# 动态规划篇: 最长公共子串问题

# 童咏昕

北京航空航天大学 计算机学院

中国大学MOOC北航《算法设计与分析》



- 子序列
  - 将给定序列中零个或多个元素(如字符)去掉后所得结果

- 示例
  - 给定序列X





- 子序列
  - 将给定序列中零个或多个元素(如字符)去掉后所得结果

- 示例
  - 给定序列X

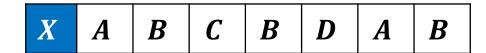


X的子序列

$X_{seq} \mid A \mid C \mid B \mid B$
---------------------------------------



- 子序列
  - 将给定序列中零个或多个元素(如字符)去掉后所得结果
- 子串
  - 给定序列中零个或多个连续的元素(如字符)组成的子序列
- 示例
  - 给定序列X

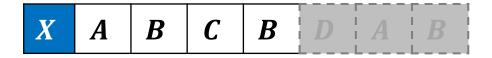


X的子序列

$X_{seq}$	4 <i>C</i>	B	В
-----------	------------	---	---



- 子序列
  - 将给定序列中零个或多个元素(如字符)去掉后所得结果
- 子串
  - 给定序列中零个或多个连续的元素(如字符)组成的子序列
- 示例
  - 给定序列X



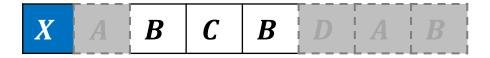
X的子序列

X的子串

$X_{seq}$	A	С	В	В	$X_{str}$	A	В	<i>C</i>	В
-----------	---	---	---	---	-----------	---	---	----------	---



- 子序列
  - 将给定序列中零个或多个元素(如字符)去掉后所得结果
- 子串
  - 给定序列中零个或多个连续的元素(如字符)组成的子序列
- 示例
  - 给定序列X



X的子序列

X的子串

$X_{seq}$	A	С	В	В	X	X <sub>str</sub>	В	С	В
-----------	---	---	---	---	---	------------------	---	---	---



• 给定两个序列X和Y





• 给定两个序列X和Y









• 给定两个序列X和Y

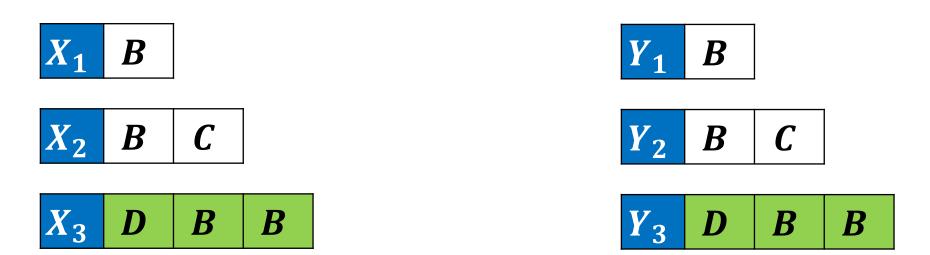






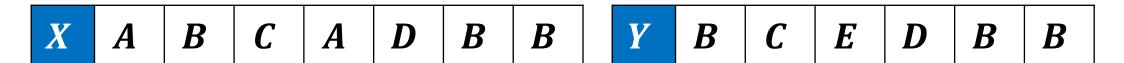
• 给定两个序列X和Y







• 给定两个序列X和Y



• 公共子串示例



问题: 如何求两个给定序列的最长公共子串?

### 问题定义



### • 形式化定义

### 最长公共子串问题

#### **Longest Common Substring Problem**

#### 输入

• 序列 $X = \langle x_1, x_2, ..., x_n \rangle$ 和序列 $Y = \langle y_1, y_2, ..., y_m \rangle$ 

### 问题定义



### • 形式化定义

### 最长公共子串问题

#### **Longest Common Substring Problem**

#### 输入

- 序列 $X = \langle x_1, x_2, ..., x_n \rangle$ 和序列 $Y = \langle y_1, y_2, ..., y_m \rangle$ 输出
- 求解一个公共子串 $Z = \langle z_1, z_2, ..., z_l \rangle$ , 令



### • 形式化定义

### 最长公共子串问题

#### **Longest Common Substring Problem**

#### 输入

- 序列 $X = \langle x_1, x_2, ..., x_n \rangle$ 和序列 $Y = \langle y_1, y_2, ..., y_m \rangle$ 输出
- 求解一个公共子串 $Z = \langle z_1, z_2, ..., z_l \rangle$ , 令

$$\max |Z|$$

$$s.t.Z = \langle x_i, x_{i+1}, ..., x_{i+l-1} \rangle = \langle y_j, y_{j+1}, ..., y_{j+l-1} \rangle$$
  
