



## 章节内容

1. 最长公共子序列问题描述
2. 最长公共子序列问题分析
3. 最长公共子序列求解实例
4. 最长公共子序列算法描述

01

## 最长公共子序列问题描述

**子序列：**给定序列  $X = \{x_1, x_2, \dots, x_m\}$  和  $Z = \{z_1, z_2, \dots, z_k\}$ ，若序列X中存在一个严格递增的下标序列  $\{i_1, i_2, \dots, i_k\}$ ，使得对所有  $j (j = 1, 2, \dots, k)$ ，均有  $z_j = x_{i_j}$ ，则称序列Z是序列X的子序列。

例1：给定序列  $X = \{A, B, C, B, D, A, B\}$  和  $Z = \{B, C, D, B\}$ 。

解：

序列X的下标序列  $\{2, 3, 5, 7\}$ ，即  $\{B, C, D, B\}$  与序列Z相同  
满足子序列的定义，故序列Z是序列X的子序列。



## 最长公共子序列问题描述

**公共子序列：**给定两个序列X和Y，当另一序列Z既是X的子序列又是Y的子序列时，称Z是序列X和Y的公共子序列。

例2. 序列 $X = \{A, B, C, B, D, A, B\}$ ， $Y = \{B, D, C, A, B, A\}$ ， $Z_1 = \{B, C, A\}$ ， $Z_2 = \{B, C, B, A\}$ 。

$Z_2$ 是X和Y的一个最长公共子序列



思考：如何求两个序列X和Y的最长公共子序列。

## 最长公共子序列问题分析

### 1) 最优子结构

设序列 $X = \{x_1, x_2, \dots, x_m\}$ 和 $Y = \{y_1, y_2, \dots, y_n\}$ 的最长公共子序列为 $Z = \{z_1, z_2, \dots, z_k\}$

(1) 若 $x_m = y_n$ , 则 $Z_k = X_m = Y_n$ , 且 $Z_{k-1}$ 是 $X_{m-1}$ 和 $Y_{n-1}$ 的最长公共子序列。

**证明:** 若 $X_{m-1}$ 和 $Y_{n-1}$ 有长度大于 $k-1$ 的公共子序列 $W$ , 则 $\{W, x_m\}$ 是 $X$ 和 $Y$ 的长度大于 $k$ 的公共子序列。矛盾!



## 最长公共子序列问题分析

### 1) 最优子结构

(2) 若  $x_m \neq y_n$   $z_k \neq x_m$ , 则  $Z$  是  $X_{m-1}$  和  $Y$  的最长公共子序列



**证明:** 若  $X_{m-1}$  和  $Y$  有长度大于  $k$  的公共子序列  $W$ , 则  $W$  也是  $X$  和  $Y$  的长度大于  $k$  的公共子序列。 **矛盾!**



## 最长公共子序列问题分析

### 1) 最优子结构

(3) 若  $x_m \neq y_n$   $z_k \neq y_n$ , 则  $Z$  是  $X_m$  和  $Y_{n-1}$  的最长公共子序列

**证明:** 若  $X$  和  $Y_{n-1}$  有长度大于  $k$  的公共子序列  $W$ , 则  $W$  也是  $X$  和  $Y$  的长度大于  $k$  的公共子序列。 **矛盾!**



最长公共子序列问题具有最优子结构性质!





## 最长公共子序列问题分析

### 2) 递归结构

- (1) 若  $x_m = y_n$ , 找出 $X_{m-1}$ 和 $Y_{n-1}$ 的最长公共子序列, 在其尾部加上 $x_m$  (或 $y_n$ ) 即是 $X$ 和 $Y$ 的最长公共子序列;
- (2) (2) 若  $x_m \neq y_n$ , 首先找出 $X_{m-1}$ 和 $Y$ 的最长公共子序列, 其次找出 $X$ 和 $Y_{n-1}$ 的最长公共子序列, 二者较长者即为 $X$ 和 $Y$ 的最长公共子序列。



## 最长公共子序列问题分析

### 2) 递归结构

(3)  $c[i][j]$ 记录序列 $X_i$ 和 $Y_j$ 的最长公共子序列的长度:

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










## 最长公共子序列问题分析

### 2) 递归结构

(4) 在计算 $c[i][j]$ 的同时, 用 $b[i][j]$ 记录该值由哪种方式产生, 也分三种情况。

产生方式	$c[i][j] = c[i-1][j-1] + 1$	$c[i][j] = c[i-1][j]$	$c[i][j] = c[i][j-1]$
记录值	$b[i][j] = 1$	$b[i][j] = 2$	$b[i][j] = 3$
			

- 
- 给定两个序列为 $X = \{A, B, C, B, D, A, B\}$  和  $Y = \{B, D, C, A, B, A\}$ ，求最长公共子序列？



