ON REFINING PARTITIONS

P. ERDŐS, RICHARD K. GUY AND J. W. MOON

1. *Introduction*

If P is a partition of a set into disjoint subsets, we *refine* P by splitting one of the subsets into two smaller subsets. Let $f(n)$ denote the number of ways of transforming a set of n indistinguishable objects into n singletons via a sequence of $n-1$ refinements. Our main object is to establish the following bounds for $f(n)$.

THEOREM 1. *There exist constants* c_1 *and* c_2 *such that*

$$c_1{}^n \, n^{\frac{1}{2}n} < f(n) < c_2{}^n \, n^{\frac{1}{2}n}.$$

We also solve exactly the corresponding problem in which the n objects are distinguished by labels. The problem to determine $f(n)$ is due to B. J. T. Morgan.

2. *Numerical Calculations*

Figure 1 illustrates the problem for $n = 7$. The numbers by the partitions are the numbers of distinct paths from the original set of 7, so that $f(7) = 33$. We have calculated the following values of $f(n)$.

n	1	2	3	4	5	6	7	8	9	10
$f(n)$	1	1	1	2	4	11	33	116	435	1832

n	11	12	13	14	15	16
$f(n)$	8167	39700	201785	1099449	6237505	37406458

3. *A lower bound*

To obtain a lower bound we count only those sequences of refinements which include the partition $1.2.3 \ldots d.r$ of n into d or $d+1$ parts, d of which are of different size, where $0 \leqslant r = n - \frac{1}{2}d(d+1) \leqslant d$, so that $\sqrt{(2n)} > d > \sqrt{(2n)} - \frac{3}{2}$. Moreover we only count sequences in which we split off 1 from each of the $d-1$ parts of different size greater than 1. These $d-1$ steps can be made in $(d-1)!$ ways and result in the partition $1^{d+1}\, 2.3 \ldots (d-1).r$ of n into $2d-1$ or $2d$ parts, $d-1$ (or possibly d) of which are of different size. We deal with this in the same way, making splits of size 1 from each of the $d-2$ parts of different size greater than 1, in $(d-2)!$ possible ways. If we continue, we see that the number of sequences of refinements is at least

$$\prod_{m=1}^{d-1} m!,$$

Received 18 July, 1973.

[J. LONDON MATH. SOC. (2), 9 (1975), 565–570]

whose logarithm, by Stirling's formula, is

$$\sum_{2}^{d-1} (m \ln m - m) + O(d \ln d) = \int_{1}^{d-1} x(\ln x - 1)dx + O(d \ln d)$$

$$= [\tfrac{1}{2}x^2 \ln x - \tfrac{3}{4}x^2]_1^{d-1} + O(d \ln d)$$

$$= \tfrac{1}{2}(d-1)^2 \ln (d-1) - \tfrac{3}{4}(d-1)^2 + O(d \ln d)$$

$$= \tfrac{1}{2} . 2n . \tfrac{1}{2} \ln (2n) - \tfrac{3}{2}n + O(\sqrt{(n)} \ln n)$$

