文章编号:1007-2861(2001)04-0365-06

称球的整数规划问题*

孙楚仁， 张连生

(上海大学 理学院,上海 200436)

摘要:分而治之算法是数据结构中解决最大最小问题、排序问题的有效算法之一. 它能够对一个问题很快地设计一种计算步骤进行计算从而解决问题,但是对于解决这些问题所需最少操作数这个问题,它并不能给出实质的解答. 该文考察分而治之算法应用的一个问题并将它一般化,通过分析对这种更一般问题建立了求解该问题所需最少操作数的整数规划模型,通过求解该模型,能够得出解决这种更一般问题所需要的最少操作数.

关键词:算法;称球问题;操作数;整数规划
中图分类号:O 221　　文献标识码:A

A Problem of Integer Programming for Classifying Balls

SUN Chu-ren，　ZHANG Lian-sheng
(School of Sciences, Shanghai University, Shanghai 200436, China)

Abstract: Dividing and ruling is an effect algorithm in solving maxminmum and sorting problems. It can effectively generate a computation procedure but can not provide the number of operations it needs to solve such a problem. This article considers a problem to which dividing and ruling algorithm applies, and then generalizes the problem. By analysing the more general problem this article constructs an integer programming model for solving the least operation numbers.

Key words: algorithm; dividing and ruling; operation numbers; integer programming

1 引　言

我们考虑这样一个一般问题:

问题 1.1　设有 n 个形状颜色都一样的小球,其中有 θ 类 m 个球与其他 $n-m$ 个球质量不一样,它们的个数和质量分别为 $s_1, s_2, \cdots, s_\theta$ 以及 $m_1, m_2, \cdots, m_\theta$,并设其它 $n-m$ 个小球每个球的质量是 m_0 (为简化问题计,设 $m_0 < m_i, i=1,2,\cdots,\theta$). 现有一架无刻度但称量范围无限大的敏感天平,问至少要称量几次才能将此 θ 类球给找出来?文献[1]给出了问题 1.1 的一个具体形式.

问题 1.2　给你一个装有 16 个硬币的袋子,16 个硬币有一个是伪造的,并且那个伪币的硬币比真的硬币要轻一些. 现有一台可以比较两组硬币质量的仪器,利用该仪器,可以知道两组硬币的重量是否相同. 现在问至少需要比较几次才能把该硬币找出来?

我们知道,在计算机科学中有一种分而治之的算法可以编程计算找出问题 1.2 中的伪币[1],但是它没有给出找出这枚伪币所需要的至少的比较次数是多少. 它那里的分而治之的方法,是假设 n 为 2 的方幂,并且简单地把硬币分成相等的 2 份,通过比较

* 收稿日期:2001-05-09　作者简介:孙楚仁(1976～),男,湖南宜章人,硕士生,主要从事数学规划的研究.

此2份硬币的重量判断出伪币属于哪一份,然后依此类推,最后找出伪币.但是这并不能说明这样的算法的比较次数是找出伪币所需要的最少的比较次数.事实上我们可能把硬币分成更多的份比如3份、4份等等而使比较次数最少.问题1.2仍然没有解决.

本文考察上述问题并对上述问题建立数学模型.通过对下述数学模型的求解,上述问题将被彻底地解决.

2 问题的一般假定和基本原理

我们对上述问题作几个基本假定:

假定1 $m \ll n$,且 $m_0 < m_i, i=1,2,\cdots,\theta$.

这不会改变问题的本质,因为为了使操作数最少,我们先分离球个数较少的类;且对于在引言中提出的问题,只要把其它 $n-m$ 个质量为 m_0 的球与这 n 个球中质量最小的球类调换一下其地位即可.以下假定 $m_0 < m_i, i=1,2,\cdots,\theta$.

假定2 天平称量过程中,可以在天平两端放这 n 个球中的任意个球的组合,亦即从这 n 个球中取球称量时每个球被取到的概率是相等的.

