国王饮水记 题目讨论

出题人：彭雨翔　验题人：余行江、王逸松

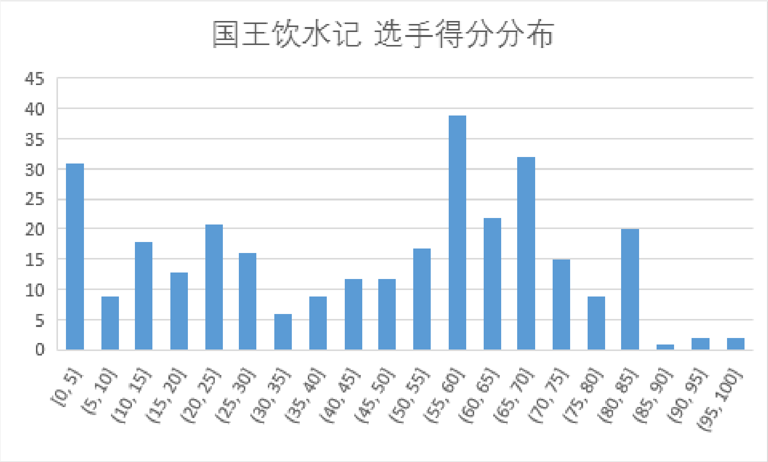
清华大学　交叉信息研究院 计算机科学与技术系

July 26, 2016

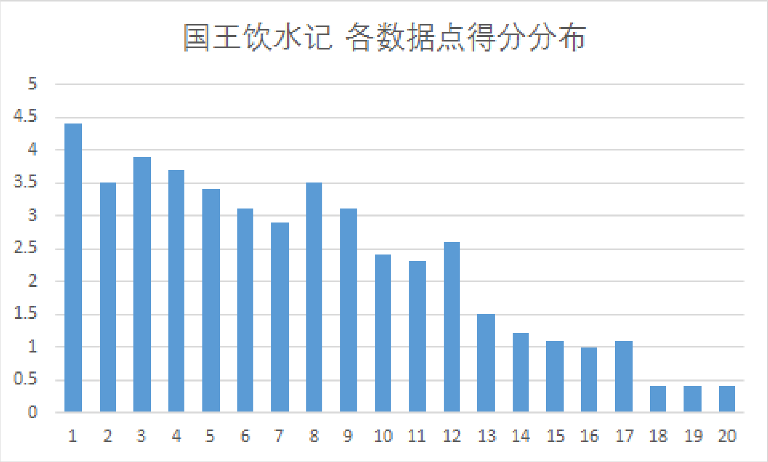
题目大意

- n 个水杯，装有不同高度的水h_i，每次可以指定任意多水杯用连通器连通一下，问k次操作之后1号水杯的最高水量。
- 由于一些原因，你需要输出p位小数，我们为你提供了一个高精度小数库。是不是很良心啊!
- $n \leq 8000, p \leq 3000, k \leq 10^9, h_i \leq 10^5$. 输入均为正整数。

统计数据


国王饮水记 选手得分分布
45
40
35
30
25
20
15
10
5
0
[0, 5]
(5, 10]
(10, 15]
(15, 20]
(20, 25]
(25, 30]
(30, 35]
(35, 40]
(40, 45]
(45, 50]
(50, 55]
(55, 60]
(60, 65]
(65, 70]
(70, 75]
(75, 80]
(80, 85]
(85, 90]
(90, 95]
(95, 100]

统计数据



国王饮水记 各数据点得分分布
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

统计数据

- 100分：辜俊儒、洪华敦。
- 91分：毛啸等2人。
- 86分：吴作凡等1辆火车。
- 85分：共15人。
- 82分：共5人。
- 平均分：45.84641
- 中位数：51

吐槽

- 本来是Day1的第一题！是不是很良心啊！

吐槽

- 本来是Day1的第一题！是不是很良心啊！
- 如果OI比赛要你写证明的话这题可以出到CTSC啊！是不是很良心啊！

吐槽

- 本来是Day1的第一题！是不是很良心啊！
- 如果OI比赛要你写证明的话这题可以出到CTSC啊！是不是很良心啊！
- 为了送分不择手段！是不是很良心啊！

吐槽

- 本来是Day1的第一题！是不是很良心啊!
- 如果OI比赛要你写证明的话这题可以出到CTSC啊!
  是不是很良心啊!
- 为了送分不择手段！是不是很良心啊!
- 为了防AC不择手段！是不是很良心啊!