(1) 数组c初始化

即 $c[0][j] = 0$     $c[i][0] = 0$



## 最长公共子序列求解实例

1)  $i = 1 \quad X_1 = \{A\}$

$$j = 1 \quad Y_1 = \{B\} \quad c[1][1] = \max\{c[0][1], c[1][0]\} = 0 \quad b[1][1] = 2$$

$$j = 2 \quad Y_2 = \{B, D\} \quad c[1][2] = \max\{c[0][2], c[1][1]\} = 0 \quad b[1][2] = 2$$

$$j = 3 \quad Y_3 = \{B, D, C\} \quad c[1][3] = \max\{c[0][3], c[1][2]\} = 0 \quad b[1][3] = 2$$

$$j = 4 \quad Y_4 = \{B, D, C, A\} \quad c[1][4] = c[0][3] + 1 = 1 \quad b[1][4] = 1$$

$$j = 5 \quad Y_5 = \{B, D, C, A, B\} \quad c[1][5] = \max\{c[0][5], c[1][4]\} = 1 \quad b[1][5] = 3$$

$$j = 6 \quad Y_6 = \{B, D, C, A, B, A\} \quad c[1][6] = c[0][5] + 1 = 1 \quad b[1][6] = 1$$



## 最长公共子序列求解实例

2)  $i = 2 \quad X_2 = \{A, B\}$

$$j = 1 \quad Y_1 = \{B\} \quad c[2][1] = c[1][0] + 1 = 1 \quad b[2][1] = 1$$

$$j = 2 \quad Y_2 = \{B, D\} \quad c[2][2] = \max\{c[1], c[2][1]\} = 1 \quad b[2][2] = 3$$

$$j = 3 \quad Y_3 = \{B, D, C\} \quad c[2][3] = \max\{c[1][3], c[2][2]\} = 1 \quad b[2][3] = 3$$

$$j = 4 \quad Y_4 = \{B, D, C, A\} \quad c[2][4] = \max\{c[1][4], c[2][3]\} = 1 \quad b[2][4] = 2$$

$$j = 5 \quad Y_5 = \{B, D, C, A, B\} \quad c[2][5] = c[1][4] + 1 = 2 \quad b[2][5] = 1$$

$$j = 6 \quad Y_6 = \{B, D, C, A, B, A\} \quad c[2][6] = \max\{c[1][6], c[2][5]\} = 2 \quad b[2][6] = 3$$



03

## 最长公共子序列求解实例

3)  $i = 3 \quad X_3 = \{A, B, C\}$

$j = 1 \quad Y_1 = \{B\} \quad c[3][1] = 1 \quad b[3][1] = 2$

$j = 2 \quad Y_2 = \{B, D\} \quad c[3][2] = 1 \quad b[3][2] = 2$

$j = 3 \quad Y_3 = \{B, D, C\} \quad c[3][3] = 2 \quad b[3][3] = 1$

$j = 4 \quad Y_4 = \{B, D, C, A\} \quad c[3][4] = 2 \quad b[3][4] = 3$

$j = 5 \quad Y_5 = \{B, D, C, A, B\} \quad c[3][5] = 2 \quad b[3][5] = 2$

$j = 6 \quad Y_6 = \{B, D, C, A, B, A\} \quad c[3][6] = 2 \quad b[2][6] = 2$



c表

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



指引构造解的方向

b表

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

箭头尾部位置对应的元素为公共子序列元素

c表

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

b表

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

$X = \langle A, B, C, B, D, A, B \rangle$

$Y = \langle B, D, C, A, B, A \rangle$

B

C

B

A

## 最长公共子序列算法描述

```

1  LCS_LENGTH(int m,int n,char *x, char **y,int **c,int **b)
2  {
3      int i,j;
4      for (i=1; i<=m; i++) //m是X序列的长度, n是Y序列的长度
5          c[i][0]=0;
6      for (j=1; j<=n; j++)
7          c[0][j]=0;
8      for (i=1; i<=m; i++)
9          for (j=1; j<=n; j++)
10             if (x[i]==y[j])
11             {
12                 c[i][j]=c[i-1][j-1]+1;
13                 b[i][j]=1;
14             }
15             else if (c[i-1][j]>=c[i][j-1])
16             {
17                 c[i][j]=c[i-1][j];
18                 b[i][j]=2;
19             }
20             else
21             {
22                 c[i][j]=c[i][j-1];
23                 b[i][j]=3;
24             }
25  } //计算复杂性: O(mn)

```

```

1  LCS(int i,int j,char *x,int **b)
2  {
3      if (i==0||j==0)
4          return;
5      if (b[i][j]==1 )
6      {
7          LCS(i-1,j-1,x,b);
8          cout<<x[i];
9      }
10     else if (b[i][j]==2)
11         LCS(i-1,j,x,b);
12     else
13         LCS(i,j-1,x,b);
14 } //计算复杂性: O(m+n)

```



## 总结

在本节中首先对最长公共子序列问题进行了介绍，其次对最长公共子序列问题进行了分析，最后通过一个实例展示了使用动态规划方法求解最长公共子序列的过程。



## 课后思考题

- (1) 两个序列的最长公共子序列唯一吗？
- (2) 如果 $c[i-1][j]=c[i][j-1]$ 相等,  $b[i][j]=2?$  or  $=3?$  不同的选择对结果有影响吗？
- (3) 序列 $X = \{a, b, c, b, d, b\}$ 和 $Y = \{a, c, b, b, a, b, d, b, b\}$ , 求最长公共子序列。