北京大学2021年研究生算法课第1次书面作业

发布时间：2021年9月27日
截止时间：2021年10月18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。

引言、算法分析基础、图

1. **题目来源：**《算法设计》第一章第6题：

题目描述：PSL公司是一家航运公司，它拥有n条船，为n个港口提供服务。每条船有一个时间表，对一个月的某一天，这个时间表会指明这条船当时会在哪个港口，或者这条船当时会在出海。假定这里的月有m天并且m>n。在这个月里每条船访问每个港口恰好1天。为了安全起见，PSL公司有这样的严格要求：

（a）同一天不能有两条船访问同一个港口。

公司在这个月想对所有的船进行维护，因此他们在考虑“截断”每条船的时间表。具体来讲，对每条船S，当它在某一天到达那一天它该到达的港口P之后，它就可以一直停留在港口P一直到月底，时间表上本来预定的后续港口都不会访问了。

证明无论每条船最初的时间表是什么样，总可以找到这样一组截断方案，使得同一天不会有两条船停在同一个港口。给出一个制定截断方案的算法。

例：假设有两条船和两个港口，这个月有4天。假设第一条船的时间表是：

港口P1，出海，港口P2，出海

第二条船的时间表是：

出海，港口P1，出海，港口P2

那么一个截断方案是让第一条船第三天起停留在P2，让第二条船第二天起停留在P1.这个问题中这是唯一满足条件的截断方案

2. **题目描述：**

设 n 是 k 的倍数，有 k 个排好序的表 $L_1, L_2, \ldots, L_k$，每个表都有 n/k 个数，假设全部 n 个数各不相同，归并长度为 m 和 n 的两个表的代价是 O(m+n)。请回答下面两个问题：

（1） 如果使用顺序归并算法按照下标从小到大的顺序依次归并这 k 个表，那么最坏情况下的时间复杂度是多少？

（2） 设计一个高效的归并顺序，说明该算法的主要思想（无需给出伪代码）并分析最坏情况下的时间复杂度。

注意：只要给出渐进意义下的复杂度即可，不用给出精确值。

3. **题目来源：**　《算法设计》第二章第 8 题

题目描述：

你正在对不同的玻璃瓶样品做强度测试以确定它们从多高的高度掉下来而仍旧不碎。对某类特定的瓶子如下设计这样一个实验，你有一个具有 n 个横挡的阶梯，你想找出最高的横挡，能使一个瓶子的样品从那里下落而不被摔碎，我们把它称为“最高的安全横挡”。

尝试二分搜索可能是一种自然的想法:使一个瓶子从中间的横挡下落，看看它是否摔碎，然后依赖于这个结果递归地从 n/4 横挡或者 3n/4 的横挡进行尝试。但是这个办法有个缺点：在找到答案之前，你可能摔碎了一大堆瓶子。

如果你的主要目标是保护瓶子，另一方面，可以尝试下面的策略。从第一个横挡开始让瓶子下落，然后第二个横档，继续下去，每次向上爬一个高度直到瓶子摔碎为止。以这种方式，你只需要一个瓶子——在它摔碎的时刻，你得到正确的答案——但是你可能不得不把它下落 n 次(而不是在二分搜索求解时的 log n 次)

于是，这里是某种权衡:似乎是如果你愿意打碎更多的瓶子，你就可以执行更少的下落。为了更好地理解这种权衡在量级上是怎样起作用的，让我们考虑给定 k⩾1 个瓶子的“预算”，怎样来运行这个实验。换句话说，你必须确定正确的答案——最高的安全横挡——并且在这个过程中至多可以用 k 个瓶子。

(a)假设只给你 k=2 个瓶子。描述一种找到最高安全横挡的策略，对某个比线性函数增长更慢的函数$f(n)$，要求一个瓶子至多落$f(n)$次(换句话说，应该满足$\lim_{n\to\infty} f(n)/n = 0$)

(b)假设你的瓶子数的预算是某个给定的 k>2。描述一种至多使用 k 个瓶子找到最高安全横挡的策略。如果$f_k(n)$表示依照你的策略需要下落瓶子的次数，那么函数$f_1, f_2, f_3, \ldots$，应该有这样的性质，每个函数渐近增长比前面的函数要慢：

对每个 k，$\lim_{n\to\infty} f_k(n)/f_{k-1}(n) = 0$

4. **题目来源：**《算法设计》第三章第 4 题：

题目描述：你有一些朋友在业余时间研究蝴蝶，一天他们带了 n 只蝴蝶回来，他们知道每一只蝴蝶都属于两个不同种类中的一种。我们称这两个种类为 A 和 B。他们想把这 n 只蝴蝶每一只都归为 A 或者 B，但是直接标记每一只蝴蝶太困难了，因此他们决定采用如下的办法：对每一对蝴蝶 i 和 j，他们会仔细研究这两只蝴蝶是不是属于同一类的。如果他们能判断，就会记下这两只蝴蝶是同类的或者是不同类的。当然，某些蝴蝶对他们也可能判断不出来是不是同一类的。

现在他们有 n 只蝴蝶，还有 m 个已经有判断的蝴蝶对，他们想知道这 m 个判断当中是不是有矛盾。具体来讲，是否能够存在一种归类方式，使得所有 m 个判断都是正确的。他们觉得你能帮他们解决这个问题。请你给出一个运行时间为 $O(m+n)$ 的算法来确定是否有矛盾。**注意：这道题假设所有的蝴蝶只能属于两个种类中的一种。**

5. **题目来源：**《算法设计》第三章第 8 题

题目描述：

在关于因特网和 Web 结构的出版物中有许多报导集中于下面的问题：在这些网络中一般的结点离开有多远？如果你细心读这些报导，你会发现许多报导关于网络的直径与平均距离之间的区别是混淆的；他们常常在这些概念之间跳来跳去，尽管他们说的是同一个东西。

正如在本书里，我们看到在图 G=(V，E)中两个结点 u 与 v 之间的距离是在与它们相交的路径上的最少边数；我们将它记为$dist(u，v)$。我们说 G 的直径是在任意一对结点之间的最大距离；并且我们将把这个量记为$diam(G)$。

让我们定义一个相关的量，我们称之为 G 中的平均点对距离(标记为$apd(G)$)。我们定义$apd(G)$是在所有$\binom{n}{2}$个两两不同的结点 u 与 v 的集合上 u 与 v 距离的平均值。即

$$\left[\sum_{\{u,v\}\subseteq V} dist(u,v)\right] \Big/ \binom{n}{2}$$

这里是一个简单的例子，它使你确信存在图 G，其中$diam(G) \neq apd(G)$。设 G 是一个图，具有三个结点 u,v,w 以及两条边$\{u,v\}$和$\{v,w\}$，那么

$$diam(G) = dist(u,w) = 2$$

而

$$apd(G) = [dist(u，v) + dist(u，w) + dist(v，w)]/3 = 4/3$$

当然，这两个数在这三个结点的图的情况下差得不远，很自然会问是否在它们之间总是存在一种接近的关系。这里有一个试图使得这个结果更为精确的论断

命题:存在一个正的自然数 c 使得对所有的连通图 G 有：

$$\frac{diam(G)}{apd(G)} \leq c$$

你认为这个论断是真还是假给出一个证明或者一个反例。