 $(1 \le i \le n-l+1; 1 \le j \le m-l+1)$ 



### • 形式化定义

#### 最长公共子串问题

#### **Longest Common Substring Problem**

#### 输入

- 序列 $X = \langle x_1, x_2, ..., x_n \rangle$ 和序列 $Y = \langle y_1, y_2, ..., y_m \rangle$ 输出
- 求解一个公共子串 $Z = \langle z_1, z_2, ..., z_l \rangle$ , 令

max |**Z**|

#### 优化目标

$$s.t.Z = \langle x_i, x_{i+1}, ..., x_{i+l-1} \rangle = \langle y_j, y_{j+1}, ..., y_{j+l-1} \rangle$$
  
 $(1 \le i \le n-l+1; 1 \le j \le m-l+1)$ 



### • 形式化定义

#### 最长公共子串问题

#### **Longest Common Substring Problem**

#### 输入

- 序列 $X = \langle x_1, x_2, ..., x_n \rangle$ 和序列 $Y = \langle y_1, y_2, ..., y_m \rangle$ 输出
- 求解一个公共子串 $Z = \langle z_1, z_2, ..., z_l \rangle$ ,令

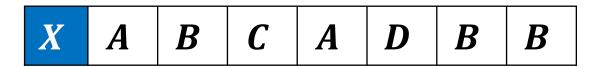
max |**Z**|

#### 优化目标

$$s.t.Z = \langle x_i, x_{i+1}, ..., x_{i+l-1} \rangle = \langle y_j, y_{j+1}, ..., y_{j+l-1} \rangle$$
  
 $(1 \le i \le n-l+1; 1 \le j \le m-l+1)$ 

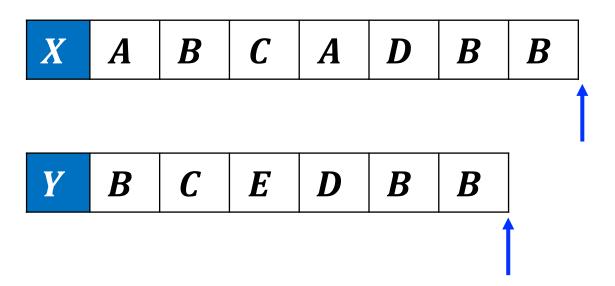
约束条件





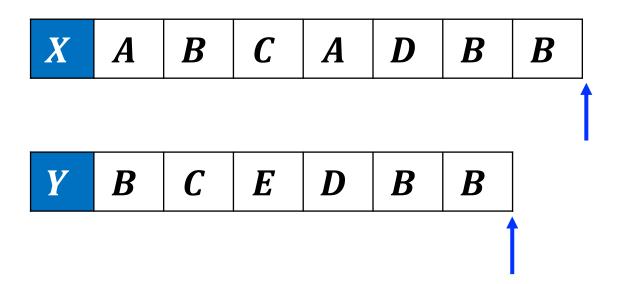
• 序列X和序列Y各选择一个位置X[i]和Y[j]





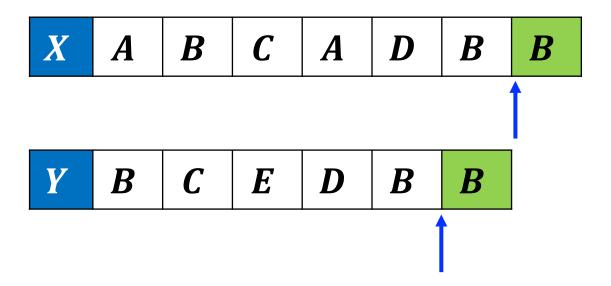
• 序列X和序列Y各选择一个位置X[7]和Y[6]





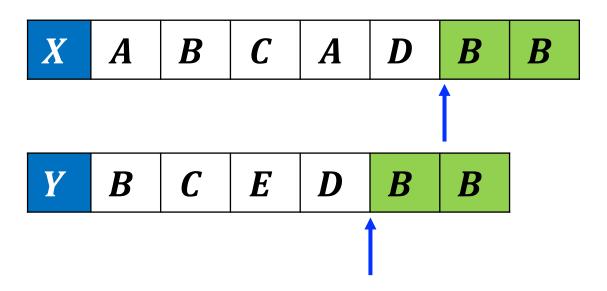
- 序列X和序列Y各选择一个位置X[7]和Y[6]
- 依次检查元素是否匹配





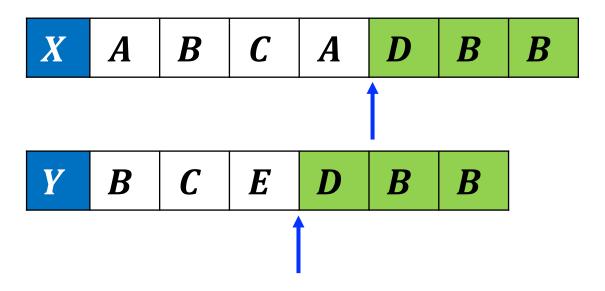
- 序列X和序列Y各选择一个位置X[7]和Y[6]
- 依次检查元素是否匹配
  - 元素相等继续匹配





