CS4102 Algorithms

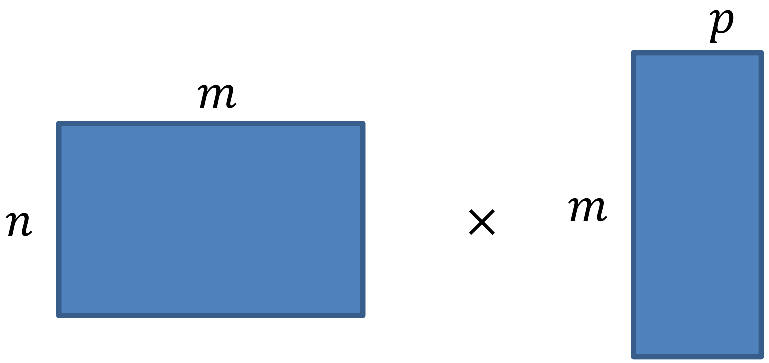
Nate Brunelle

Spring 2018

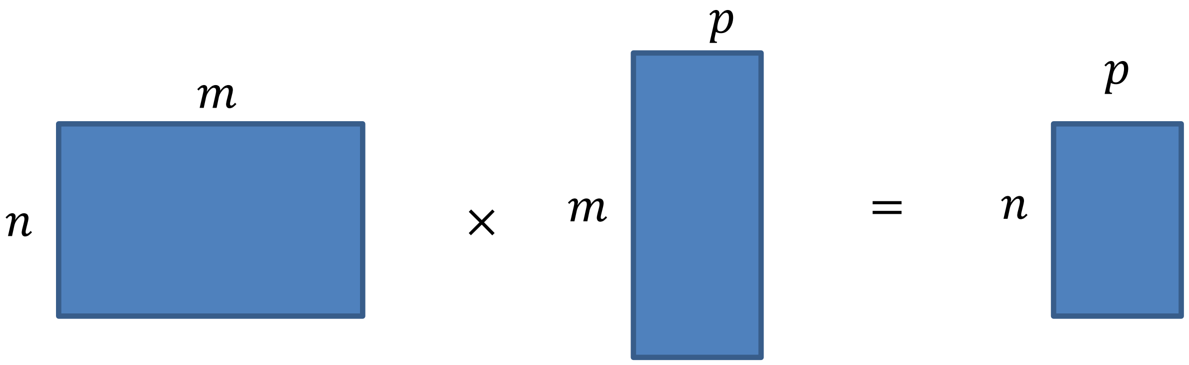
Warm up

How many arithmetic operations are required to multiply a $n \times m$ Matrix with a $m \times p$ Matrix?

(don't overthink this)



How many arithmetic operations are required to multiply a $n \times m$ Matrix with a $m \times p$ Matrix?



- m multiplications and additions per element
- $n \cdot p$ elements to compute
- Total cost: $m \cdot n \cdot p$

Today’s Keywords

- Dynamic Programming
- Matrix Chaining
- Longest Common Subsequence

CLRS Readings

- Chapter 15

Homeworks

- Hw4 due 11pm Friday March 16
  - Sorting
  - Written

Midterm

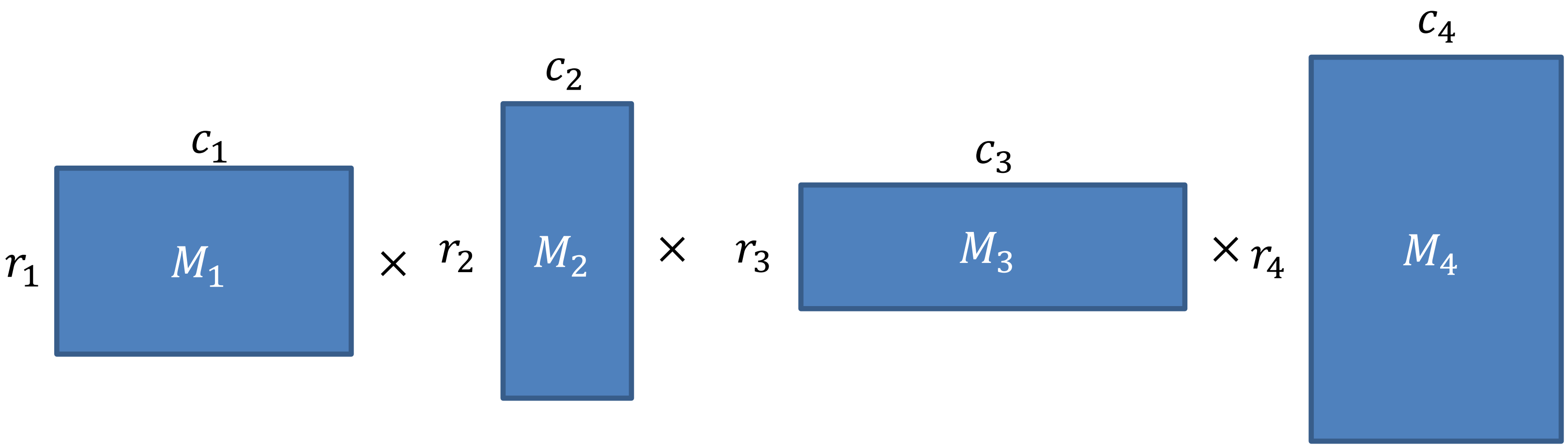
- Monday March 19 in class
  - Covers all content through sorting
  - We will have a review session the weekend before

Dynamic Programming

- Requires Optimal Substructure
  - Solution to larger problem contains the solutions to smaller ones
- Idea:
  1. Identify recursive structure of the problem
     - What is the “last thing” done?
  2. Select a good order for solving subproblems
     - Usually smallest problem first