选手讨论与吐槽

吐槽+1

- 为了讲题的流畅性，我们在讨论时仅给出定理，定理的证明我们挪到最后一起给出。
- (kan)尽(shi)量(jian)把所有定理都证完!

吐槽+1

- 为了讲题的流畅性，我们在讨论时仅给出定理，定理的证明我们挪到最后一起给出。
- (kan)尽(shi)量(jian)把所有定理都证完！
- 欢迎选手和我同台竞技，一会帮我证定理！

5分做法

- 首先，$n = 2$怎么做呢？

5分做法

- 首先，$n = 2$怎么做呢？
- 我会手算！

15分做法

- 那么，$n \le 4$怎么做呢？

15分做法

- 那么，$n \le 4$怎么做呢？
- 我会手算！

15分做法

- 那么，$n \leq 4$怎么做呢？
- 我会手算！
- 好吧我手算不出来，那我写个搜索吧！
- 直接枚举每一轮选哪几个城市，然后算个答案。
- 不会用库，那就用个double嘛。

15分做法

- 那么，$n \leq 4$怎么做呢？
- 我会手算！
- 好吧我手算不出来，那我写个搜索吧！
- 直接枚举每一轮选哪几个城市，然后算个答案。
- 不会用库，那就用个double嘛。
- 复杂度：$O(2^{nk})$。其实可以通过25分。
- 是不是很良心！

25分做法

- 考虑$k = 1$的点，此时的连接一定包含了1号城市。

25分做法

- 考虑$k = 1$的点，此时的连接一定包含了1号城市。
- 一个直接的idea是，选水量大的总是比水量小的要优。
- 排序后枚举选最大的i个，算一下答案即可。

定理1

- 所有水量小于等于1号城市水量的城市都不对答案产生影响。

定理2&定理3

- 以下叙述略去条件:“在最优方案中”。

定理2&定理3

- 以下叙述略去条件:“在最优方案中”。
- 定理2：除1号城市之外，每一个城市至多被连通一次。

定理2&定理3

- 以下叙述略去条件:“在最优方案中”。
- 定理2：除1号城市之外，每一个城市至多被连通一次。
- 定理3：每一次连通都一定和1号城市连通。

35分做法

- 证(cai)出定理2之后，我们可以轻易地提出一个状压DP的方法。

35分做法

- 证(cai)出定理2之后，我们可以轻易地提出一个状压DP的方法。
- 直接以n位二进制数表示每个城市是否使用过，写出状压DP的方程就好了。
- 注意到实际上是将两个子状态并出一个状态，可以在复杂度分析上用一些小技巧。

35分做法

- 证(cai)出定理2之后，我们可以轻易地提出一个状压DP的方法。
- 直接以n位二进制数表示每个城市是否使用过，写出状压DP的方程就好了。
- 注意到实际上是将两个子状态并出一个状态，可以在复杂度分析上用一些小技巧。
- 复杂度：$O(k3^n)$。

定理4

- 当$k \to \infty$时，最优方案为排序后由小到大将比1号城市大的城市与1号城市依次相接。

定理4

- 当$k \to \infty$时，最优方案为排序后由小到大将比1号城市大的城市与1号城市依次相接。
- 我们由此可以看出，$k > n$时，只需将k设置为n即可。以下令$K = \min\{k, n\}$。