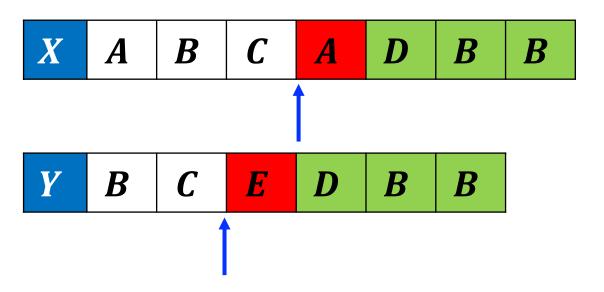
- 序列X和序列Y各选择一个位置X[7]和Y[6]
- 依次检查元素是否匹配
  - 元素相等继续匹配





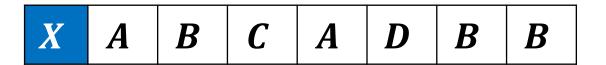
- 序列X和序列Y各选择一个位置X[7]和Y[6]
- 依次检查元素是否匹配
  - 元素相等继续匹配





- 序列X和序列Y各选择一个位置X[7]和Y[6]
- 依次检查元素是否匹配
  - 元素相等继续匹配
  - 元素不等(或某序列已达端点)匹配终止





- 枚举所有的X[i],Y[j]
- 求以其为结尾的尽可能长的公共子串







最长公共子串长度为3

- 枚举所有的X[i],Y[j]
- 求以其为结尾的尽可能长的公共子串
- 记录最长公共子串长度



X	A	B	<b>C</b>	A	D	B	B
	Y	B	C	E	D	B	B



X	A	B	<b>C</b>	A	D	B	B
	Y	В	<i>C</i>	E	D	B	B
	•	-	-	-	-	-	-
X	A	В	C	A	D	B	В

**----**



X	A	B	<b>C</b>	A	D	B	B
	Y	B	<b>C</b>	E	D	B	B
W		D	<u> </u>	A		D	D
X	A	B	C	A	D	B	B
	Y	B	<b>C</b>	E	D	B	B
				_			
X	A	$oldsymbol{B}$		A	D	B	B
	Y	B	C	E	D	B	B

<del>---</del>



X	A	B	<b>C</b>	A	D	B	B
	Y	B	C	E	D	B	B
X	A	B	<b>C</b>	A	D	B	B
	Y	В	С	E	D	B	B
X	A	B	<b>C</b>	A	D	B	B
	Y	B	<b>C</b>	E	D	B	B



X	A	B	C	A	D	B	B
	Y	B	С	E	D	B	B
		_					
X	A	B	<b>C</b>	A	D	B	B
	Y	B	С	E	D	B	B
X	A	B	<i>C</i>	A	D	B	B
	Y	В	С	E	D	B	B

• 可能存在最优子结构和重叠子问题

问题: 如何利用动态规划求解?

## 问题结构分析



- 给出问题表示
  - C[i,j]
    - X[1..i]和Y[1..j]中,以 $x_i$ 和 $y_j$ 结尾的最长公共子串Z[1..l]的长度

$X_i$	$x_1$	$x_2$	 $x_{i-1}$	$x_i$
$Y_j$	$y_1$	<b>y</b> <sub>2</sub>	 $y_{j-1}$	$y_j$



问题结构分析



递推关系建立



自底向上计算



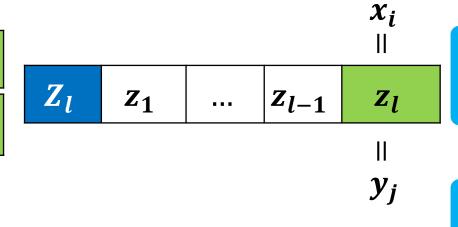
## 问题结构分析



### • 给出问题表示

- C[i,j]
  - X[1..i]和Y[1..j]中,以 $x_i$ 和 $y_j$ 结尾的最长公共子串Z[1..l]的长度

$X_i$	$x_1$	$x_2$	 $x_{i-1}$	$x_i$
$Y_j$	$y_1$	<b>y</b> <sub>2</sub>	 $y_{j-1}$	$y_j$



问题结构分析



递推关系建立



自底向上计算



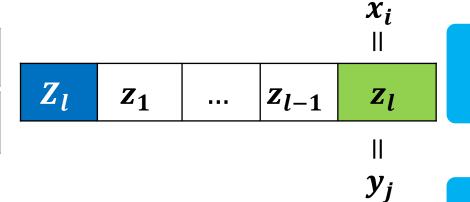
### 问题结构分析



### • 给出问题表示

- C[i,j]
  - X[1..i]和Y[1..j]中,以 $x_i$ 和 $y_j$ 结尾的最长公共子串Z[1..l]的长度

$X_i$	$x_1$	$x_2$	 $x_{i-1}$	$x_i$
$Y_j$	$y_1$	<b>y</b> <sub>2</sub>	 $y_{j-1}$	$y_j$



### • 明确原始问题

- $p_{max} = \max_{1 \leq i \leq n, 1 \leq j \leq m} \{C[i, j]\}$ 
  - X[1..n]和Y[1..m]中最长公共子串的长度