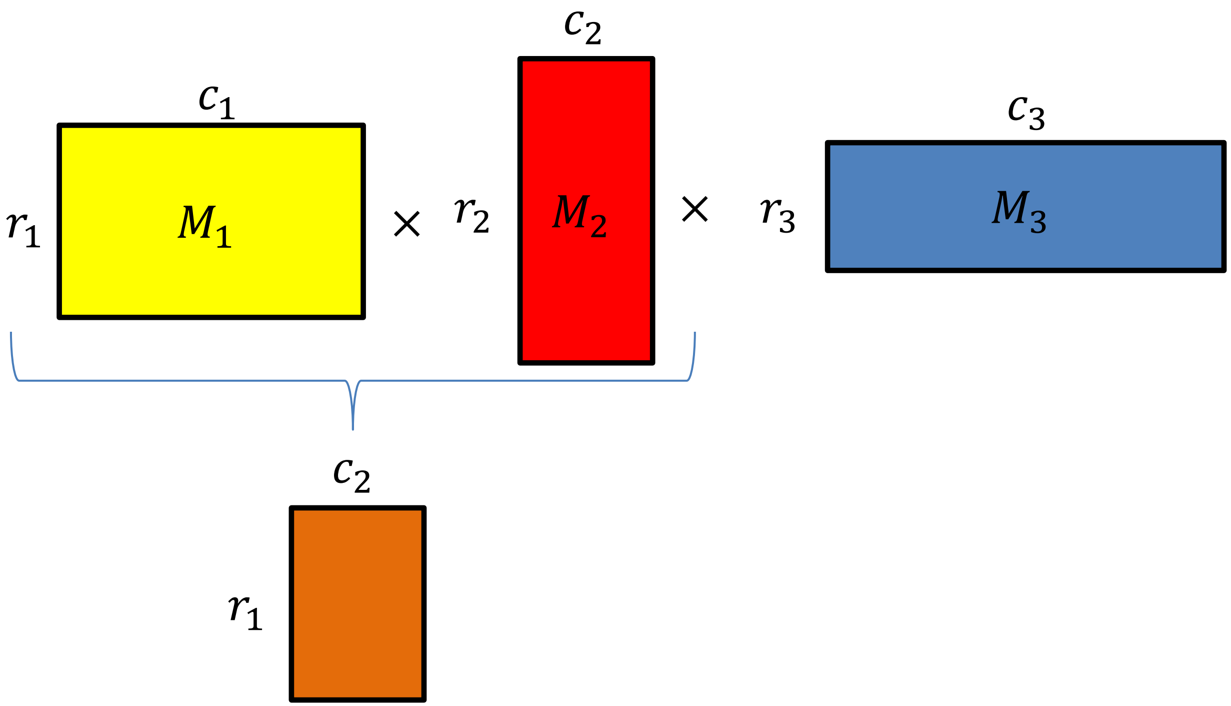
Matrix Chaining

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



$c_1 = r_2$
$c_2 = r_3$

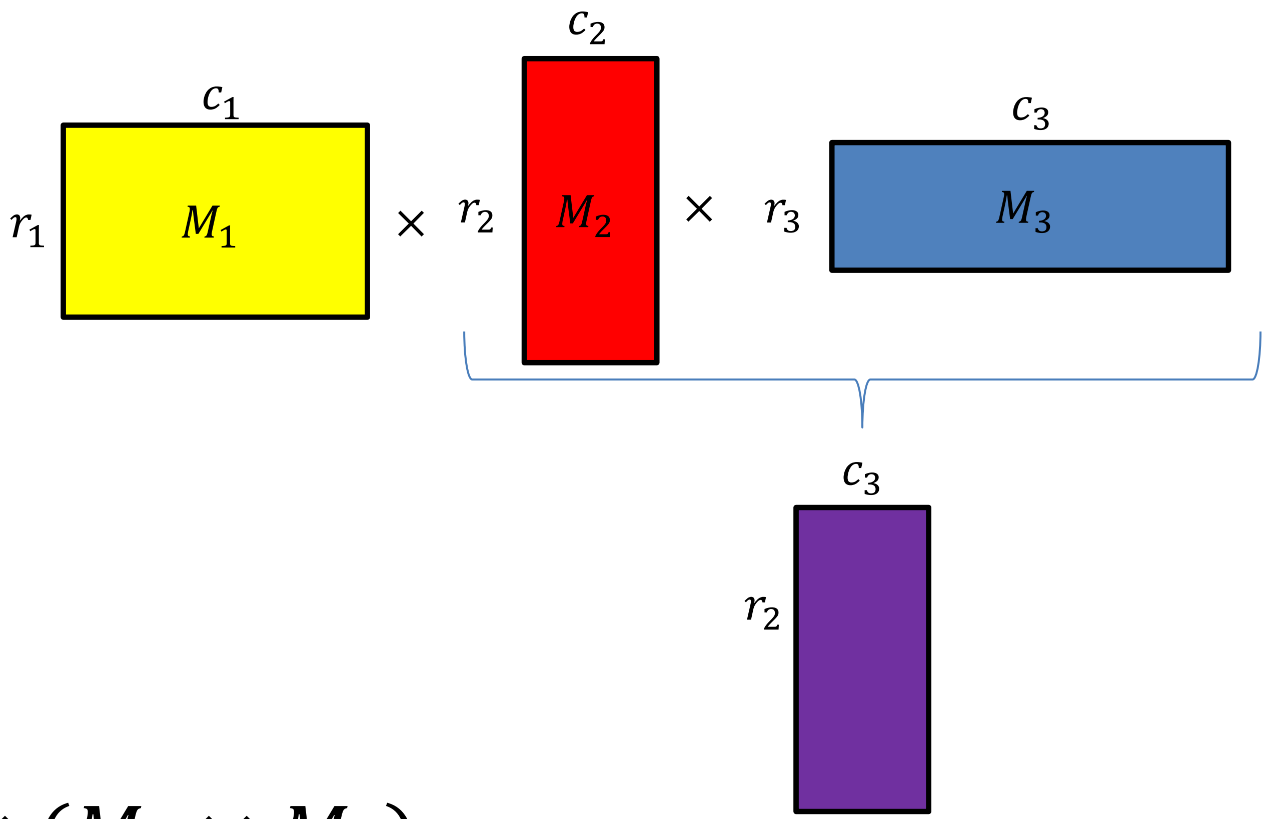
Order Matters!



- $(M_1 \times M_2) \times M_3$
  - uses $(c_1 \cdot r_1 \cdot c_2) + \mathrm{c}_2 \cdot r_1 \cdot c_3$ operations

$c_1 = r_2$
$c_2 = r_3$

Order Matters!