显然,为找出与这 n 个球中 $n-m$ 个球质量不同的其它球,就必须对球进行比较称量.例如,我们可以两个球两个球地称量,通过比较找出 n 个球中质量最小的球,然后再以此球为基准球与其他球比较即可找出质量最小的球,在重复进行上述步骤即可把这 $\theta+1$ 类球给区分开来.注意到在每一步操作中,根据天平的属性,当且仅当天平两边的球数相同时能够判断哪边的球堆具有更多的不属于 $n-m$ 个球中的球,当且仅当天平两边只有一个球时能够知道两个球谁的质量小一些.但是我们不知道这样做是否就是所需操作数最小的策略,因为也可以 k 个球 k 个球地称量.我们的任务事实上就是如何称量才能是把这 $\theta+1$ 类球区分出来所需的步数最小.

以这样的观点看待我们的称量方式:对于把球区分开来的所取球的操作,可以看成先把这 n 个球进行分堆后才进行比较称量.这样的观点是合理的.因为取球称量是随机的,对球分堆自然也是随机的.但是球一旦分堆,每个球所属的堆即确定.

假定3 在分球过程中球可以任意组合,即每个球可以分到每一堆中;分出的球堆可以作编号.

要解决的就是如何对球进行分堆.需要下述基本事实:

鸽巢原理[2] 设 A 是 $n(n\geqslant 2)$ 元素的集合,$A_i \subset A(i=1,2,\cdots,m)$ 且 $\bigcup_{i=1}^{m} A_i = A$,则必有正整数 $k(1\leqslant k\leqslant m)$ 使得 $|A_k| \geqslant \left[\frac{n-1}{m}\right]+1$.

整数分解定理 设 $n,m\in N$,则$\exists\ k,r\in N$ 使得 $n=km+r$,其中 $0\leqslant r<m$.

3 $\theta=1$ 的情况

$\theta=1$ 时,n 个球中只有2类球.显然此时最简单的一种称量方式就是两个球两个球地称,只要天平两边的球属于不同类,就立刻可以把该二球分离开来,但是至少需要 $\left[\frac{n}{2}+1\right]$ 次.当 n 很大时,这是一个很大的数.事实上是有其它称量方法的.例如:当 $n=11, m=1$ 时,最少的称量方式所需要的称量次数是3次,但是两个球两个球地称需要6次.因此需要找出另外的更好的称量方式.

为了把这两类球分离出来我们在称量时需要对球进行分堆,但是根据天平的特点对球分堆时又该如何分堆才能使把这两类球分离出来所需操作数最少呢?

设把 n 个球分成 s 堆,根据假定3,这样不带来实际操作数的增加,设其编号分别是 $1,2,\cdots,s$,每堆球的个数分别是 $s_i, i=1,2,\cdots,s$.相应地有以下定理:

定理3.1 能够把球分离当且仅当 $s>m$.

证明:

充分性 $s>m$ 时,根据鸽巢原理,至少有一堆不含有质量为 m_1 的球.此时根据天平的性质,能够找出这样的球堆.

必要性 假如能够把这两类球分离出来而 $s\leqslant m$,则由假定2,可能这 s 堆球中每堆球都含有质量为 m_1 的球,由天平的性质,不能分离这两类球.这样就必须再对这 n 个球分堆.重复上述过程,s 将逐步增大至 $s>m$ 为止,所以 s 必将会大于 m. □

定理3.2 为使把质量分为 m_0 与 m_1 的两类球分离开来所用操作数最少,在每次分堆后需要把不含质量为 m_1 的球堆全部分离开来.