问题结构分析



递推关系建立



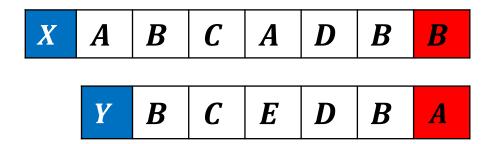
自底向上计算



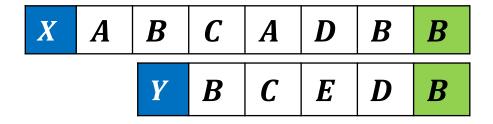
## 递推关系建立:分析最优(子)结构



• 情况1:  $x_7 \neq y_6$ 



• 情况2:  $x_7 = y_6$ 



问题结构分析



递推关系建立



自底向上计算

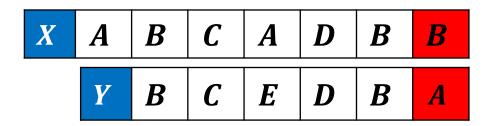


## 递推关系建立:分析最优(子)结构



• 情况1:  $x_7 \neq y_6$ 

C[7,6]



问题结构分析



递推关系建立



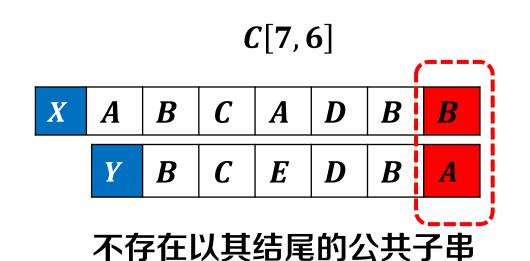
自底向上计算



## 递推关系建立:分析最优(子)结构



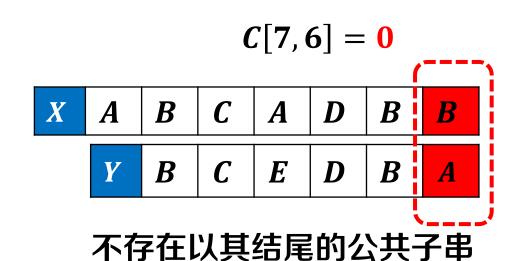
• 情况1:  $x_7 \neq y_6$ 







• 情况1:  $x_7 \neq y_6$ 



问题结构分析



递推关系建立



自底向上计算





• 情况1:  $x_i \neq y_j$ 

$$C[i,j]=0$$

X	$x_1$	$x_2$	 $x_{i-1}$	$x_i$
Y	<i>y</i> <sub>1</sub>	$y_2$	 $y_{j-1}$	$y_j$

问题结构分析



递推关系建立



自底向上计算





• 情况1:  $x_i \neq y_j$ 

$$C[i,j]=0$$

X	$x_1$	$x_2$	 $x_{i-1}$	$x_i$
Y	$y_1$	$y_2$	 $y_{j-1}$	$y_j$

无子问题

问题结构分析



递推关系建立



自底向上计算





情况2: x<sub>7</sub> = y<sub>6</sub>

C[7,6]

X	A	B	<i>C</i>	A	D	B	B
	Y	В	C	E	D	В	В

问题结构分析



递推关系建立



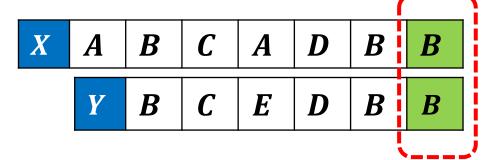
自底向上计算





情况2: x<sub>7</sub> = y<sub>6</sub>

C[7,6]



