to find the best order

“Remember” which choice of k was the minimum at each cell. Intuitively this was the best place to “split” for that range (i,j) .

$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

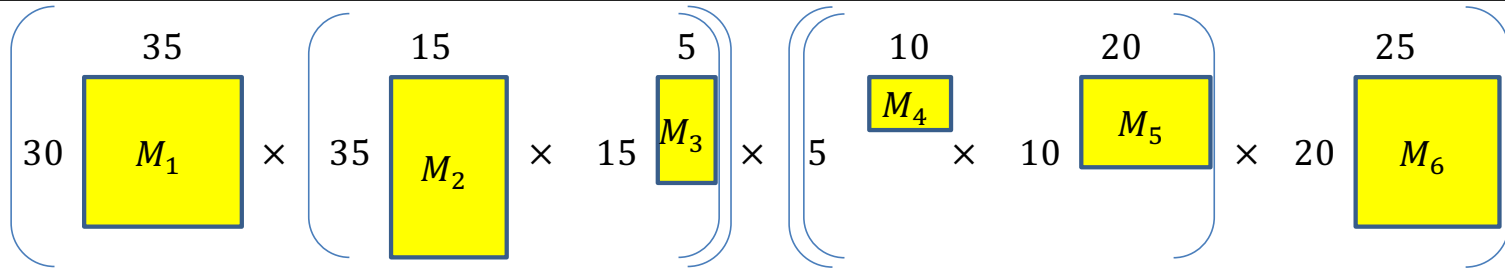
$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	$= i$
	0	15750	7875	9375	11875	15125	1
		0	2625	4375	7125	10500	2
			0	750	2500	5375	3
				0	1000	3500	4
					0	5000	5
						0	6

$Best(1,6) = \min$

- $Best(1,1) + Best(2,6) + r_1 r_2 c_6$
- $Best(1,2) + Best(3,6) + r_1 r_3 c_6$
- $Best(1,3) + Best(4,6) + r_1 r_4 c_6$
- $Best(1,4) + Best(5,6) + r_1 r_5 c_6$
- $Best(1,5) + Best(6,6) + r_1 r_6 c_6$

Matrix Chaining



$$Best(i, j) = \min_{k=i}^{j-1} (Best(i, k) + Best(k+1, j) + r_i r_{k+1} c_j)$$

$$Best(i, i) = 0$$

	$j = 1$	2	3	4	5	6	$= i$
1	0	15750	7875 ₁	9375	11875	15125 ₃	1
2		0	2625	4375	7125	10500	2
3			0	750	2500	5375	3
4				0	1000	3500 ₅	4
5					0	5000	5
6						0	6

$Best(1,6) = \min$

- $Best(1,1) + Best(2,6) + r_1 r_2 c_6$
- $Best(1,2) + Best(3,6) + r_1 r_3 c_6$
- $Best(1,3) + Best(4,6) + r_1 r_4 c_6$
- $Best(1,4) + Best(5,6) + r_1 r_5 c_6$
- $Best(1,5) + Best(6,6) + r_1 r_6 c_6$

Storing and Recovering Optimal Solution

- Maintain table **Choice**[i,j] in addition to **Best** table
 - **Choice**[i,j] = k means the best “split” was right after M_k
 - Work backwards from value for whole problem, **Choice**[1,n]
 - Note: **Choice**[i,i+1] = i because there are just 2 matrices
- From our example:
 - **Choice**[1,6] = 3. So $[M_1 M_2 M_3] [M_4 M_5 M_6]$
 - We then need **Choice**[1,3] = 1. So $[(M_1) (M_2 M_3)]$
 - Also need **Choice**[4,6] = 5. So $[(M_4 M_5) M_6]$
 - Overall: $[(M_1) (M_2 M_3)] [(M_4 M_5) M_6]$

Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

Longest Common Subsequence

Given two sequences X and Y , find the length of their longest common subsequence

Example:

$X = ATCTGAT$

$Y = TGCATA$

$LCS = TCTA$

$X = AT \ C \ TGAT$

$Y = \ TGCAT \ A$

Brute force: Compare every subsequence of X with Y : $\Omega(2^n)$



Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

1. Identify Recursive Structure

Let $LCS(i, j)$ = length of the LCS for the first i characters of X , first j character of Y

Find $LCS(i, j)$:

Case 1: $X[i] = Y[j]$

$X = ATCTGCGT$

$Y = TGCATAT$

$$LCS(i, j) = LCS(i - 1, j - 1) + 1$$

Case 2: $X[i] \neq Y[j]$

$X = ATCTGCGA$

$Y = TGCATAG$

$$LCS(i, j) = LCS(i, j - 1)$$

$X = ATCTGCGA$

$Y = TGCATAG$

$$LCS(i, j) = LCS(i - 1, j)$$

$$LCS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i - 1, j - 1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j - 1), LCS(i - 1, j)) & \text{otherwise} \end{cases}$$

Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

1. Identify Recursive Structure

Let $LCS(i, j)$ = length of the LCS for the first i characters of X , first j character of Y

Find $LCS(i, j)$:

Case 1: $X[i] = Y[j]$

$X = ATCTGCGT$

$Y = TGCATAT$

$$LCS(i, j) = LCS(i - 1, j - 1) + 1$$

Case 2: $X[i] \neq Y[j]$

$X = ATCTGCGA$

$Y = TGCATAT$

$$LCS(i, j) = LCS(i, j - 1)$$

$X = ATCTGCGT$

$Y = TGCATAC$

$$LCS(i, j) = LCS(i - 1, j)$$

$$LCS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i - 1, j - 1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j - 1), LCS(i - 1, j)) & \text{otherwise} \end{cases}$$

↑ Save to $M[i, j]$
↖ Read from $M[i, j]$ if present

Dynamic Programming

- Requires **Optimal Substructure**
 - Solution to larger problem contains the solutions to smaller ones
- Avoid extra work due to **overlapping subproblems**
- Idea:
 1. Identify the recursive structure of the problem
 - What is the “last thing” done?
 2. Save the solution to each subproblem in memory
 3. Select a good order for solving subproblems
 - “Top Down”: Solve each recursively
 - “Bottom Up”: Iteratively solve smallest to largest

3. Solve in a Good Order

$$LCS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i - 1, j - 1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j - 1), LCS(i - 1, j)) & \text{otherwise} \end{cases}$$

$X =$			A	T	C	T	G	A	T
$Y =$		0	1	2	3	4	5	6	7
0		0	0	0	0	0	0	0	0
T	1	0	0	1	1	1	1	1	1
G	2	0	0	1	1	1	2	2	2
C	3	0	0	1	2	2	2	2	2
A	4	0	1	1	2	2	2	3	3
T	5	0	1	2	2	3	3	3	4
A	6	0	1	2	2	3	3	4	4

To fill in cell (i, j) we need cells $(i - 1, j - 1)$, $(i - 1, j)$, $(i, j - 1)$
 Fill from Top->Bottom, Left->Right (with any preference)

Run Time?

$$LCS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i - 1, j - 1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j - 1), LCS(i - 1, j)) & \text{otherwise} \end{cases}$$

$X =$			A	T	C	T	G	A	T
$Y =$		0	1	2	3	4	5	6	7
	0	0	0	0	0	0	0	0	0
T	1	0	0	1	1	1	1	1	1
G	2	0	0	1	1	1	2	2	2
C	3	0	0	1	2	2	2	2	2
A	4	0	1	1	2	2	2	3	3
T	5	0	1	2	2	3	3	3	4
A	6	0	1	2	2	3	3	4	4

Run Time: $\Theta(n \cdot m)$ (for $|X| = n, |Y| = m$)

Reconstructing the LCS

$$LCS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i - 1, j - 1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j - 1), LCS(i - 1, j)) & \text{otherwise} \end{cases}$$

X =			A	T	C	T	G	A	T
Y =		0	1	2	3	4	5	6	7
	0	0	0	0	0	0	0	0	0
T	1	0	0	1	1	1	1	1	1
G	2	0	0	1	1	1	2	2	2
C	3	0	0	1	2	2	2	2	2
A	4	0	1	1	2	2	2	3	3
T	5	0	1	2	2	3	3	3	4
A	6	0	1	2	2	3	3	4	4

Start from bottom right,
 if symbols matched, print that symbol then go diagonally
 else go to largest adjacent

Reconstructing the LCS

$$LCS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i - 1, j - 1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j - 1), LCS(i - 1, j)) & \text{otherwise} \end{cases}$$

	X =		A	T	C	T	G	A	T
Y =		0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0	0
T	1	0	0	1	1	1	1	1	1
G	2	0	0	1	1	1	2	2	2
C	3	0	0	1	2	2	2	2	2
A	4	0	1	1	2	2	2	3	3
T	5	0	1	2	2	3	3	3	4
A	6	0	1	2	2	3	3	4	4

Start from bottom right,
 if symbols matched, print that symbol then go diagonally
 else go to largest adjacent

Reconstructing the LCS

$$LCS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i - 1, j - 1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j - 1), LCS(i - 1, j)) & \text{otherwise} \end{cases}$$

	X =		A	T	C	T	G	A	T
Y =		0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0	0
T	1	0	0	1	1	1	1	1	1
G	2	0	0	1	1	1	2	2	2
C	3	0	0	1	2	2	2	2	2
A	4	0	1	1	2	2	2	3	3
T	5	0	1	2	2	3	3	3	4
A	6	0	1	2	2	3	3	4	4

Start from bottom right,
 if symbols matched, print that symbol then go diagonally
 else go to largest adjacent