45分做法

- 定理4也比较好猜出来，在猜出定理4之后，可以直接计算答案。

45分做法

- 定理4也比较好猜出来，在猜出定理4之后，可以直接计算答案。
- 写一个取最大K个的贪心。等等，为什么我的得分比45高这么多？

45分做法

- 定理4也比较好猜出来，在猜出定理4之后，可以直接计算答案。
- 写一个取最大K个的贪心。等等，为什么我的得分比45高这么多？
- 注意到当答案的精度有5位小数时，我们可以拿40%的分数。
- 考虑计算答案的式子：$a_i = (a_{i-1} + H_i) \times \frac{1}{2}$.
- 这和秦九韶算法中的多项式非常像，那么在多项式高次部分出现的项对答案的贡献被$\frac{1}{2}$的幂压得非常小了。
- 对答案产生了显著贡献的部分为最大的几个水杯。

45分做法

- 定理4也比较好猜出来，在猜出定理4之后，可以直接计算答案。
- 写一个取最大K个的贪心。等等，为什么我的得分比45高这么多？
- 注意到当答案的精度有5位小数时，我们可以拿40%的分数。
- 考虑计算答案的式子：$a_i = (a_{i-1} + H_i) \times \frac{1}{2}$.
- 这和秦九韶算法中的多项式非常像，那么在多项式高次部分出现的项对答案的贡献被$\frac{1}{2}$的幂压得非常小了。
- 对答案产生了显著贡献的部分为最大的几个水杯。
- 所以有很多额外的分！之后的题目分析中不再加上这一部分分数。

定理5、6、7

- 之后所有定理的叙述都不包含1号城市，且已经将城市按水量排序。

定理5、6、7

- 之后所有定理的叙述都不包含1号城市，且已经将城市按水量排序。
- 定理5：每次操作选择的城市的最小水量一定大于前一次操作所选择的城市的最大水量。

定理5、6、7

- 之后所有定理的叙述都不包含1号城市，且已经将城市按水量排序。
- 定理5：每次操作选择的城市的最小水量一定大于前一次操作所选择的城市的最大水量。
- 定理6：每次操作选择的城市均是选择排序后连续的一段城市（选择一个区间）。

定理5、6、7

- 之后所有定理的叙述都不包含1号城市，且已经将城市按水量排序。
- 定理5：每次操作选择的城市的最小水量一定大于前一次操作所选择的城市的最大水量。
- 定理6：每次操作选择的城市均是选择排序后连续的一段城市（选择一个区间）。
- 定理7：任意两次相邻操作选择的区间之间不存在城市（排列紧密）。

60分做法

- 有了定理6、7之后，我们终于可以清楚地看到一个非常传统的DP模型了！

60分做法

- 有了定理6、7之后，我们终于可以清楚地看到一个非常传统的DP模型了！
- 可以轻易地写出DP方程。令$f_{i,j}$表示到第i个城市，用了j次连通器后1号城市的最高水位。

$$f_{i,j} = \max_{k<i} \left\{ \frac{f_{k,j-1} + S_i - S_k}{i - k + 1} \right\}.$$

- 其中S_i表示排序之后城市水量的前缀和。
- 转移方程即枚举两次操作的分界点。

60分做法

- 有了定理6、7之后，我们终于可以清楚地看到一个非常传统的DP模型了！
- 可以轻易地写出DP方程。令$f_{i,j}$表示到第i个城市，用了j次连通器后1号城市的最高水位。

$$f_{i,j} = \max_{k<i} \left\{ \frac{f_{k,j-1} + S_i - S_k}{i - k + 1} \right\}.$$

- 其中S_i表示排序之后城市水量的前缀和。
- 转移方程即枚举两次操作的分界点。
- 复杂度：$O(n^2 Kp)$.

