

《暴躁的排序鸽》解题报告

太原市第五中学李嘉图

2018 年 10 月 28 日

1 题目大意

这是一道交互题：给定一个长度为 n 的数组和一个参数 k ，你拥有一个一次能将不超过 k 个数排序的排序器，请你用尽可能的次数给原数组排序。

每个子任务有三个评分参数 $L1, L2, L4$ ，设你使用的比较次数为 cmp ：

- 当 $cmp \leq L1$ 时，可以获得满分
- 当 $L1 < cmp \leq L2$ 时，可以获得一半的分数
- 当 $L2 < cmp \leq L4$ 时，可以获得四分之一的分数
- 当 $L4 < cmp$ 时，不能获得任何分数。

2 命题思路和期望得分分布

这道题的命题灵感来源于清华大学数据结构与算法课的一道课后作业：给定一个三元比较器，将1000000个数在不超过10003000 次操作内排序。而这道题将三元比较器扩展到了4元和16元。

作为一套题目中的简单题，本题期望有约四分之一的选手可以获得满分，至少三分之二的选手可以得到超过60分的成绩，全部选手可以获得超过30分的成绩。

3 解法

3.1 Subtask1

在这个测试点中, $n = 1024, k = 4$, 即要使用4元排序器给长度为1024的数组排序。

对于 $L2$ 和 $L4$ 的部分只要在普通的排序算法上稍加处理就可以通过, 因此不再赘述。考虑 $L1 = 5700$ 如何做。不妨将四元比较器看成两个二元比较器, 将序列分成两半, 同时做两边即可。这样排序所用的比较次数就是将512个数排序的比较次数, 加上将两边归并到一起的次数。

归并到一起需要1024次比较, 而每个排序不妨使用归并排序, 操作次数为 $1024 + 512 * \log_2 512 = 5632$ 次。

3.2 Subtask2

首先考虑如何做 $L4$ 。考虑改进快速排序, 将把序列分成两部分改为分成16路, 即随机选择15个分隔元素, 每次将一个元素与15个元素比较, 即可将序列划分成16段。那么操作次数大约是 $n \log_{16} n$ 次, 可以通过 $L4$ 这一档。

对于 $L1$ 和 $L2$, 我们发现操作次数的限制和 n 类似甚至比 n 要小, 那么每次取出一个的归并排序显然是行不通的, 考虑每次取出多个。考虑这样的算法: 当序列长度 ≤ 16 时, 直接排序; 否则将序列划分成两部分递归排序然后归并, 一次取出左边最靠前的8 个元素和右边最靠前的8 个元素, 整个数列的前八个元素一定出现在这16 个元素之中, 直接取出即可。总的比较次数是:

$$\frac{n}{8} \times \log_2 \left(\frac{n}{16} \right) + \frac{n}{16} = 832$$

另外, 还可以采用分两段的快速排序: 即每次取15个数和分割数比较, 分到序列两侧。这种方法虽然期望次数要略好一些, 但由于测试点很多, 出错的概率仍然不小。但达到 $L2$ 的界还是非常轻松的。

上面的两个算法都可以直接套用Subtask1。

3.3 Subtask3

直接套用上一个子任务的标准做法即可得到 $L2$ 的解。但在这个部分使

用分两段快速排序只能通过 $L4$ 的部分。

在这个子任务中最紧的比较次数是 $n + c$ 级别的，使用归并是不行的。考虑维护数据结构来做排序，最常见的即堆和线段树。堆的维护操作是自顶向下的，每次维护的复杂度很高，无法通过 $L1$ 部分；而线段树就好很多——考虑从顶到下修改一个元素的过程中，计算变化相关的元素只有 $\log_2(n)$ 个。如果我们能将这些元素一起计算，总操作次数就是 $O(n)$ 级别的了。

由于树高是 $\log_2 n + 1 = 20$ ，直接比较至少需要两次才能计算好。加上线段树的操作次数，总操作次数是刚好能通过 $L4$ 的部分。

但是如果我们结合Subtask2，先将序列按照长度为16分割成若干段，就变成了长度为 2^{15} 的线段树，这样就可以用一个比较器比较出来了，总的操作次数是 n +建树次数。建树也可以每次处理连续16个元素，总共的操作次数约为560000，可以获得满分。