存在以其结尾的公共子串

问题结构分析



递推关系建立



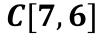
自底向上计算

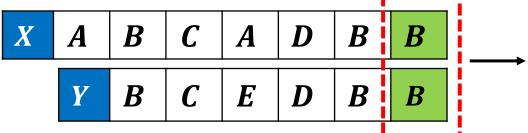




• 情况2:  $x_7 = y_6$ 

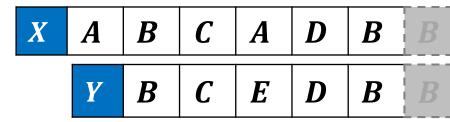






存在以其结尾的公共子串

$$C[7,6] = C[7-1,6-1] + 1$$



递推关系建立



自底向上计算





• 情况2:  $x_i = y_j$ 



X	$x_1$	$x_2$	 $x_{i-1}$	$x_i$
Y	<i>y</i> <sub>1</sub>	$y_2$	 $y_{j-1}$	$y_j$

问题结构分析



递推关系建立

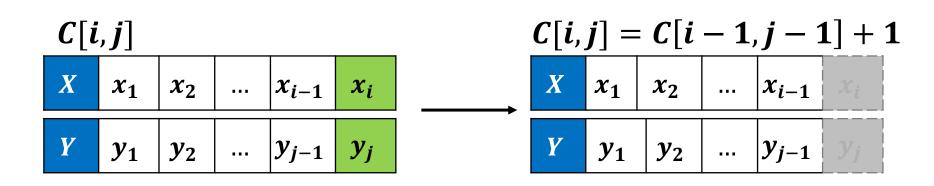


自底向上计算





• 情况2:  $x_i = y_j$ 



问题结构分析



递推关系建立

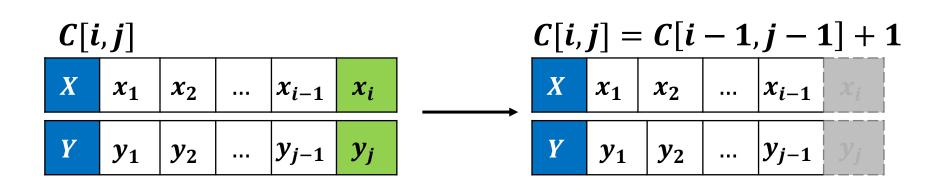


自底向上计算





• 情况2:  $x_i = y_i$ 



问题结构分析

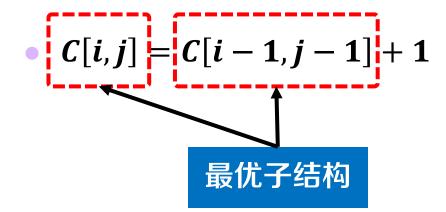


递推关系建立



自底向上计算

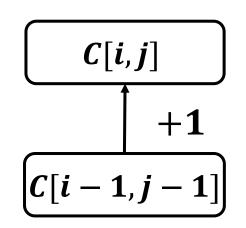




### 递推关系建立: 构造递推公式



• 
$$C[i,j] = \begin{cases} 0 & , x_i \neq y_j \\ C[i-1,j-1] + 1 & , x_i = y_j \end{cases}$$



#### 问题结构分析



递推关系建立



自底向上计算



### 自底向上计算:确定计算顺序



- 初始化
  - C[i, 0] = C[0, j] = 0
    - 。 某序列长度为0时,最长公共子串为0

问题结构分析



递推关系建立



自底向上计算



### 自底向上计算:确定计算顺序



- 初始化
  - C[i, 0] = C[0, j] = 0
    - 。 某序列长度为0时,最长公共子串为0

C[i,j]	j = 0	<i>j</i> = 1	j=2		j=m	初始化
i = 0	0	0	0	0	0	
i = 1	0					
i = 2	0					
	0					
i = n	0					

问题结构分析



递推关系建立



自底向上计算



### 自底向上计算:确定计算顺序



#### • 初始化

- C[i, 0] = C[0, j] = 0
  - 某序列长度为0时,最长公共子串为0
- 递推公式

• 
$$C[i,j] = \begin{cases} 0 & , x_i \neq y_j \\ C[i-1,j-1] + 1 & , x_i = y_j \end{cases}$$

C[i,j]	j = 0	j = 1	j=2		j=m
i = 0	0	0	0	0	0
i = 1	0				
i = 2	0				
***	0			C[i,j]	
i = n	0				

问题结构分析



递推关系建立



自底向上计算



### 自底向上计算: 依次求解问题



#### • 初始化

- C[i, 0] = C[0, j] = 0
  - 。 某序列长度为0时,最长公共子串为0
- 递推公式

• 
$$C[i,j] = \begin{cases} 0 & , x_i \neq y_j \\ C[i-1,j-1] + 1 & , x_i = y_j \end{cases}$$

C[i,j]	j = 0	j = 1	j=2		j = m	
i = 0	0	0	0	0	自	底向上计算
i = 1	0					
i = 2	0	<del></del>		=		
	0	<del>+</del>				
i = n	0	<del>+</del>			<b>→</b>	

问题结构分析



递推关系建立



自底向上计算



### 自底向上计算: 依次求解问题



#### • 初始化

- C[i, 0] = C[0, j] = 0
  - 。 某序列长度为0时,最长公共子串为0

#### • 原始问题

• 
$$p_{max} = \max_{1 \leq i \leq n, 1 \leq j \leq m} \{C[i, j]\}$$

C[i,j]	j = 0	<i>j</i> = 1	j=2		j=m
i = 0	0	0	0	0	0
i = 1	0				
i = 2	0		*_		
•••	0			最优解	
i = n	0				