70分做法

- 稍有常识的人都知道，这是一个明显的斜率方程！整理之后写出：

$$f_{i,j}=\max_{k<i}\left\{\frac{S_i-(S_k-f_{k,j-1})}{i-(k-1)}\right\}.$$

70分做法

- 稍有常识的人都知道，这是一个明显的斜率方程！整理之后写出：

$$f_{i,j} = \max_{k<i}\left\{\frac{S_i - (S_k - f_{k,j-1})}{i-(k-1)}\right\}.$$

- 凸包的横坐标为$k-1$，纵坐标为$S_k - f_{k,j-1}$。
- 我们只需要求经过点(i, S_i)的凸包的切点即可。
- 维护上一层的凸包的同时，用三分求一个极值点即可。

70分做法

- 稍有常识的人都知道，这是一个明显的斜率方程！整理之后写出：

$$f_{i,j} = \max_{k<i}\left\{\frac{S_i - (S_k - f_{k,j-1})}{i - (k-1)}\right\}.$$

- 凸包的横坐标为$k-1$，纵坐标为$S_k - f_{k,j-1}$。
- 我们只需要求经过点(i, S_i)的凸包的切点即可。
- 维护上一层的凸包的同时，用三分求一个极值点即可。
- 复杂度：$O(nKp\log n)$.

定理8

- 当对于$f_{i,j}$，决策在k点比决策在$l < k$点优时，对于$f_{i+1,j}$，决策在k点也比决策在l优。

85分做法

- 稍有常识的人都知道，这时候我们可以打表找规律。

85分做法

- 稍有常识的人都知道，这时候我们可以打表找规律。
- 通过证(zhao)定(gui)理(lv)，我们可以发现在同一层内DP的决策是单调的。

85分做法

- 稍有常识的人都知道，这时候我们可以打表找规律。
- 通过证(zhao)定(gui)理(lv)，我们可以发现在同一层内DP的决策是单调的。
- 利用决策单调性在凸包上进行一个单调性优化，就可以将每一次决策的复杂度降到均摊$O(1)$。
- 复杂度：$O(nKp)$.

85分做法

- 稍有常识的人都知道，这时候我们可以打表找规律。
- 通过证(zhao)定(gui)理(lv)，我们可以发现在同一层内DP的决策是单调的。
- 利用决策单调性在凸包上进行一个单调性优化，就可以将每一次决策的复杂度降到均摊$O(1)$。
- 复杂度：$O(nKp)$.
- 什么？复杂度还可以优化？

85分做法

- 稍有常识的人都知道，这时候我们可以打表找规律。
- 通过证(zhao)定(gui)理(lv)，我们可以发现在同一层内DP的决策是单调的。
- 利用决策单调性在凸包上进行一个单调性优化，就可以将每一次决策的复杂度降到均摊$O(1)$。
- 复杂度：$O(nKp)$.
- 什么？复杂度还可以优化？
- 什么？题面里还有一个条件我们没用上？

85分做法

- 稍有常识的人都知道，这时候我们可以打表找规律。
- 通过证(zhao)定(gui)理(lv)，我们可以发现在同一层内DP的决策是单调的。
- 利用决策单调性在凸包上进行一个单调性优化，就可以将每一次决策的复杂度降到均摊$O(1)$。
- 复杂度：$O(nKp)$.
- 什么？复杂度还可以优化？
- 什么？题面里还有一个条件我们没用上？
- 什么？为什么所有水量高度互不相同？

定理9&定理10

- 定理9：每一次操作的区间长度一定不比上一次操作的区间长度长。

定理9&定理10

- 定理9：每一次操作的区间长度一定不比上一次操作的区间长度长。
- 定理10：在所有水量高度互不相同的情况下，长度大于1的区间仅有$O(\log \frac{nh}{\Delta})$个，其中$\Delta = \min_i\{h_i - h_{i-1}\}$.