- $M_1 \times (M_2 \times M_3)$
  - uses $\mathrm{c}_1 \cdot \mathrm{r}_1 \cdot c_3 + (\mathrm{c}_2 \cdot r_2 \cdot c_3)$ operations

$c_1 = r_2$
$c_2 = r_3$

$c_1 = 10$
$c_2 = 20$
$c_3 = 8$
$r_1 = 7$
$r_2 = 10$
$r_3 = 20$

Order Matters!

- $(M_1 \times M_2) \times M_3$
  - uses $(c_1 \cdot r_1 \cdot c_2) + \mathrm{c}_2 \cdot r_1 \cdot c_3$ operations
  - $(10 \cdot 7 \cdot 20) + 20 \cdot 7 \cdot 8 = 2520$
- $M_1 \times (M_2 \times M_3)$
  - uses $c_1 \cdot r_1 \cdot c_3 + (\mathrm{c}_2 \cdot r_2 \cdot c_3)$ operations
  - $10 \cdot 7 \cdot 8 + (20 \cdot 10 \cdot 8) = 2160$

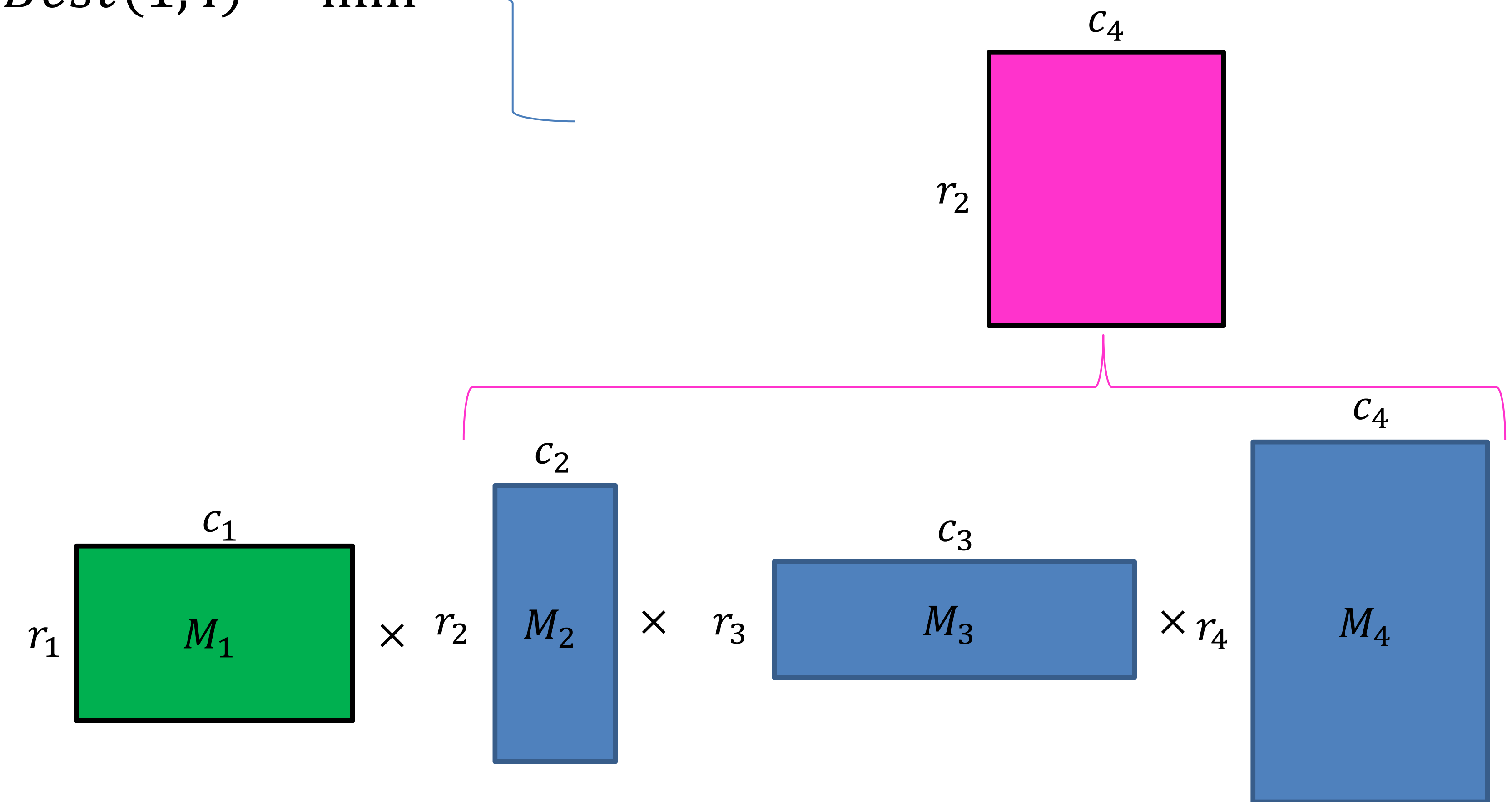
Dynamic Programming

- Idea:
  1. Identify recursive structure of the problem
     - What is the "last thing" done?
  2. Select a good order for solving subproblems
     - Usually smallest problem first
     - "Bottom up"