$$= \tfrac{1}{2}n \ln (2n/e^3) + O(\sqrt{(n)} \ln n),$$

so that, for sufficiently large n,

$$f(n) > c_1{}^n n^{\frac{1}{2}n}.$$

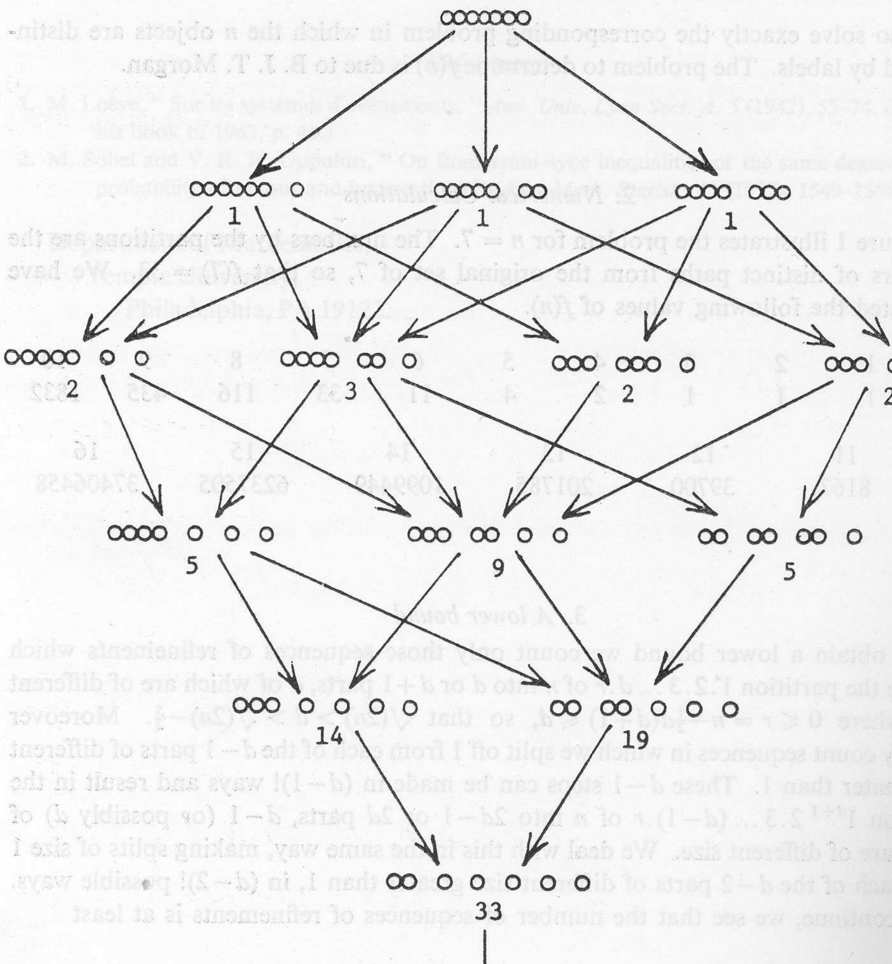

Fig. 1

where $c_1 \geqslant (2/e^3)^{\frac{1}{2}} > 0{\cdot}31555$. Our numerical results suggest that $\{f(n)\}^{1/n}/n^{\frac{1}{2}}$ increases steadily after a minimum at $n = 3$; so probably c_1 is $1/\sqrt{(3)} > 0{\cdot}57735$, and, for $n > 16$, may be replaced by $0{\cdot}75$.

4. *An upper bound*

Each refinement partitions a set of cardinality $s+t$ into 2 sets of cardinalities s and t, where we may assume $s \leqslant t$; such a refinement will be called a *split of size* s. We show that in a sequence of $n-1$ refinements, at most

$$[n/s]-1 \tag{1}$$

of them are splits of size s or larger. For, imagine the n members of the original set to be ordered from left to right, and adopt the convention that in any split, the smaller set will be the one on the left. At each split of size s or larger, we mark the s members immediately to the left of the split. We will never mark a member twice, so, after u such splits, regardless of what smaller splits are made, there will be exactly us marked members and at least s unmarked on the right; so $us+s \leqslant n$ and the result follows.

We say that a sequence of $n-1$ refinements is of *type* $(s_1, s_2, \ldots, s_{n-1})$ if the ith refinement is a split of size s_i. If there are parts of d different sizes in a partition, then $1+2+\ldots+d \leqslant n$ and $d < \sqrt{(2n)}$. It follows that if there are S sequences of $n-1$ refinements of a given type, then

$$S < (\sqrt{(2n)})^{n-1} < (2n)^{\frac{1}{2}n}, \tag{2}$$

since at each stage in the refinement we have to choose one part to split, and we have fewer than $\sqrt{(2n)}$ distinguishably different parts to choose from.

We say that a type $(s_1, s_2, \ldots, s_{n-1})$ has *pattern* $(n_0, n_1, \ldots)$ if exactly n_j of the numbers s_i satisfy the inequality

$$2^j \leqslant s_i < 2^{j+1} \tag{3}$$

for $j = 0, 1, \ldots$. Note that $\sum_j n_j = n-1$, and that by (1),

$$n_k \leqslant \sum_{j \geqslant k} n_j \leqslant \max([n/2^k]-1, 0) = m_k, \tag{4}$$

say, for $k = 0, 1, \ldots$; in particular $n_k = m_k = 0$ for $k > \log_2 n - 1$, i.e. for $k > l = [\log_2 n]-1$.

For a fixed j there are just 2^j distinct possible integer values that s_i may take in the interval (3). If there are P types $(s_1, s_2, \ldots, s_{n-1})$ with a given pattern $(n_0, n_1, \ldots, n_l)$, then

$$P \leqslant \binom{n-1}{n_0, n_1, \ldots, n_l} 2^{0n_0+1n_1+\ldots+ln_l}. \tag{5}$$

The exponent satisfies

$$\sum_{j=0}^{l} jn_j = \sum_{k=1}^{l} \sum_{j=k}^{l} n_j \leqslant \sum_{k=1}^{l} ([n/2^k]-1) < n,$$

by (4); the multinomial coefficient

$$\binom{n-1}{n_0, n_1, \ldots, n_l} = \binom{n-1}{n_0}\binom{n-1-n_0}{n_1}\binom{n-1-n_0-n_1}{n_2}\cdots\binom{n-1-n_0-\ldots-n_{l-1}}{n_l}$$

is at most

$$\binom{m_0}{n_0}\binom{m_1}{n_1}\binom{m_2}{n_2}\cdots\binom{m_l}{n_l},$$

also by (4), since

$$n-1-\sum_{j=0}^{k-1} n_j = \sum_{j=k}^{l} n_j \leqslant m_k.$$

So

$$P < \binom{m_0}{n_0}\binom{m_1}{n_1}\cdots\binom{m_l}{n_l} 2^n. \tag{6}$$

The total number of sequences of refinements is equal to the number of sequences of a given type, summed over all possible types, so

$$f(n) < \sum S < (2n)^{\frac{1}{2}n} T$$

by (2), where T is the number of types. In turn, T is the number of types of a given pattern, summed over all patterns $(n_0, n_1, \ldots, n_l)$;

$$T = \sum P < 2^n \sum \binom{m_0}{n_0}\binom{m_1}{n_1}\cdots\binom{m_l}{n_l}$$

by (6). Now each term of the sum appears in the product

$$\left(\binom{m_0}{0}+\ldots+\binom{m_0}{m_0}\right)\left(\binom{m_1}{0}+\ldots+\binom{m_1}{m_1}\right)\cdots\left(\binom{m_l}{0}+\ldots+\binom{m_l}{m_l}\right);$$

so $T < 2^{n+m_0+m_1+\ldots+m_l} < 2^{3n}$ and $f(n) < 2^{7n/2} n^{n/2}$, i.e.