证明 既然我们在分离质量为 m_0、m_1 的两类球需要对球进行分堆,而分堆数 s 要满足 $s>m$ 时方可保证把这两类球分离开来所需要的操作数最少,这样一次分堆是不能够把这两类球完全分开的,所以需要多次分堆.每分堆一次,需要把不含质量为

m_1 的球堆全部分离开来方可保证把这两类球分离开来总操作数最少. 事实上,因为要把不含质量为 m_1 的球堆分离开来,在最坏的情况下,可能需要比较完所有的球堆. 当我们找出一堆不含质量为 m_1 的球堆,则至多再需要比较 $s-1$ 次即可以把所有的不含有质量为 m_1 的球堆分离出来. 若不进行比较不把所有不含质量为 m_1 的球堆分离出来,则这些球堆将会参与下一次分堆与比较称量. 为了使把两类球分离出来所用的操作数最少,需要前一次分堆和比较称量与后一次分堆和比较称量在不论分堆初始球的情况下操作过程一样,每一次分堆与比较称量分离出来的每个球所用的操作数最小. 这样如果前一次分离中把不含质量为 m_1 的球堆分离开来分离出球堆的每个球所用的操作数最少而不必把所有的不含质量为 m_1 的球的球堆完全分离开来,则在后一次分离中也需要不把不含质量为 m_1 的球的球堆完全分离出来. 这样的结果是总会有质量为 m_0 的球留下来,从而不能够把两类球给分离开来. 这与我们的要求矛盾. □

定义 3.1　对球进行一次分 $s(s>m)$ 堆后把这些堆中所有的不含质量为 m_1 的小球分离开来的一些操作组成的集合称为一个分离过程.

定理 3.3　要使对球分堆时分球操作数最少,s 需满足 $s>m, n=(s-1)k+r, 0\leqslant r<k, k,r,s\in N$,且 $\exists\ j, 0<j\leqslant s$ 使 $s_j=r$ 而 $s_i=k, i=1,2,\cdots,s, i\neq j$.

证明　反证法. 根据天平的性质,为判断天平两边哪边含质量为 m_1 的小球数多需当天平两端个小球个数相等. 假如至少有 2 组球个数不等,不妨就设为 2 组. 由于每个小球都可能落入每一堆球中,可以设这两堆球中都含有质量为 m_1 的小球. 对其它 $s-2$ 堆球,因为其个数相等,可以任意取两堆称量,并把称量的结果记录下来. 显然,经过若干次操作之后,我们能够对各堆球的质量大小作一排序. 同时也可以在这 2 堆球中取两个球数相等的子堆进行称量,并把称量结果记录下来. 为了能尽量减少称量次数,就要使这两堆球在取出两个子堆的小球称量之后所剩的球个数最小,从而需重新对这两堆球分堆,最后分成两堆球个数相等的堆和只剩下一个小球的堆. 前面 $s-2$ 堆球称量完毕得出了质量大小的次序后,仍需要对另外两堆称量. 由假设 2,这两堆球可能质量相等,如果它们质量相等,不能判断出这两堆球哪堆球含有更多的质量为 m_1 的小球. 若 $s=m+1$ 或 $s=m+2$ 而其他 $s-2$ 堆小球质量相等,并假设剩下的一个小球是质量为 m_0 的小球,则我们不能判断出除去剩下的一个小球后的球堆哪堆不含有质量为 m_1 的小球,从而问题事实上没有获得任何进展. 若 $s>m+2$,则前面 $s-2$ 堆球至少有一堆球不含有质量为 m_1 的小球,当称量完毕得出各球堆的质量大小的次序后,我们即可找出这样的一堆球,以此为基准堆即可以找出这 $s-2$ 堆球中所有的不含质量为 m_1 的小球. 假如另外两堆新分出来的堆质量仍然相等,我们仍然不能对哪堆球不含有质量为 m_1 做出判断,这样,为排除最多的质量为 m_1 的球,就需要重新对这两堆球加剩下的那个小球分堆,使得能够从它们中分出与前面 $s-2$ 堆球个数相等的堆以使能与基准球堆进行比较. 假如这两堆球加剩下的一个球其总球数小于前面 $s-2$ 堆球每堆的球数,由整数分解定理问题即可获证. 不然,则根据前面所述分堆,直至分到所剩下的球数小于前面所述 $s-2$ 堆每堆的球数为止. 根据整数分解定理,并根据分堆不算入一个操作之中,这正是定理所要证明的结论. □

对球分堆,如果不考虑分堆的中间过程而仅仅考虑分堆的最终结果,我们只有对球按定理 3.2 所述的方式分堆才能使把这两类球分离出来所用的操作数最少. 问题在于,我们究竟要对这些球分多少堆才能使操作数最少呢?