问题结构分析



递推关系建立



自底向上计算



### 最优方案追踪



- 记录决策过程
  - 最长公共子串末尾位置为 $p_{max}$
  - 最长公共子串长度为*l<sub>max</sub>*

问题结构分析



递推关系建立



自底向上计算

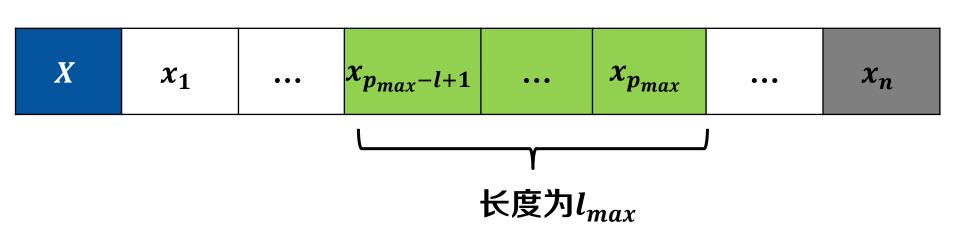


### 最优方案追踪



- 记录决策过程
  - 最长公共子串末尾位置为 $p_{max}$
  - 最长公共子串长度为*l<sub>max</sub>*

- 输出最优方案
  - 最长公共子串< $x_{p_{max}-l+1}, x_{p_{max}-l+2}, \dots, x_{p_{max}}$ >



问题结构分析



递推关系建立



自底向上计算





	1	2	3	4	5	6	7
$X_i$	A	B	C	A	D	В	В
$Y_{j}$	В	C	E	D	В	$\boldsymbol{B}$	



	1	2	3	4	5	6	7
$X_{i}$	A	В	<b>C</b>	A	D	B	B
$Y_{j}$	В	<b>C</b>	E	D	$\boldsymbol{B}$	B	

位置 $p_{max} = 0$ 长度 $l_{max} = 0$ 

j	0	1	2	3	4	5	6
0							
1							
2							
3							
4							
5							
6							
7							



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	<b>C</b>	E	D	В	В	

$\boldsymbol{\mathcal{C}}$		1
	_	_

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0		初出4				
2	0		初始化	Ĺ			
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	C	A	D	В	B
$Y_j$	В	λ	$x_i \neq y_j$	j	В	В	

<b>C</b> []	ر ر	$x_i \neq$	- <i>y<sub>j</sub></i>				
j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0						
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	В	λ	$c_i \neq y$	j	В	В	

<b>C</b> []	\	$x_i \neq$	- <i>y</i> j	B			
j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0					
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	$\boldsymbol{B}$	С	E	D	В	В	

	Γ	7
C	ı	ı

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0				
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



			3				
$X_i$	A	В	<i>C</i>	A	D	В	В
			E			В	

$\boldsymbol{\Gamma}$	Γ	٦
L		

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0			
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_{i}$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	C	E	D	В	В	

位置 $p_{max} = 0$ 长度 $l_{max} = 0$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0		
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



					5		
$X_{i}$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	B	C	E	D	В	В	

	Γ	٦
U	L	

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	В
$Y_j$	В	С	E	D	В	В	

位置 $p_{max} = 0$ 长度 $l_{max} = 0$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	C	$x_i =$	$= y_j$	}	В	

C[]			$x_i - y_j$				
j $i$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	$\begin{bmatrix} 0 \end{bmatrix}$	0	0	0	0	0	0
2	0						
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	В	C	$x_i =$	$= y_j$	3	В	

<b>C</b> []			$x_i - y_j$				
j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1		0	0	0	0	0	0
2	0	$\begin{bmatrix} 1 \end{bmatrix}$					
3	0	·					
4	0						
5	0						
6	0						
7	0						



				4			7
$X_{i}$	A	В	<b>C</b>	A	D	B	B
$Y_{j}$	· ·	С		D	В	В	

位置 $p_{max}=2$ 长度 $l_{max}=1$ 

$\boldsymbol{\Gamma}$	Γ	7
C		ı

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0				
3	0						
4	0						
5	0						
6	0						
7	0						



			3				
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	С	E	D	В	В	

位置 $p_{max}=2$ 长度 $l_{max}=1$ 

j $i$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0			
3	0						
4	0						
5	0						
6	0						
7	0						



					5		
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	С	E	D	В	В	

位置 $p_{max}=2$ 长度 $l_{max}=1$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0		
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	С	E	D	В	В	

位置 $p_{max}=2$ 长度 $l_{max}=1$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	C	E	D	В	В	

位置 $p_{max}=2$ 长度 $l_{max}=1$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0						
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_{i}$	A	В	С	A	D	В	B
$Y_{j}$	B	C	E	D	В	В	

位置 $p_{max}=2$ 长度 $l_{max}=1$ 

	Γ	7
C	ı	ı

j $i$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0					
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	С	A	D	В	B
$Y_{j}$	В	С	E	D	B	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	$\begin{bmatrix} 1 \end{bmatrix}$	0	0	0	1	1
3	0	0	2				
4	0						
5	0						
6	0						
7	0						



			3				
$X_i$	A	В	С	A	D	В	В
	В	<b>C</b>	E	D	B	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0			
4	0						
5	0						
6	0						
7	0						



			3				
$X_i$	A	В	C	A	D	В	B
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0		
4	0						
5	0						
6	0						
7	0						



					5		
$X_{i}$	A	В	С	A	D	В	В
	B	C	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	
4	0						
5	0						
6	0						
7	0						



			3				
$X_i$	A	В	С	A	D	В	В
			E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0						
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	B	<b>C</b>	A	D	В	B
$Y_j$	В	С	E	D	B	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0					
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_{i}$	A	B	<b>C</b>	A	D	В	B
$Y_{j}$	B	С	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0				
5	0						
6	0						
7	0						