1. Identify the Recursive Structure of the Problem

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

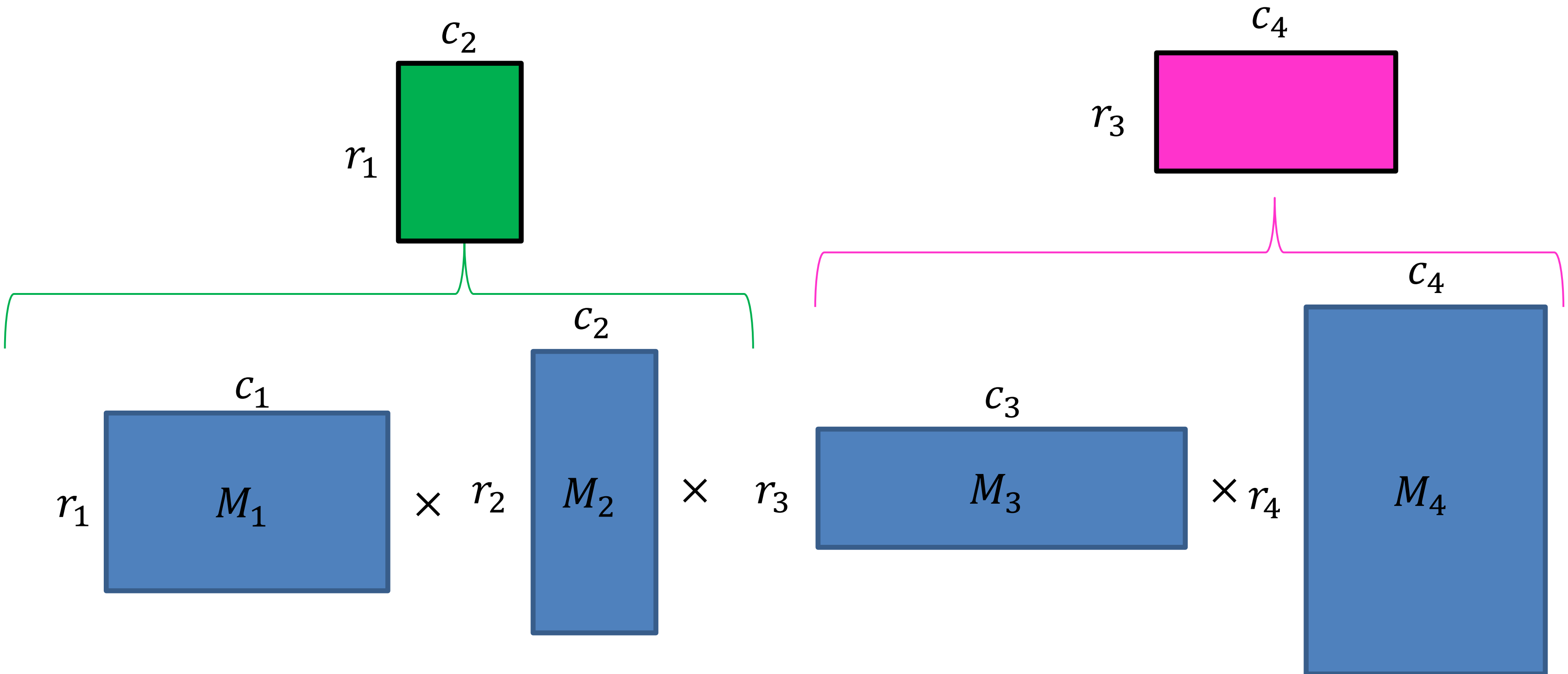
$$Best(1,4) = \min \begin{cases} Best(2,4) + r_1 r_2 c_4 \\ \\ \\ \end{cases}$$



1. Identify the Recursive Structure of the Problem

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

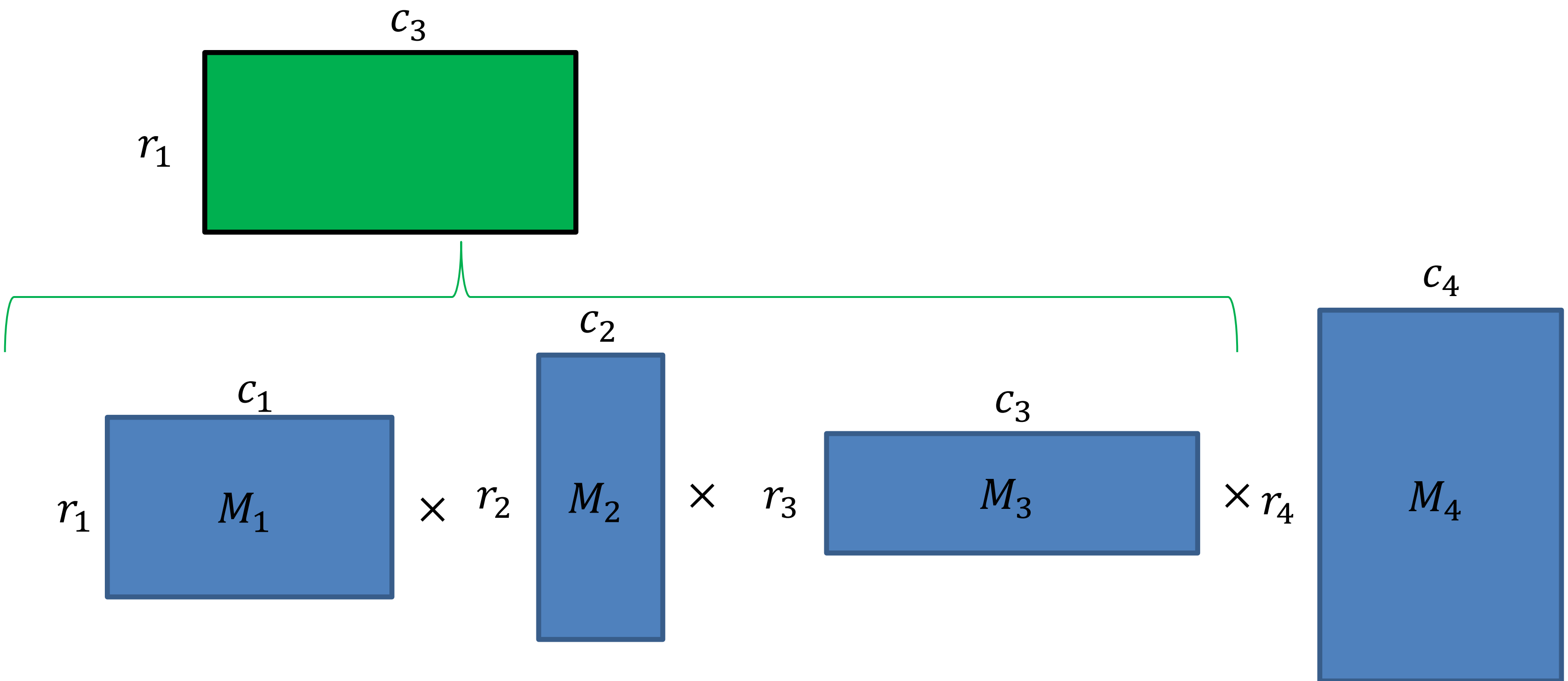
$$Best(1{,}4) = \min \begin{cases} Best(2{,}4) + r_1 r_2 c_4 \\ Best(1{,}2) + Best(3,4) + r_1 r_3 c_4 \\ \end{cases}$$



