

最长公共子串问题的实现计算机等级考试 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/645/2021_2022__E6_9C_80_E9_95_BF_E5_85_AC_E5_c97_645870.htm 最长公共子串问题：

一个给定序列的子序列是在该序列中删去若干元素后得到的序列。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。最长公共子串就是求给定两个序列的一个最长公共子序列。例如，X=“ ABCBDAB ”，Y=“ BCDB ”是X的一个子序列。问题分析：给定两个序列A和B，称序列Z是A和B的公共子序列，是指Z同是A和B的子序列。问题要求已知两序列A和B的最长公共子序列。如采用列举A的所有子序列，并一一检查其是否又是B的子序列，并随时记录所发现的子序列，最终求出最长公共子序列。这种方法因耗时太多而不可取。考虑最长公共子序列问题如何分解成子问题，设A=“ a_0, a_1, \dots, a_{m-1} ”，B=“ b_0, b_1, \dots, b_{n-1} ”，并Z=“ z_0, z_1, \dots, z_{k-1} ”为它们的最长公共子序列。不难证明有以下性质：（1）如果 $a_{m-1}=b_{n-1}$ ，则 $z_{k-1}=a_{m-1}=b_{n-1}$ ，且“ z_0, z_1, \dots, z_{k-2} ”是“ a_0, a_1, \dots, a_{m-2} ”和“ b_0, b_1, \dots, b_{n-2} ”的一个最长公共子序列；（2）如果 $a_{m-1} \neq b_{n-1}$ ，则若 $z_{k-1} \neq a_{m-1}$ ，蕴涵“ z_0, z_1, \dots, z_{k-1} ”是“ a_0, a_1, \dots, a_{m-2} ”和“ b_0, b_1, \dots, b_{n-1} ”的一个最长公共子序列；（3）如果 $a_{m-1} \neq b_{n-1}$ ，则若 $z_{k-1} \neq b_{n-1}$ ，蕴涵“ z_0, z_1, \dots, z_{k-1} ”是“ a_0, a_1, \dots, a_{m-1} ”和“ b_0, b_1, \dots, b_{n-2} ”的一个最长公共子序列。这样，在找A和B的公共子序列时，如有 $a_{m-1}=b_{n-1}$ ，则进一步解决一个子问题，找“ a_0, a_1, \dots, a_{m-2} ”和“ $b_0, b_1,$

... , b_{m-2} ” 的一个最长公共子序列；如果 $a_{m-1} \neq b_{n-1}$, 则要解决两个子问题, 找出 “ a_0, a_1, \dots, a_{m-2} ” 和 “ b_0, b_1, \dots, b_{n-1} ” 的一个最长公共子序列和找出 “ a_0, a_1, \dots, a_{m-1} ” 和 “ b_0, b_1, \dots, b_{n-2} ” 的一个最长公共子序列, 再取两者中较长者作为A和B的最长公共子序列。为了节约重复求相同子问题的时间, 引入一个数组, 不管它们是否对最终解有用, 把所有子问题的解存于该数组中, 这就是动态规划法所采用的基本方法, 具体说明如下。定义 $c[i][j]$ 为序列 “ a_0, a_1, \dots, a_{i-2} ” 和 “ b_0, b_1, \dots, b_{j-1} ” 的最长公共子序列的长度, 计算 $c[i][j]$ 可递归地表述如下: (1) $c[i][j] = 0$ 如果 $i=0$ 或 $j=0$; (2) $c[i][j] = c[i-1][j-1] + 1$ 如果 $i, j > 0$ 且 $a_{i-1} = b_{j-1}$; (3) $c[i][j] = \max\{c[i-1][j], c[i][j-1]\}$ 如果 $i, j > 0$ 且 $a_{i-1} \neq b_{j-1}$ 。100Test 下载频道开通, 各类考试题目直接下载。详细请访问 www.100test.com