95分做法

- 如果你试图将方案打出来，你会发现，很快所有的决策都会收敛至长度为1。
- 于是我们继续通过打(zheng)表(ming)找(ding)规(li)律，可以将复杂度降下来。

95分做法

- 如果你试图将方案打出来，你会发现，很快所有的决策都会收敛至长度为1。
- 于是我们继续通过打(zheng)表(ming)找(ding)规(li)律，可以将复杂度降下来。
- 由于在此题中水量都是互不相等的正整数，$\Delta \geq 1$.
- 那么我们只需要对前$O(\log nh)$层（进行精确的计算后可以得知为14层）进行DP，之后用DP的结果将答案算出来即可。

95分做法

- 如果你试图将方案打出来，你会发现，很快所有的决策都会收敛至长度为1。
- 于是我们继续通过打(zheng)表(ming)找(ding)规(li)律，可以将复杂度降下来。
- 由于在此题中水量都是互不相等的正整数，$\Delta \geq 1$.
- 那么我们只需要对前$O(\log nh)$层（进行精确的计算后可以得知为14层）进行DP，之后用DP的结果将答案算出来即可。
- 复杂度：$O(np\log nh)$.

95分做法

- 如果你试图将方案打出来，你会发现，很快所有的决策都会收敛至长度为1。
- 于是我们继续通过打(zheng)表(ming)找(ding)规(li)律，可以将复杂度降下来。
- 由于在此题中水量都是互不相等的正整数，$\Delta \geq 1$.
- 那么我们只需要对前$O(\log nh)$层（进行精确的计算后可以得知为14层）进行DP，之后用DP的结果将答案算出来即可。
- 复杂度：$O(np \log nh)$.
- 什么？复杂度还可以更低？

100分做法

- 注意到上述算法的巨量时间消耗均在高精度小数的计算上。
- 然而由于DP只进行了$O(\log nh)$层，根据DP式可以看出来每一位上的结果至多曾经进行了$O(\log nh)$次与低精度数进行的运算。

100分做法

- 注意到上述算法的巨量时间消耗均在高精度小数的计算上。
- 然而由于DP只进行了$O(\log nh)$层，根据DP式可以看出来每一位上的结果至多曾经进行了$O(\log nh)$次与低精度数进行的运算。
- 假如我们不再使用定点高精度小数，而是使用分数类，可以发现所得到的分子分母是$O(n^{\log nh})$级别的。
- 那么使用分数类进行计算时，复杂度就将每次运算降到了$O(\log^2 nh)$.（理论上可以用FFT优化到$O(\log nh \log\log nh)$，但由于长度非常小，意义不大。）

100分做法

- 注意到上述算法的巨量时间消耗均在高精度小数的计算上。
- 然而由于DP只进行了$O(\log nh)$层，根据DP式可以看出来每一位上的结果至多曾经进行了$O(\log nh)$次与低精度数进行的运算。
- 假如我们不再使用定点高精度小数，而是使用分数类，可以发现所得到的分子分母是$O(n^{\log nh})$级别的。
- 那么使用分数类进行计算时，复杂度就将每次运算降到了$O(\log^2 nh)$.（理论上可以用FFT优化到$O(\log nh \log\log nh)$，但由于长度非常小，意义不大。）
- 这样的复杂度就最终降到了：$O(n(\log^3 nh + p))$.

来自场外的想法

- 机智的xyz大爷注意到了一个更厉害的性质。（不过我没有证明）

来自场外的想法

- 机智的xyz大爷注意到了一个更厉害的性质。（不过我没有证明）
- 实际上，每个状态的决策点的改变对答案影响很大，而前几层决策对这一层的决策影响较小。（即上两层的较远决策的影响不如上一层的较近决策影响大）
- 我们可以利用高度互不相同来界出影响，则实际上大概只需要保留几十位小数，就能得到正确的转移。
- 之后再根据转移算出答案即可。

Q&A&Q

Q&A&Q

你也可以趁机吐槽。

Q&A&Q

你也可以趁机吐槽。

接下来我们来证(cai)定(jie)理(lun)吧！

定理1

所有水量小于等于1号城市水量的城市都不对答案产生影响。

- 首先，如果凭空变大一个城市的水量，答案一定不更差。

定理1

所有水量小于等于1号城市水量的城市都不对答案产生影响。

- 首先，如果凭空变大一个城市的水量，答案一定不更差。
- 若某次操作包含这些城市，将这些城市替换成1号城市。
- 若1号城市在某次操作中出现多次，则只保留一次。