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:

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

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

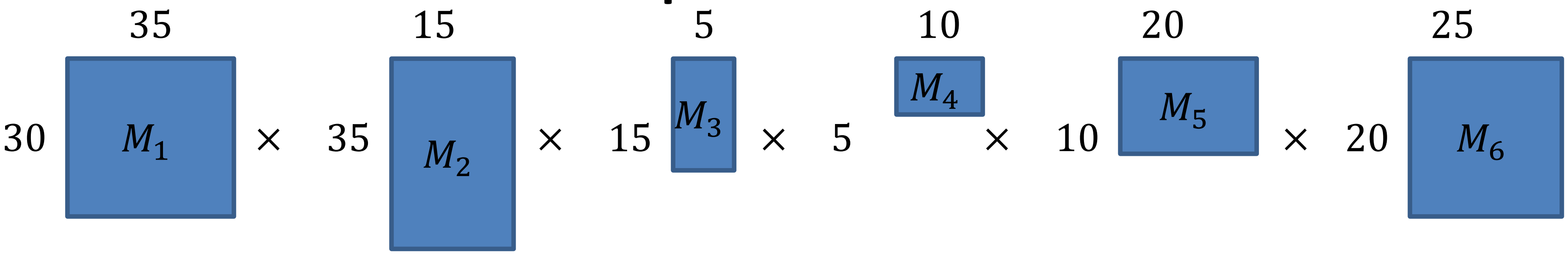
$$Best(i,j) = \min_{k=i}^{j-1} \left(Best(i,k) + Best(k+1,j) + r_i r_{k+1} c_j \right)$$

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

Dynamic Programming

- Idea:
  1. Identify recursive structure of the problem
     - What is the "last thing" done?
  2. Select a good order for solving subproblems
     - Usually smallest problem first
     - "Bottom up"

2. Select a good order for solving subproblems

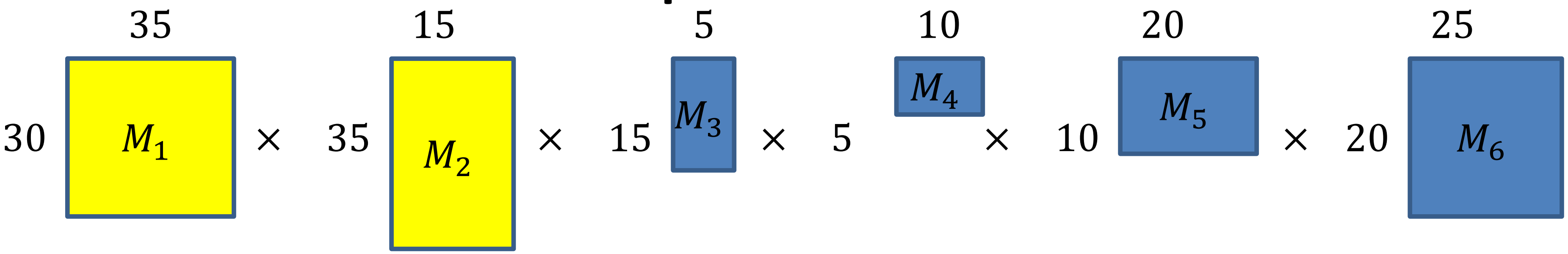


$$Best(i,j) = \min_{k=i}^{j-1}\left(Best(i,k) + Best(k+1,j) + r_i r_{k+1} c_j\right)$$

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

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

2. Select a good order for solving subproblems



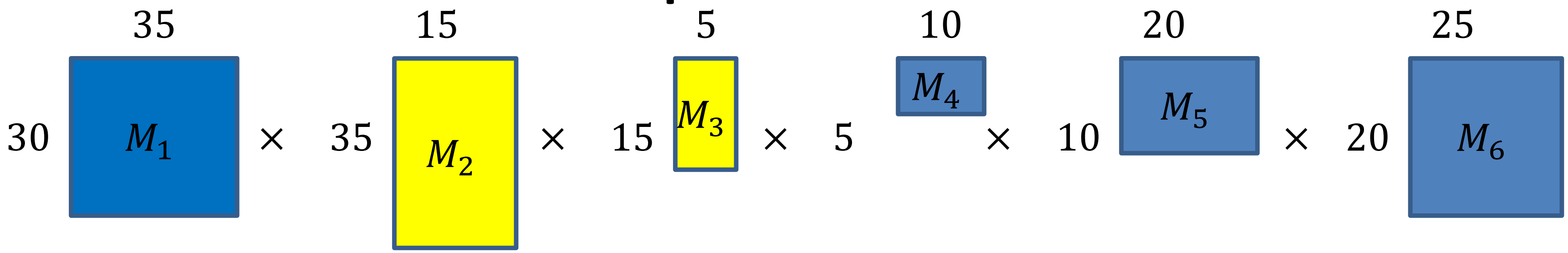
$$Best(i,j) = \min_{k=i}^{j-1}\left(Best(i,k) + Best(k+1,j) + r_i r_{k+1} c_j\right)$$

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

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

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

2. Select a good order for solving subproblems



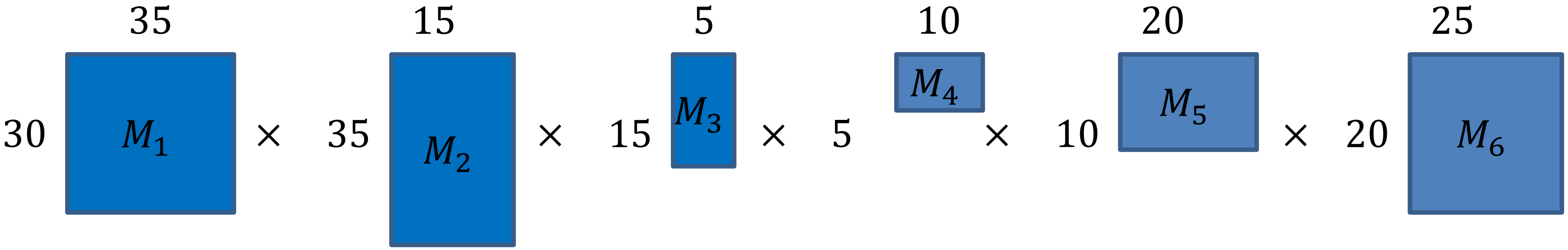
$$Best(i,j) = \min_{k=i}^{j-1}\left(Best(i,k) + Best(k+1,j) + r_i r_{k+1} c_j\right)$$