假如已经把球分成定理 3.3 所述的情况. 我们来看为把不含有质量为 m_1 的球堆给找出来需要的操作数.

对于球个数相等的 $s(s>m)$ 堆球,为把质量为 m_1 的球给找出来,在最坏的情况下,可能要称量所有球堆. 无论哪种称量方式,由天平只能判别两端哪端含有质量为 m_1 的球多而不能判别出到底多多少个小球的性质,我们可能需要比较完所有的球堆. 因为天平每次只能比较两堆小球,所以事实上需要对这些球堆进行排序,而原来的问题变成如何排序才能使操作数最少. 这样,只要解决两个问题:其一为如何对球分堆, 其二为如何排序.

我们来建立在一次分离出来不含有质量为 m_1 的小球的球堆所需最少操作数的数学模型.

首先对于这 $bp_1=s$ 堆球进行编号,使得能够区分每一个球堆,记此 s 堆编号之后的球堆组成的集合为 S_1. 从中任意取两堆称量,这样两两比较,所需的操作数为 $\left[o_1=\dfrac{bp_1}{2}\right]$,比较的结果显然只有两种:

天平两端平衡或是不平衡. 记第一次比较的结果为：使天平两端平衡的球堆为 $2a_1$；使天平两端不平衡的球堆为 $2b_1$. 因为使天平两端不平衡的球堆中一定含有质量为 m_1 的小球，所以可以把重的一端放到一边，在第二次比较称量中不参与称量. 而对于使天平平衡的球堆也把一端放在一边不参与第二次的称量. 并设使天平两端平衡的球堆组成的集合为 B_1，使天平两端不平衡的球堆组成的集合为 UB_1，从而有 $S_1 = B_1 \cup UB_1$. 这样，如果 s 为偶数，则第二次比较称量时开始的球堆就为 $bp_2 = a_1 + b_1$，如果 s 为奇数，则第二次比较称量时开始的球堆数就为 $bp_2 = a_1 + b_1 + 1$，记第二次参与称量的球堆组成的集合为 S_2，重复第一次比较称量的操作，有第二次所需的操作数为 $\left[o_2 = \dfrac{bp_2}{2}\right]$，并记第二次操作完毕时使天平两端平衡的球堆为 $2a_2$，其集合为 B_2，使天平两端不平衡的球堆为 $2b_2$，其集合为 UB_2. 其中使天平两端平衡的球堆组成的集合分别为 B_{21}, B_{22}，其中含有 B_1 的球堆的个数分别为 c_{21}, c_{22}，含有 UB_1 的球堆的个数为 d_{21}, d_{22}，使天平两端不平衡的较重的球堆组成的集合分别为 UB_{21}，其中含有 B_1, UB_1 的球堆的个数分别为 e_{21}, e_{22}，使天平两端不平衡的较轻的集合为 UB_{22}，其中含有 B_1, UB_1 的球堆的个数分别为 f_{21}, f_{22}. 从而有 $S_2 = B_2 \cup UB_2, B_2 = B_{21} \cup B_{22}, UB_2 = UB_{21} \cup UB_{22}, 2a_2 = c_{21} + c_{22} + d_{21} + d_{22}, 2b_2 = e_{21} + e_{22} + f_{21} + f_{22}, S_2 \subset S_1$. 显然，在判别出 UB_{21} 中含有质量为 m_1 的小球时，由我们对球堆进行的编号也可以判断出 B_1 中 e_{21} 个球含有质量为 m_1 的小球，因而也可以把它们放在一边. 这样在第三次称量时初始球堆数就变成了：当 $0 = bp_2 \bmod 2$ 时，$bp_3 = a_2 + b_2 - e_{21}$；当 $1 = bp_2 \bmod 2$ 时，$bP_3 = a_2 + b_2 + 1 - e_{21}$. 依次进行下去，最终可以把含有质量为 m_1 的球堆全部给分离出来，设总共经过 l 次比较称量完成不含质量为 m_1 的球堆的分离，并记第 i 次开始时的球堆为 bp_i，第 i 次比较称量的操作数为 o_i，第 i 次比较称量完毕使天平两端平衡的球堆数为 $2a_i$，其集合为 B_i，使天平两端不平衡的球堆数为 $2b_2$，其集合为 UB_i. 其中使天平两端平衡的球堆组成的集合分别为 B_{i1}, B_{i2}，含有 B_{i-1} 的球堆的个数分别为 c_{i1}, c_{i2}，含有 UB_{i-1} 的球堆的个数为 d_{i1}, d_{i2}，使天平两端不平衡的较重的球堆组成的集合分别为 UB_{i1}，其中含有 B_i，UB_i 的球堆的个数分别为 e_{i1}, e_{i2}，使天平两端不平衡的较轻的集合为 UB_{i2}，其中含有 B_i, UB_i 的球堆的个数分别为 f_{i1}, f_{i2}. 从而有 $S_i = B_i \cup UB_i, B_i = B_{i1} \cup B_{i2}, UB_i = UB_{i1} \cup UB_{i2}, S_i \subset S_{i-1}, 2a_i = c_{i1} + c_{i2} + d_{i1} + d_{i2}, 2b_i = e_{i1} + e_{i2} + f_{i1} + f_{i2}$. 显然，在判别出 UB_{i1} 中含有质量为 m_1 的小球时，由我们对球堆的编号也可以判断出 B_i 中 e_{i1} 个球含有质量为 m_1 的小球，因而也可以把它们放在一边. 第 i 次可以分离出的球堆数为 $b_i + e_{i-1,1}$. 依次进行下去，然后返回比较我们在称量过程中未确定是否含有质量为 m_1 的小球的球堆，最终可以把不含有质量为 m_1 的小球的球堆全部分离出来. 则有：

