

# Task3-Q2

2019年12月1日 17:31

## • Q2

用动态规划算法求最长公共子序列问题。

## • 最长公共子序列 (LCS)

最长公共子序列 (LCS) 是一个在一个序列集合中 (通常为两个序列) 用来查找所有序列中最长子序列的问题。这与查找最长公共子串的问题不同的地方是: 子序列不需要在原序列中占用连续的位置。而最长公共子串 (要求连续) 和最长公共子序列是不同的。

另外在计算机科学中, **最长递增子序列**是指, 在一个给定的数值序列中, 找到一个子序列, 使得这个子序列元素的数值依次递增, 并且这个子序列的长度尽可能地大。最长递增子序列中的元素在原序列中不一定是连续的。许多与数学、算法、随机矩阵理论 (random matrix theory)、表示论相关的研究都会涉及最长递增子序列。解决最长递增子序列问题的**算法最低要求 $O(n \log n)$ 的时间复杂度**, 这里 $n$ 表示输入序列的规模。

## • 算法设计思路

动态规划的一个计算两个序列的最长公共子序列的方法如下:

以两个序列  $X$ 、 $Y$  为例子:

设有二维数组  $f[i,j]$  表示  $X$  的  $i$  位和  $Y$  的  $j$  位之前的最长公共子序列的长度, 则有:

$f[1][1] = \text{same}(1,1);$

$f[i,j] = \max\{f[i-1][j-1] + \text{same}(i,j), f[i-1][j], f[i,j-1]\};$

其中,  $\text{same}(a,b)$  当  $X$  的第  $a$  位与  $Y$  的第  $b$  位相同时为 "1", 否则为 "0"。

此时, 二维数组中最大的数便是  $X$  和  $Y$  的最长公共子序列的长度, 依据该数组回溯, 便可找出最长公共子序列。

该算法的空间、时间复杂度均为  $O(n^2)$ , 经过优化后, 空间复杂度可为  $O(n)$ 。

		$j$						
		0	1	2	3	4	5	6
$i$	$y_j$		B	D	C	A	B	A
0	$x_i$	0	0	0	0	0	0	0
1	A	0	↑	↑	↑	↖1	←1	↖1
2	B	0	↖1	←1	←1	↑1	↖2	←2
3	C	0	↑1	↑1	↖2	←2	↑2	↑2
4	B	0	↑1	↑1	↑2	↑2	↖3	←3
5	D	0	↑1	↖2	↑2	↑2	↑3	↑3
6	A	0	↑1	↑2	↑2	↖3	↑3	↖4
7	B	0	↖1	↑2	↑2	↑3	↖4	↑4

## • 源代码

```
# -*- coding: utf-8 -*-
# @Time : 2019/12/1 17:32
```

```

# @Author : BaoBao
# @Mail : baobaotql@163.com
# @File : LCS.py
# @Software: PyCharm

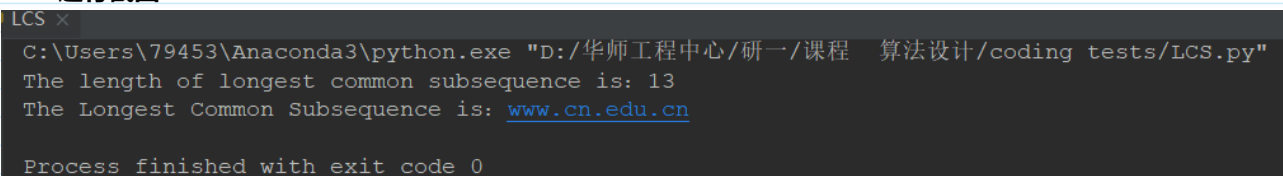
import numpy as np

def Longest_Common_Subsequence(a, b):
    '''
    :param a:字符串a
    :param b:字符串b
    :return:length
            subseq
    '''
    length = 0
    subseq = ''
    cell = np.zeros(shape=(len(a), len(b)))
    for i in range(len(a)):
        for j in range(len(b)):
            if a[i] == b[j]:
                cell[i][j] = cell[i - 1][j - 1] + 1
                if cell[i][j] > length:
                    length = cell[i][j]
                    subseq += a[i]
            else:
                cell[i][j] = max(cell[i-1][j], cell[i][j-1])
    return length, subseq

if __name__ == "__main__":
    a, b = Longest_Common_Subsequence('www.ccnu.edu.cn', 'www.neu.edu.cn')
    print('The length of longest common subsequence is: %d\nThe Longest Common Subsequence is: %s' % (a, b))

```

#### • 运行截图



```

LCS x
C:\Users\79453\Anaconda3\python.exe "D:/华师工程中心/研一/课程 算法设计/coding tests/LCS.py"
The length of longest common subsequence is: 13
The Longest Common Subsequence is: www.cn.edu.cn

Process finished with exit code 0

```