					5		
$X_{i}$	A	В	C	A	D	В	B
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

	Γ	7
C	ı	ı

j $i$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0			
5	0						
6	0						
7	0						



					5		
$X_i$	A	B	<b>C</b>	A	D	В	В
		<b>C</b>		D		В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0		
5	0						
6	0						
7	0						



				4			
$X_{i}$	A	В	<b>C</b>	A	D	В	B
	В	C	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0						
6	0						
7	0						



	1	2	3	4	5	6	7
$X_{i}$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	С	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

$\boldsymbol{\Gamma}$	Γ	7
C		ı

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0					
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$						B	B
$Y_{j}$	В	С	E	D	$\boldsymbol{B}$	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

$\boldsymbol{\Gamma}$	Γ	٦
L		

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0				
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	C	A	D	B	В
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0			
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

	Γ	7
U	ı	1

j $i$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1		
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	B	<b>C</b>	A	D	B	В
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	
6	0						
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	B	B
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

	Γ	7
L	L	

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0						
7	0						



	1	2	3	4	5	6	7
	A				D	В	В
$Y_{j}$	$oldsymbol{B}$	С	E	D	B	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

$\boldsymbol{\Gamma}$	Г	7
L	ı	

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5		0	0	0	1	0	0
6	0	1					
7	0						



	1	2	3	4	5	6	7
$X_i$	A	B	C	A	D	В	В
$Y_{j}$	В	С	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

$\boldsymbol{\Gamma}$	Γ	٦
L		

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0				
7	0						



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	B	<b>C</b>	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j $i$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0			
7	0						



	1	2	3	4	5	6	7
$X_{i}$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0		
7	0						



					5		
$X_i$	A	В	<b>C</b>	A	D	В	B
$Y_{j}$	A B	<b>C</b>	E	D	В	В	

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0		0	0
6	0	1	0	0	0	2	
7	0						



				4			
$X_{i}$	A	В	C	A	D	В	B
	В						

位置 $p_{max} = 3$ 长度 $l_{max} = 2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0	2	1
7	0						



	1	2	3	4	5	6	7
$X_{i}$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	В	С	E	D	B	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0	2	1
7	0	1					



	1	2	3	4	5	6	7
$X_i$	A	B	C	A	D	В	В
$Y_{j}$	В	С	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0	2	1
7	0	1	0				



	1	2	3	4	5	6	7
$X_i$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

	Γ	7
L	L	

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0	2	1
7	0	1	0	0			



	1	2	3	4	5	6	7
$X_{i}$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0	2	1
7	0	1	0	0	0		



	1	2	3	4	5	6	7
$X_{i}$	A	В	<b>C</b>	A	D	В	В
$Y_{j}$	В	<b>C</b>	E	D	В	В	

位置 $p_{max}=3$ 长度 $l_{max}=2$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	$\begin{bmatrix} 0 \end{bmatrix}$	2	1
7	0	1	0	0	0	1	



	1	2	3	4	5	6	7
$X_{i}$	A	В	C	A	D	В	В
$Y_{j}$	В	C	E	D	В	В	

位置 $p_{max} = 7$ 长度 $l_{max} = 3$ 

 $\emph{\textbf{C}}[\ ]$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0	2	1
7	0	1	0	0	0	1	3



	1	2	3	4	5	6	7
$X_{i}$	A	В	C	A	D	B	В
$Y_{j}$	В	<b>C</b>	E	D	B	В	

位置 $p_{max} = 7$ 长度 $l_{max} = 3$ 

	Γ	7
L	L	

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	^	1		0
6	0	1	0	最长公	共子串	〈度	1
7	0	1	0	0	0	1	3



	1	2	3	4	5	6	7
$X_{i}$	A	B	C	A	D	В	В
$Y_{j}$	B	<b>C</b>	E	D	В	В	

位置 $p_{max} = 7$ 长度 $l_{max} = 3$ 

j	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	1	0	0	0	1	1
3	0	0	2	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0
6	0	1	0	0	0	2	1
7	0	1	0	0	0	1	3



```
输入: 两个字符串X, Y
 输出: X和Y的最长公共子串
//初始化
n \leftarrow \operatorname{length}(X)
m \leftarrow \operatorname{length}(Y)
新建二维数组C[0..n,0..m]
 l_{max} \leftarrow 0
p_{max} \leftarrow 0
 for i \leftarrow 0 to n do
  C[i,0] \leftarrow 0
 end
for j \leftarrow 0 to m do
  C[0,j] \leftarrow 0
 end
```

序列长度

### 伪代码



#### Longest-Common-Substring(X, Y)

```
输入: 两个字符串X, Y
 输出: X和Y的最长公共子串
//初始化
n \leftarrow \operatorname{length}(X)
m \leftarrow \operatorname{length}(Y)
新建二维数组C[0..n,0..m]
l_{max} \leftarrow 0
p_{max} \leftarrow 0
for i \leftarrow 0 to n do
C[i,0] \leftarrow 0
end
for j \leftarrow 0 to m do
 C[0,j] \leftarrow 0
lend
```