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

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

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

2. Select a good order for solving subproblems

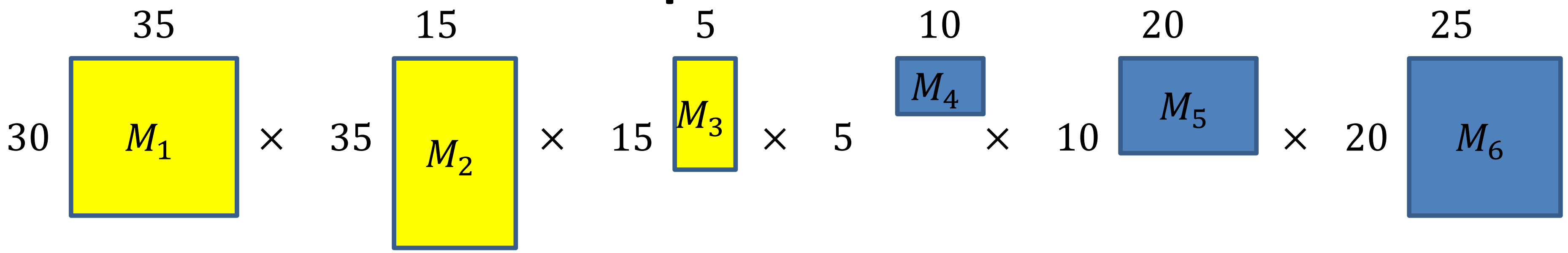


$$Best(i,j) = \min_{k=i}^{j-1}\left({\color{green}Best(i,k)} + {\color{magenta}Best(k+1,j)} + r_i r_{k+1} c_j \right)$$

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

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

2. Select a good order for solving subproblems



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

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

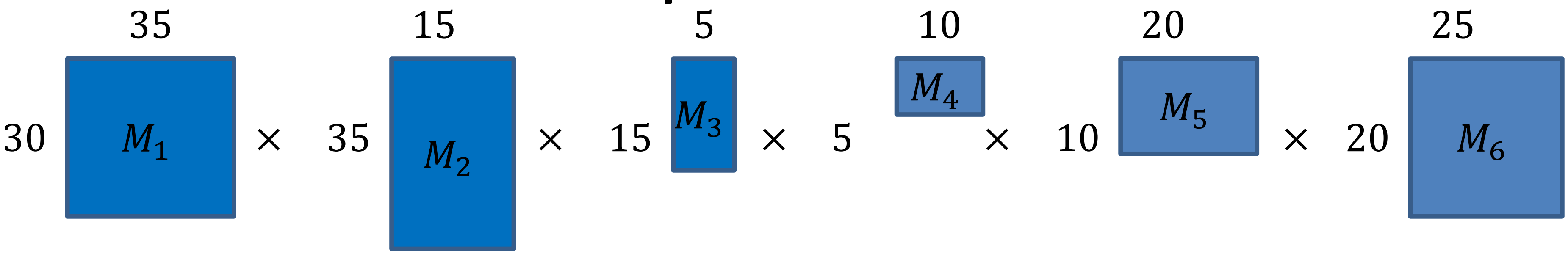
$j =$	1	2	3	4	5	6	$= i$
	0	15750	7875				1
		0	2625				2
			0	750			3
				0	1000		4
					0	5000	5
						0	6

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

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

$$Best(1,3) = \min\begin{cases} \overset{0}{Best(1,1)} + \overset{2625}{Best(2,3)} + r_1 r_2 c_3 \\ \underset{15750}{Best(1,2)} + \underset{0}{Best(3,3)} + r_1 r_3 c_3 \end{cases}$$

2. Select a good order for solving subproblems



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

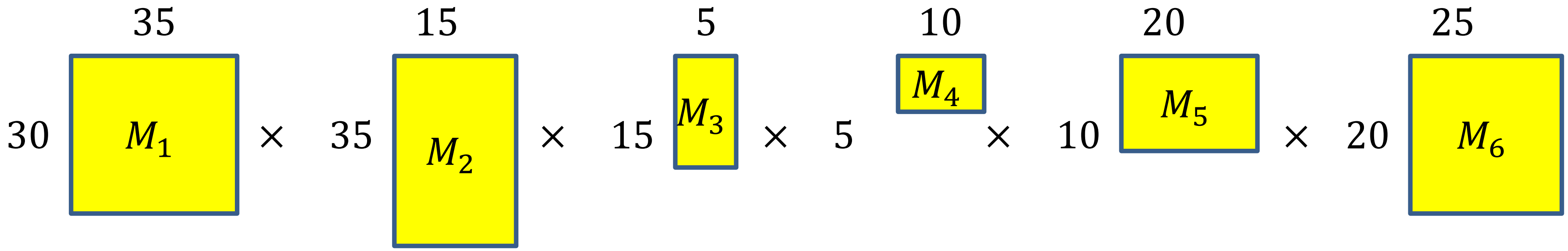
$$Best(i,i) = 0$$

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

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

Conclusion: solve in order of diagonal

Longest Common Subsequence



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

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

$j =$ 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

$$Best(1,6) = \min\begin{cases} Best(1,1) + Best(2,6) + r_1 r_2 c_6 \\ Best(1,2) + Best(3,6) + r_1 r_3 c_6 \\ Best(1,3) + Best(4,6) + r_1 r_4 c_6 \\ Best(1,4) + Best(5,6) + r_1 r_5 c_6 \\ Best(1,5) + Best(6,6) + r_1 r_6 c_6 \end{cases}$$

Run Time

1. Initialize $Best[i,i]$ to be all 0s
2. Starting at the main diagonal, working to the upper-right, fill in each cell using:
   1. $Best[i,i] = 0$
   2. $Best[i,j] = \min_{k=i}^{j-1}\left(Best(i,k) + Best(k+1,j) + r_i r_{k+1} c_j\right)$