$o_i = \dfrac{bp_i}{2}$；当 $0 = bp_{i-1} \bmod 2$ 时，$bp_i = a_{i-1} + b_{i-1} - e_{i1}$；当 $1 = bp_{i-1} \bmod 2$ 时，$bp_i = a_{i-1} + b_{i-1} + 1 - e_{i1}$；而到第 i 次为止总共所需的操作数为 $so_i = o_1 + o_2 + \cdots + o_i$，我们可以分离出来的不含质量为 m_1 的球堆数为 $sep_i = b_1 + b_2 + \cdots + b_i + e_{21} + e_{31} + e_{(i-1)1}$.

经过 l 次这样的过程，可以至少找出一个不含有质量为 m_1 的球堆，然后返回比较在称量过程中未确定是否含有质量为 m_1 的小球的球堆，既可以把所有不含有质量为 m_1 的小球的球堆找出来. 返回比较所需要的操作数为 $s - b_1 - b_2 - \cdots - b_l - e_{21} - e_{31} - \cdots - e_{l1}$. 可以分离出 $sep_l = b_1 + b_2 + \cdots + b_l + e_{21} + e_{31} + \cdots + e_{l1}$ 堆不含有质量为 m_1 的小球的球堆，球的个数为 $k \times sep_l$，而所要的总操作数为 $so_l = o_1 + o_2 + \cdots + o_l + s - b_1 - b_2 - \cdots - b_l - e_{21} - e_{31} - \cdots - e_{l1}$，从而为分离出一个质量为 m_0 的球，所要的操作数为 $\dfrac{so_l}{k \times sep_l}$. 由我们的问题，需要寻找这样的过程使 $\dfrac{so_l}{k \times sep_l}$ 最小. 总结以上过程有如下定理.