#### 初始化最优解



```
//动态规划
for i \leftarrow 1 to n do
                                                      依次计算子问题
    for j \leftarrow 1 to m do
      if X_i \neq Y_i then
          C[i,j] \leftarrow 0
        end
        else
            C[i,j] \leftarrow C[i-1,j-1] + 1
            if C[i,j] > l_{max} then
              l_{max} \leftarrow C[i,j]
               p_{max} \leftarrow i
            end
        end
    end
end
return l_{max}, p_{max}
```



```
//动态规划
for i \leftarrow 1 to n do
    for j \leftarrow 1 to m_{do}
      if X_i \neq Y_i then
                                                            末尾不等
      C[i,j] \leftarrow 0
      \mathbf{end}
        else
            C[i,j] \leftarrow C[i-1,j-1] + 1
            if C[i,j] > l_{max} then
             l_{max} \leftarrow C[i,j]
               p_{max} \leftarrow i
             end
        end
    end
end
return l_{max}, p_{max}
```



```
//动态规划
for i \leftarrow 1 to n do
     for j \leftarrow 1 to m do
           if X_i \neq Y_i then
              C[i,j] \leftarrow 0
           end
          \mathbf{else}
                                                                                 末尾相等
             \begin{vmatrix} C[i,j] \leftarrow C[i-1,j-1] + 1 \\ \textbf{if } C[i,j] > l_{max} \textbf{ then} \end{vmatrix} 
                    l_{max} \leftarrow C[i,j]
                     p_{max} \leftarrow i
                 end
           end
     end
end
return l_{max}, p_{max}
```



```
//动态规划
for i \leftarrow 1 to n do
    for j \leftarrow 1 to m do
        if X_i \neq Y_i then
          C[i,j] \leftarrow 0
        end
        else
            C[i,j] \leftarrow C[i-1,j-1] + 1
        if C[i,j] > l_{max} then
                                                    记录最长公共子串
            l_{max} \leftarrow C[i,j]
              p_{max} \leftarrow i
            end
        \mathbf{end}
    end
end
return l_{max}, p_{max}
```

### 伪代码



• Print-LCS(X,  $l_{max}$ ,  $p_{max}$ )

```
输入: 字符串 X, l_{max}, p_{max}
输出: X 和 Y的最长公共子串
if l_{max}=0 then
-1 return NULL
end
for i\leftarrow (p_{max}-l_{max}+1) to p_{max} do
-1 print X_i
end
```

### 伪代码



#### • Print-LCS(X, $l_{max}$ , $p_{max}$ )

```
输入: 字符串 X, l_{max}, p_{max}
输出: X 和 Y的最长公共子串
if l_{max} = 0 then
| return NULL
end
for i \leftarrow (p_{max} - l_{max} + 1) to p_{max} do
| print X_i
end
```

#### 追踪最优解

### 时间复杂度分析

return  $l_{max}, p_{max}$ 



#### Longest-Common-Substring(X, Y)

```
//动态规划
for i \leftarrow 1 to n do
    for j \leftarrow 1 to m do
        if X_i \neq Y_i then
         C[i,j] \leftarrow 0
        end
        else
            C[i,j] \leftarrow C[i-1,j-1] + 1
            if C[i,j] > l_{max} then
              l_{max} \leftarrow C[i,j]
               p_{max} \leftarrow i
             end
        end
    end
end
```

时间复杂度:  $O(n \cdot m)$ 



X	D	B	D	A	B
Y	С	A	B	B	

X	D	B	D	A	B
Y	С	A	B	B	



X	D	В	D	A	B
Y	C	A	B	B	

### 最长公共子串

X	D	B	D	A	$oxedsymbol{B}$
Y	C	A	B	B	

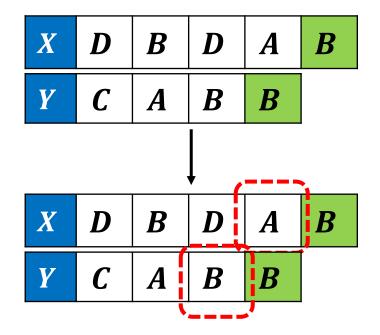
情况2:  $x_5 = y_4$ 



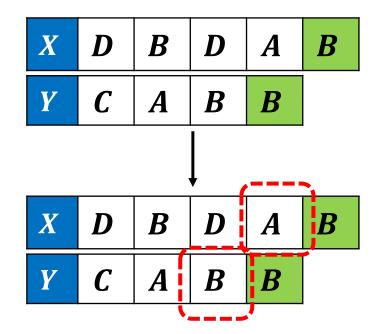
X	D	B	D	A	B
Y	C	A	B	B	

X	D	B	D	A	B
Y	<b>C</b>	A	B	B	





### 最长公共子串



情况1:  $x_4 \neq y_3$ 