$\Theta(n^2)$ cells in the Array

$\Theta(n)$ options for each cell

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

Backtrack to find the best order

“remember” which choice of k was the minimum at each cell

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

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

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

$$Best(1,6) = \min \begin{cases} Best(1,1) + Best(2,6) + r_1 r_2 c_6 \\ Best(1,2) + Best(3,6) + r_1 r_3 c_6 \\ \boxed{Best(1,3) + Best(4,6) + r_1 r_4 c_6} \\ Best(1,4) + Best(5,6) + r_1 r_5 c_6 \\ Best(1,5) + Best(6,6) + r_1 r_6 c_6 \end{cases}$$

Longest Common Subsequence

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

Example:
$X = ATCTGAT$
$Y = TGCATA$
$LCS = TCTA$

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

Dynamic Programming

- Idea:
  1. Identify recursive structure of the problem
     - What is the "last thing" done?
  2. Select a good order for solving subproblems
     - Usually smallest problem first
     - "Bottom up"

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

$$X = ATCTGCGA$$
$$Y = TGCATAT$$
$$LCS(i,j) = LCS(i,j-1)$$

$$X = ATCTGCGG$$
$$Y = TGCATAC$$
$$LCS(i,j) = LCS(i-1,j)$$

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

Dynamic Programming

- Idea:
  1. Identify recursive structure of the problem
     - What is the “last thing” done?
  2. Select a good order for solving subproblems
     - Usually smallest problem first
     - “Bottom up”

2. 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?

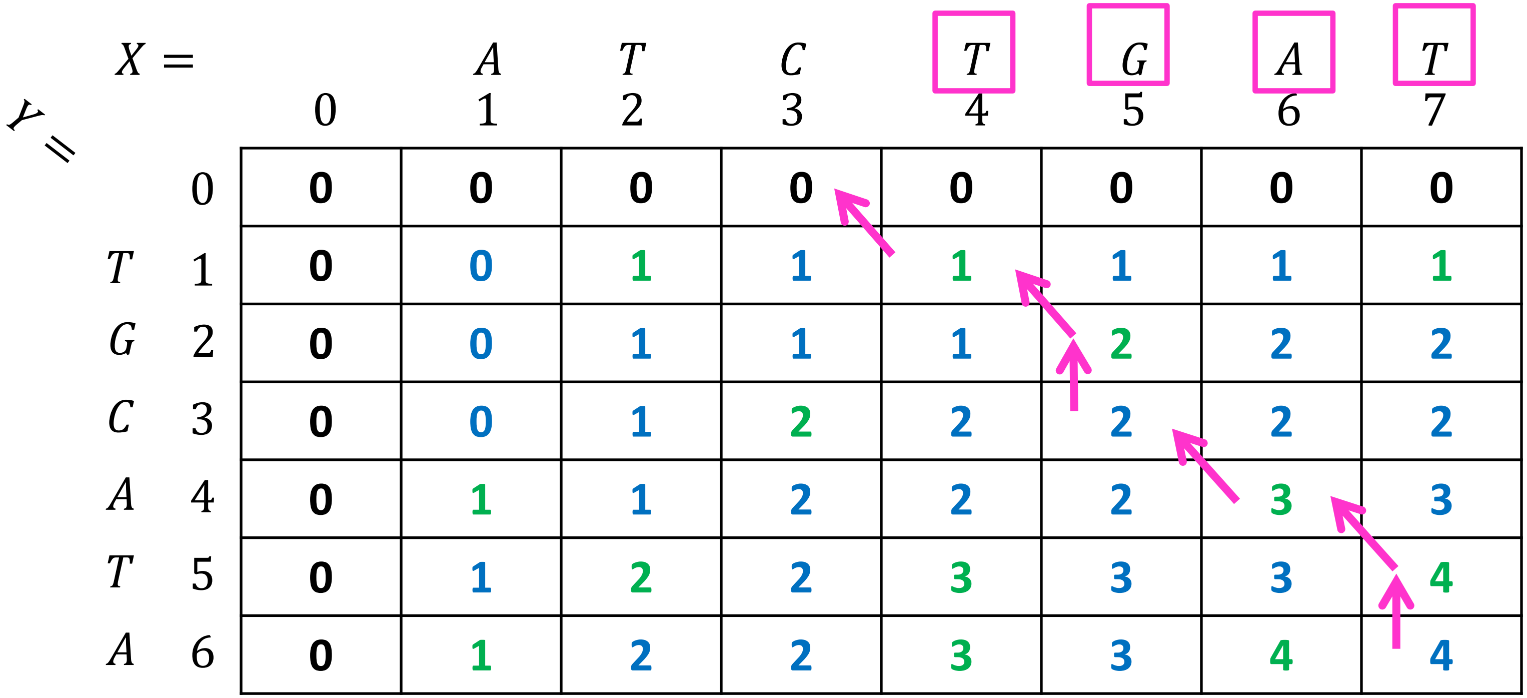
$$LCS(i,j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCS(i-1, j-1) + 1 & \text{if } X[i] = Y[j] \\ \max(LCS(i, j-1), LCS(i-1, j)) & \text{otherwise} \end{cases}$$

$Y =$	$X =$		A	T	C	T	G	A	T
		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	$\boxed{T}$	$\boxed{G}$	$\boxed{A}$	$\boxed{T}$
$Y =$		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**