定理 3.4 把质量为 m_0 与 m_1 的小球分离开来所需要最少操作数的问题归成如下所述的整数规划问题：

$$
\begin{aligned}
& \min \frac{so_l}{k \times sep_l}, \\
& \text{s.t.} \\
& so_l = \sum_{i=1}^{l} o_i + 2 - \sum_{i=1}^{l} b_i - \sum_{i=2}^{l} e_{i1}; \\
& sep_l = \sum_{i=1}^{l} b_i + \sum_{i=2}^{l} e_{i1}; \\
& n = sk + r, \\
& 0 \leqslant r < k, n, s, k, r \in N; \\
& 2a_i = c_{i1} + c_{i2} + d_{i1} + d_{i2}, \\
& 2b_i = e_{i1} + e_{i2} + f_{i1} + f_{i2},
\end{aligned}
$$

$$i=1,2,\cdots,l;$$
$$S_i=B_i\cup UB_i,$$
$$B_i=B_{i1}\cup B_{i2},$$
$$i=1,2,\cdots,l;$$
$$UB_i=UB_{i1}\cup UB_{i2},$$
$$S_i\subset S_{i-1},$$
$$i=1,2,\cdots,l;$$
$$bp_i=a_{i-1}+b_{i-1}-e_{i1},$$
$$0=bp_{i-1}\bmod 2,$$
$$i=1,2,\cdots,l;$$
$$bp_i=a_{i-1}+b_{i-1}+1-e_{i1},$$
$$1=bp_{i-1}\bmod 2,$$
$$i=1,2,\cdots,l;$$
$$o_i=\frac{bp_i}{2},$$
$$i=1,2,\cdots,l;$$

其中,$B_i,UB_i,i=1,2,\cdots,l,S_i,B_{i1},B_{i2},UB_{i1},UB_{i2}$的含义如上分析所述.

可以看到,上述问题事实上是一个随机规划问题.这是因为,上述问题中 $B_i,UB_i,i=1,2,\cdots,l$ 的选取是随机的,从而 $S_i,B_{i1},B_{i2},UB_{i1},UB_{i2}$的选取也是随机的.但是解是确定的,我们所要求的是在一个分离过程中使操作数最少的分堆数以及操作数.

定理 3.5　上述规划问题有最优解 s^*.

证明:由定理 3.1,$s>m$.又根据上述的分离过程,我们总能够分离出来不含有质量为 m_1 的小球的球堆.又由该整数规划问题的中小球的有限性,问题一定存在最优值,从而一定存在最优解.　□

对球完成一次分离的分堆数 s 即为上述规划问题的最优解,而其最优目标值即为完成一次分离所需要的最小操作数.当完成一次分离之后,问题又归结成上述问题,故上述问题是一个嵌套问题.若记第 j 个分离过程的初始球数为 n_j,因为上述问题其他操作相同,是一个迭代过程,故其最优目标值为第 j 个分离过程的初始球数的函数,不妨记此函数为 f,则有第一个分离过程的最少操作数为 $f(n)$,第 j 个分离过程的最优解 $f(n_j)$.不妨设总共经过 t 个分离过程完成这 n 个球的两类球的分离.这样,为完成两类球的分离,问题便归结成下述的规划问题:

$$\min\sum_{j=1}^{t}f(n_j),$$

s.t.

$$n_{j+1}=mk_j+r_j,$$
$$j=1,2,\cdots,t;$$
$$n_1=n,\ k_t=1,\ r_t=0;$$

其中,$f(n_j)$由以下规划问题解得:

$$\min f(n_j)=\frac{so_l}{k\times sep_l},$$

s.t.