定理1

所有水量小于等于1号城市水量的城市都不对答案产生影响。

- 首先，如果凭空变大一个城市的水量，答案一定不更差。
- 若某次操作包含这些城市，将这些城市替换成1号城市。
- 若1号城市在某次操作中出现多次，则只保留一次。
- 在这样的情况下，答案不会更差。

定理2&定理3

定理2：除1号城市之外，每一个城市至多被连通一次。

定理3：每一次连通都一定和1号城市连通。

- 这两个定理是等价的。

定理2&定理3

定理2：除1号城市之外，每一个城市至多被连通一次。

定理3：每一次连通都一定和1号城市连通。

- 这两个定理是等价的。
- 定理2 $\Rightarrow$ 定理3：一次连接如果未和1号城市连接，则之后这些城市就没有意义了。可以删去这次操作。
- 定理3 $\Rightarrow$ 定理2：一个城市若和1号城市连通一次，此时其水量等于1号城市，不会出现第二次。

定理2&定理3

- 我们采用归纳法来同时证明这两个定理。当只有一次操作时显然。
- 不妨假设对于最后m个操作，他们均成立，接下来只要考虑倒数第$m+1$个操作。

定理2&定理3

- 我们采用归纳法来同时证明这两个定理。当只有一次操作时显然。
- 不妨假设对于最后m个操作，他们均成立，接下来只要考虑倒数第$m+1$个操作。
- 若该操作包含了1号城市，则由等价性可直接得出结论。

定理2&定理3

- 我们采用归纳法来同时证明这两个定理。当只有一次操作时显然。
- 不妨假设对于最后m个操作，他们均成立，接下来只要考虑倒数第$m+1$个操作。
- 若该操作包含了1号城市，则由等价性可直接得出结论。

定理2&定理3

- 若该操作不包含1号城市，注意到在这次操作之后，与本次操作相关联的水杯都变为了同一高度，即我们可以任意交换他们在之后的编号。
- 假设存在一些城市在之后也用到了。由于后m个操作都满足这两个定理，我们可以发现：通过交换编号，可以使得这些城市的水量自出现顺序从后向前而由大到小变化。且此时删去第$m+1$次操作所得答案更优。
- 证毕。

定理4

当$k \to \infty$时，最优方案为排序后由小到大将比1号城市大的城市与1号城市依次相接。

- 注意到有无穷多次操作时，任意一个操作都可以拆成无数次某城市与1号城市的连接。此时只需要考虑所有选择1号城市和某城市的操作即可。
- 显然一个城市连接完之后就废了，且任意两个城市一定是水量小的先连接。那么只要按顺序连接一下就好了。

定理5、6、7

每次操作选择的城市的最小水量一定大于前一次操作所选择的城市的最大水量。每次操作选择的城市均是选择排序后连续的一段城市（选择一个区间）。任意两次相邻操作选择的区间之间不存在城市（排列紧密）。

- 定理5证明：否则，交换两个逆序城市以进行调整，显然可以变得更优。

定理5、6、7

每次操作选择的城市的最小水量一定大于前一次操作所选择的城市的最大水量。每次操作选择的城市均是选择排序后连续的一段城市（选择一个区间）。任意两次相邻操作选择的区间之间不存在城市（排列紧密）。

- 定理5证明：否则，交换两个逆序城市以进行调整，显然可以变得更优。
- 定理6证明：否则，将最小水量的城市删去并选择一个断点处的城市，显然可以更优。

定理5、6、7

每次操作选择的城市的最小水量一定大于前一次操作所选择的城市的最大水量。每次操作选择的城市均是选择排序后连续的一段城市（选择一个区间）。任意两次相邻操作选择的区间之间不存在城市（排列紧密）。

- 定理5证明：否则，交换两个逆序城市以进行调整，显然可以变得更优。
- 定理6证明：否则，将最小水量的城市删去并选择一个断点处的城市，显然可以更优。
- 定理7证明：否则，将靠左侧的区间右移一位，显然更优。