$$f(n) < c_2{}^n n^{\frac{1}{2}n},$$

where $c_2 = 8\sqrt{2} < 11{\cdot}31371$.

5. *A generalization*

Let $1^g 2^h 3^i \ldots$ denote a partition of $n = g+2h+3i+\ldots$ into g parts of size 1, h parts of size 2, etc. and let $f(1^g 2^h 3^i \ldots)$ denote the number of sequences of $n-g-h-i-\ldots$ refinements of the partition into n singletons. Clearly $f(1^g 2^h \ldots) = f(2^h \ldots)$ and it is easy to write a general recurrence relation for $f(1^g 2^h \ldots)$; this was the basis of one of our methods of calculation. We can only solve the relation in simple cases.

THEOREM 2. *If h, i are non-negative*

$$f(2^h 3^i) = \binom{h+2i}{i} - \binom{h+2i}{i-1},$$

$$f(2^h 3^i 4^1) = (i+1)\left\{\binom{h+2i+4}{i+2} - \binom{h+2i+4}{i+1}\right\},$$

$$f(2^h 4^2) = \tfrac{1}{12}(h+1)(h^3+19h^2+118h+228),$$

$$f(2^h 3^1 4^2) = \tfrac{1}{24}(h+1)(h+6)(h^3+26h^2+225h+636),$$

$$f(2^h 5^1) = \tfrac{1}{6}(h+1)(h+3)(h+8),$$

$$f(2^h 3^1 5^1) = \tfrac{1}{24}(h+1)(3h^3+59h^2+358h+648),$$

$$f(2^h 3^2 5^1) = \tfrac{1}{40}(h+1)(2h^4+63h^3+717h^2+3458h+5800),$$

$$f(2^h 4^1 5^1) = \tfrac{1}{24}(h+1)(h^4+30h^3+323h^2+1458h+2376),$$

$$f(2^h 6^1) = \tfrac{1}{24}(h+1)(h+3)(h^2+22h+88),$$

$$f(2^h 3^1 6^1) = \tfrac{1}{30}(h+1)(h^4+34h^3+386h^2+1784h+2910),$$

$$f(2^h 7^1) = \tfrac{1}{120}(h+1)(h^4+49h^3+606h^2+2764h+3960).$$

6. *The labelled case*

Let $g(n)$ denote the number of ways of transforming a set of distinguished objects into n singletons via a sequence of $n-1$ refinements.

THEOREM 3. $g(n) = n!(n-1)!/2^{n-1}$.

Proof. If we classify the sequences according to the size of the subsets in the first refinement, we find that

$$g(n) = \tfrac{1}{2}\sum_{r=1}^{n-1}\binom{n}{r}\binom{n-2}{r-1} g(r)g(n-r)\,;$$

for, there are

$$\binom{n}{r} = \tfrac{1}{2}\left(\binom{n}{r}+\binom{n}{n-r}\right)$$

ways to split the original set into subsets of sizes r and $n-r$ and there are $\binom{n-2}{r-1}$ orders in which the remaining $n-2$ refinements may be carried out, $r-1$ on the first part, and $n-r-1$ on the second. The formula for $g(n)$ follows by induction.

The following derivation is even easier. We *consolidate* a partition by replacing any two parts by their union; then $g(n)$ is the number of ways of transforming a collection of n distinguished singletons into one set via a sequence of $n-1$ consolidations. There are $\binom{n}{2}$ ways of performing the first consolidation, so

$$g(n) = \binom{n}{2} g(n-1)$$

and the result follows by iteration or induction.

More generally, if P denotes a partition of n distinguished objects into k subsets of sizes $s_1, \ldots, s_k$, then it is not difficult to show that there are

$$(n-k)! \prod_{j=1}^{k} s_j!/2^{s_j-1}$$

ways to transform P into n singletons via a sequence of $n-k$ refinements.

The lattice of partitions is the subject of a monograph by Kreweras [1], but he does not give the results of the present paper. We are indebted to E. C. Milner and

the referee for clarifying discussions and remarks. Preparation of the paper was assisted by grants from the National Research Council of Canada.

Reference

1. Germain Kreweras, " Sur une classe de problèmes de dénombrement liés au treillis des partitions des entiers ", Cahiers du Bureau Universitaire de Recherche Opérationnelle # 6 (Paris, 1965).

Mathematical Institute,
University of Oxford.

The University of Calgary,
Alberta, Canada.

University of Alberta,
Edmonton.