$$so_l=\sum_{i=1}^{l}o_i+2-\sum_{i=1}^{l}b_i-\sum_{i=2}^{l}e_{i1};$$
$$sep_l=\sum_{i=1}^{l}b_i+\sum_{i=2}^{l}e_{i1};$$
$$n_j=sk_j+r_j,$$
$$0\leqslant r_j<s,\quad n_j,s,k_j,r_j\in N;$$
$$2a_i=c_{i1}+c_{i2}+d_{i1}+d_{i2},$$
$$2b_i=e_{i1}+e_{i2}+f_{i1}+f_{i2},\quad i=1,2,\cdots,l;$$
$$S_i=B_i\cup UB_i,$$
$$B_i=B_{i1}\cup B_{i2},\quad i=1,2,\cdots,l;$$
$$UB_i=UB_{i1}\cup UB_{i2},$$
$$S_i\subset S_{i-1},\quad i=1,2,\cdots,l;$$
$$bp_i=a_{i-1}+b_{i-1}-e_{i1},$$
$$0=bp_{i-1}\bmod 2,\quad i=1,2,\cdots,l;$$
$$bp_i=a_{i-1}+b_{i-1}+1-e_{i1},$$
$$1=bp_{i-1}\bmod 2,\quad i=1,2,\cdots,l;$$
$$o_i=\frac{bp_i}{2},\quad i=1,2,\cdots,l.$$

4　$\theta>1$ 的情况

当 $\theta>1$ 时,问题可以以类似于 $\theta=1$ 的情况解决.假设这 $\theta+1$ 类球的质量和个数分别为 $m_0,m_1,\cdots,m_\theta$ 及 $n-m,s_1,\cdots,s_\theta$,并设 $m_0<m_1<\cdots<m_\theta$.我们的想法是一类球一类球地分离,先分离质量最大的一类球,待分离完之后,再分离质量次大的一类球,依次进行下去,直至把这 $\theta+1$ 类球全部分离出来.这是由天平的性质决定的.天平只能判别出天平两端哪端重,这样,就不可能从分离质量介于 m_0 与 m_θ 之间的球类.从而只能从质量最大或者最小的球类分离开始.又由假设 $m\ll n$,从分离质量最大的球开始方使分离的操作数最小.

为分离质量最大的一类球(在这里,质量最大的一类球是第 $\theta+1$ 类球),我们仍然要对这些球分堆.仿照定理 3.1,3.2,可以证明,在一个分离过程中,至少需要分成 $s>s_\theta+1$ 堆球才能分离出质量为 m_θ 的球;为使一个分离过程操作最少以及能够判别出哪堆球含有质量为 m_θ 的球,各堆球需具有相同个

数.同时仿照第3节对问题建立模型的过程,我们也可以建立这种情况下问题的整数规划模型.可以看到,球类的增加不改变模型的实质.这里就不重复了.

对一些特例的计算结果如表1所示.

表1　利用定理3.4计算一些特例的结果

Tab.1　The result of solving some special cases by Theorem 3.4

n	m	s	一次分离过程分离出来的堆数	总操作数
18	1	3	2	3
15	1	3	2	3
101	1	3	2	5
579	1	3	2	6
10001	1	3	2	9
8	2	4	2	5
18	2	4	2	8

5　注　记

分而治之的算法对于一些排序问题、最大最小问题很有效果,但是这种算法并不能给出解决这样一些问题到底最少需要多少操作数的问题.本文对这样一些问题构建了数学模型,并进行解决.如上述,在确定了最优的分堆数以后,问题的操作过程即已经确定.操作的过程事实上就只是简单地随机选取两堆比较称量了.这样就基本上解决了在第1节中提出的问题.

从以上的分析中可以看到,在本文中所要求的把质量不相同的不同类球分离开来的最小操作数事实上是在分堆时球分布的最坏可能的情况下所要的最小操作数.我们在第3节中建立的模型是一个随机规划模型,如果把分堆后不同类球的分布看作是一组随机变量,则操作数在分堆数确定后也是关于这一组随机变量的随机变量,因而存在一个数学期望值.这就意味着我们在把这$\theta+1$类质量不同的球分离开来时有一个平均的最小操作数,即在分堆与分离操作时各种可能的情况下操作数的最优的数学期望.那么,这个期望值如何求、是多少呢?这个问题仍然没有解决.

参考文献:

[1]　Sartaj Sahni.数据结构算法与应用——C++语言描述[M].北京:机械工业出版社,1999.434-466.

[2]　曹汝成.组合数学[M].广州:华南理工大学出版社,2000.170-176.

[3]　王晓东.计算机算法设计与分析[M].北京:电子工业出版社,2001.6-32.