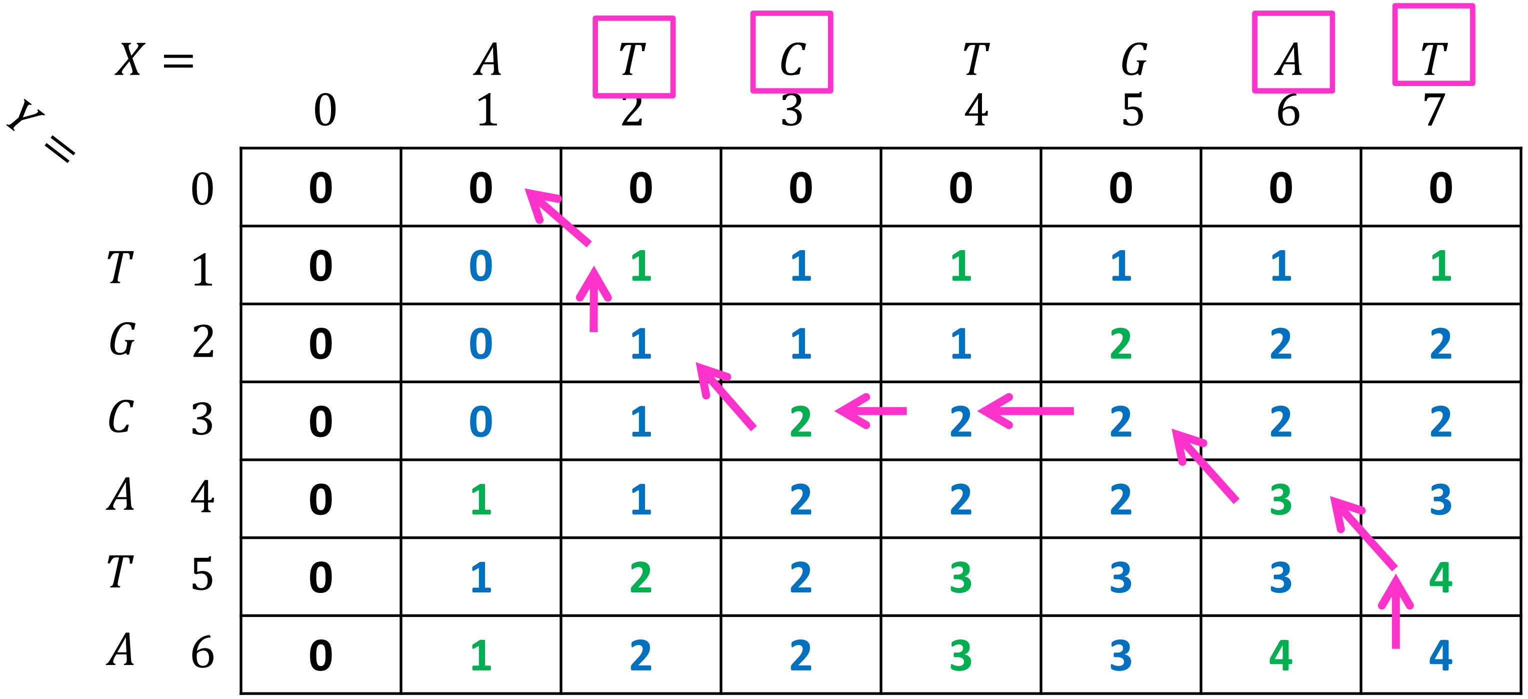
Start from bottom right,

if symbols matched, then go diagonally and print that symbol

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**
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**

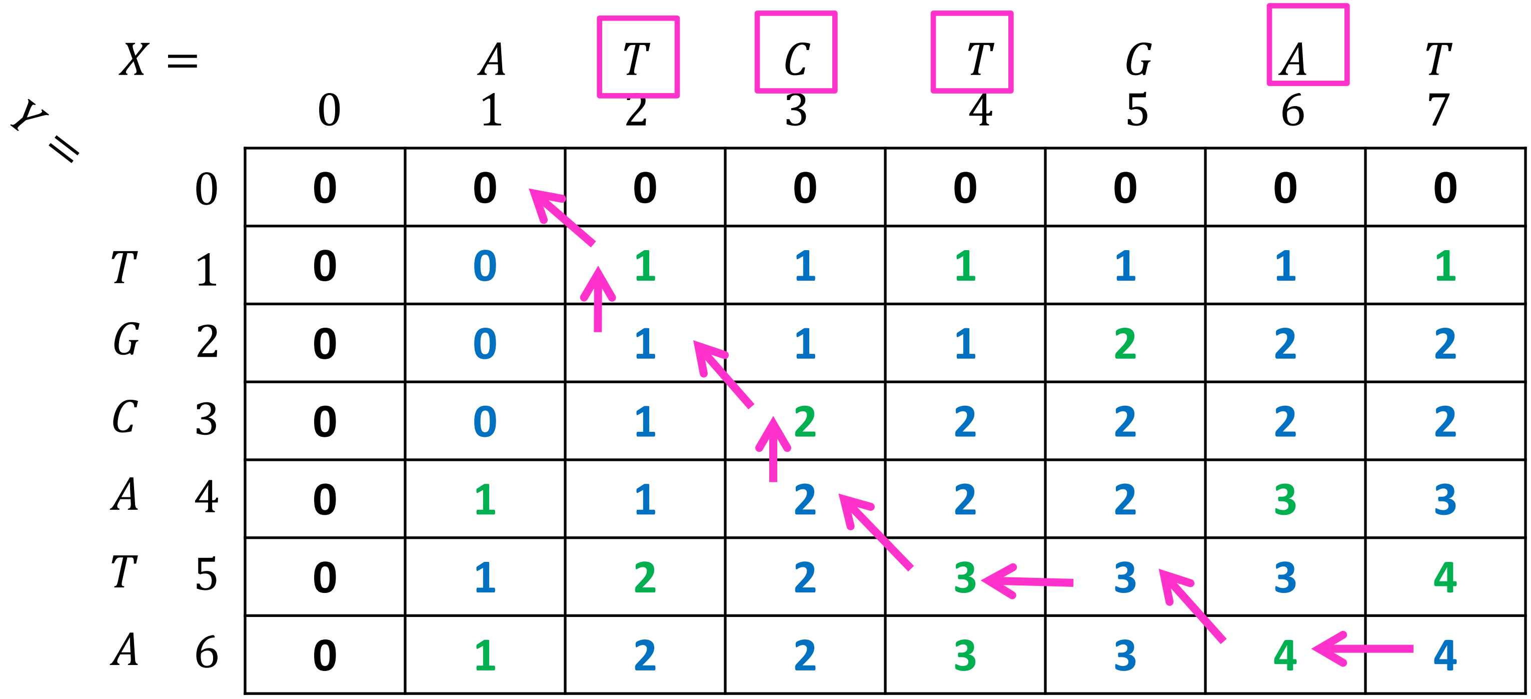
Start from bottom right,

if symbols matched, then go diagonally and print that symbol

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
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, then go diagonally and print that symbol

else go to largest adjacent