定理8

当对于$f_{i,j}$，决策在k点比决策在$l<k$点优时，对于$f_{i+1,j}$，决策在k点也比决策在l优。

- 注意到DP方程式为：

$$f_{i,j}=\max_{k<i}\left\{\frac{f_{k,j-1}+S_i-S_k}{i-k+1}\right\}.$$

- 即此时有不等式：

$$(i-l+1)(S_i-S_k+f_{k,j-1})\geq(i-k+1)(S_i-S_l+f_{l,j-1})$$

- 整理得：

$$(k-l)(S_i-S_k+f_{k,j-1})\geq(i-k+1)(S_k-S_l-(f_{k,j-1}-f_{l,j-1}))$$

定理8

当对于$f_{i,j}$，决策在k点比决策在$l < k$点优时，对于$f_{i+1,j}$，决策在k点也比决策在l优。

- 由定义，$f_{k,j-1} \geq f_{l,j-1}$。且由于已经排序，$h_{i+1} \geq h_i \geq h_{i-1} \geq \cdots \geq h_k$。不等式两边加上$(k-l)h_{i+1}$，得：

$$
\begin{aligned}
(k-l)(S_{i+1} - S_k + f_{k,j-1}) \geq \quad & (i-k+1)h_k + (k-l)h_{i+1} \\
& - (i-k+1)(f_{k,j-1} - f_{k-1,j-1}) \\
\geq & (i-l+1)(h_k - (f_{k,j-1} - f_{k-1,j-1})).
\end{aligned}
$$

- 证毕。

定理9

每一次操作的区间长度一定不比上一次操作的区间长度长。

- 当存在相邻两个操作的区间长度上升时，将后一个区间的第一个城市放入前一个区间当中，答案会更优。

定理10

在所有水量高度互不相同的情况下，长度大于1的区间仅有$O(\log \frac{nh}{\Delta})$个，其中$\Delta = \min_i\{h_i - h_{i-1}\}$.

- 考虑长度大于1的所有区间，假设为l段。我们比较此时的方案与去掉第一个区间、将最后一个区间的最后一个城市单独拿出来的方案。

定理10

在所有水量高度互不相同的情况下，长度大于1的区间仅有$O(\log \frac{nh}{\Delta})$个，其中$\Delta = \min_i\{h_i - h_{i-1}\}$.

- 考虑长度大于1的所有区间，假设为l段。我们比较此时的方案与去掉第一个区间、将最后一个区间的最后一个城市单独拿出来的方案。
- 列出不等式，利用所有高度互不相等，可知最后一个城市之前的任意一段的平均值都小于最后一个城市。

定理10

在所有水量高度互不相同的情况下，长度大于1的区间仅有$O(\log \frac{nh}{\Delta})$个，其中$\Delta = \min_i\{h_i - h_{i-1}\}$.

- 将不等式整理之后可以得到：

$$\frac{\Sigma_1}{\prod_{i=1}^{l}(L_i + 1)} \le \frac{L_l - 1}{L_l}\Delta.$$

- 由于$L \ge 2$，我们可以利用上述不等式将l给界住，即：

$$l \le \log_3 \frac{2\sum_i h_i}{\Delta} \le 14.$$

定理10

在所有水量高度互不相同的情况下，长度大于1的区间仅有$O(\log \frac{nh}{\Delta})$个，其中$\Delta = \min_i\{h_i - h_{i-1}\}$.

- 将不等式整理之后可以得到：

$$\frac{\Sigma_1}{\prod_{i=1}^{l}(L_i+1)} \le \frac{L_l - 1}{L_l}\Delta.$$

- 由于$L \ge 2$，我们可以利用上述不等式将l给界住，即：

$$l \le \log_3 \frac{2\sum_i h_i}{\Delta} \le 14.$$

- 即仅有$O(\log \frac{nh}{\Delta})$个长度大于1的区间。

Q&A&Q