

北京大学 2014 年研究生算法课第 5 次作业

发布时间：2014 年 12 月 1 日

截止时间：2014 年 12 月 15 日课前

注意事项：

- 作业应独立完成，严禁抄袭。作业必须使用统一规定的模板。
- 在截止日期那天，直接把纸质版的作业交给任课老师。
- 如果因生病等特殊原因不能按时完成作业的，那么应在截止日期前一天向任课老师请假。
- 作业和考试中的动态规划算法设计都要说明状态转移方程、转移条件、初始状态、返回值、时间复杂度以及空间复杂度。

动态规划

1. 题目来源：《算法导论》第三版 15-9

题目描述：

某种字符串处理语言允许程序员将一个字符串拆分成两段，由于此操作需要复制字符串，因此需要花费 n 个时间单位来将一个 n 个字符的字符串拆为两段。假定一个程序员希望将一个字符串拆分为多段，拆分的顺序会影响所花费的总时间。例如，假定这个程序员希望将一个 20 个字符的字符串在第 2 个、第 8 个以及第 10 个字符后进行拆分（字符从左到右，从 1 开始升序编号）。如果按由左向右的顺序进行拆分，则第一次拆分花费 20 个时间单位，第二次花费 18 个时间单位，第三次花费 12 个时间单位，共花费 50 个时间单位。但如果从右向左，只需要花费 38 个时间单位。当然我们可以选择每次都从不同的方向来拆分。

请你设计算法，对给定的拆分位置，确定最小代价的拆分顺序。更形式化的：给定一个 n 个字符的字符串 S 和一个保存 m 个拆分点的数组 $L[1..m]$ ，计算拆分的最小代价，以及最优拆分序列。

2. 题目来源：《算法导论》第三版 15-7

题目描述：

我们可以通过在有向图 $G = (V, E)$ 上使用动态规划方法来实现语音识别，对每条边 $(u, v) \in E$ 打上一个声音标签 $\sigma(u, v)$ ，该声音来自于有限声音集 Σ ，这样的标签图就成为一个特定人说限定语言的形式化模型。图中从特定顶点 $v_0 \in V$ 开始的每条路径都对应模型产生的一个可能的声音序列。对于一条有向路径，我们定义其标签为路径中边的标签的简单连结。

(a) 设计高效算法，对给定的带边标签的图 G 、特定点 v_0 以及 Σ 上的声音序列 $s = (\sigma_1, \sigma_2, \dots, \sigma_k)$ ，返回 G 中从 v_0 开始的一条路径， s 为该路径的标签（如果存在这样的路径）。否则，算法应返回 NO-SUCH-PATH。分析算法的时间复杂度。

现在，假定每条边 $(u, v) \in E$ 都关联一个非负概率 $p(u, v)$ ，它表示从顶点 u 开始，经过边 (u, v) ，产生对应的声音的概率。任何顶点的出射边的概率之和为 1。一条路径的概率定义为路径的边的概率之积。对于从 v_0 开始的一条路径，我们可以将其概率看做为从 v_0 开始随机游走（指当位于定点 u 时，随机选择一条出射边前进，每条边被选中的概率就是它所关联的概率），最后恰巧经过这条路径的概率。

(b) 扩展 (a) 的算法，使得返回的路径是从 v_0 开始且标签为 s 的路径中概率最大者，分析算法的时间复杂性。