JOURNAL OF COMBINATORIAL THEORY, Series A 70, 323–333 (1995)

Note

Enumerating Nested and Consecutive Partitions

F. K. HWANG AND C. L. MALLOWS

AT&T Bell Laboratories, 600 Mountain Avenue, P.O. Box 636, Murray Hill, New Jersey 07974

Communicated by the Managing Editors

Received March 4, 1993

1. INTRODUCTION

We consider partitions of an ordered set of n objects, which we take to be $N_n = \{1, 2, ..., n\}$. Using the principle of inclusion and exclusion, the number $P(n, p)$ of different p-part partitions is

$$P(n, p) = \frac{1}{p!} \sum_{k=0}^{p-1} (-1)^k \binom{p}{k} (p-k)^n. \tag{1}$$

Even for $p = 2$, this number $(2^{n-1} - 1)$ is exponential in n. Thus it is very time-consuming, in searching for an optimal partition under some cost function, to examine all these partitions. One way to deal with this "size" problem is to confine attention to a small subset of partitions. When the elements can be linearly ordered, an approach that is popular in the operations research literature [1, 3–6] is to work with cost functions such that the optimal partition will be consecutive, where a *consecutive* partition is one where every subset of the partition consists of consecutive elements. The number of consecutive partitions is

$$C(n, p) = \binom{n-1}{p-1}, \tag{2}$$

the number of ways of inserting $p - 1$ commas in the $n - 1$ spaces between adjacent elements. Thus even a brute force search needs only to examine a polynomial number of such partitions. Dynamic programming has been proposed to further cut down the computation.

Other subsets of partitions have been studied. Consider order-consecutive partition sequences, where an ordered partition $(S_1, ..., S_p)$ of N_n is *order-consecutive* iff for $k = 1, ..., p$, $\bigcup_{i=1}^{k} S_i$ is a consecutive subset of N_n.

0097-3165/95 $6.00

Clearly a consecutive partition gives an order-consecutive partition sequence. Such an ordered partition can be represented by a completely nested set of pairs of parentheses, i.e., one with all the left parentheses occurring before the first right parenthesis, as for example in

$$(1((2((3)\,45)\,6)\,78)\,9) \tag{3}$$

which represents the order-consecutive partition sequence

$$S_1=\{3\}, s_2=\{4, 5\}, S_3=\{2, 6\}, S_4=\{7, 8\}, S_5=\{1, 9\}.$$

Chakravarty *et al.* [4] gave conditions under which there exists an order-consecutive optimal partition, where *order-consecutive* (they called it *semi-consecutive*) means that the subsets of the partition can be ordered so that it has the order-consecutive property. For example, suppose $n=5$, $p=3$. Then $(S_1=\{2, 3\},\ S_2=\{4\},\ S_3=\{1, 5\})$ and (S_2, S_1, S_3) are both order-consecutive partition sequences; so the unordered set $\{S_1, S_2, S_3\}$ is order-consecutive. None of these partitions is consecutive. Clearly a consecutive partition is order-consecutive.

Boros and Hammer [2] gave conditions under which there exists a nested optimal partition, where *nested* means that there do not exist four elements $a<b<c<d$ with a and c in one subset and b and d in another. Clearly, an order-consecutive partition is nested since a partition that is not nested cannot have the order-consecutive property. An example of a nested partition that is not order-consecutive is $\{S_1=\{2\},\ S_2=\{4\},\ S_3=\{1, 3, 5\}\}$.

Let the numbers of order-consecutive partition sequences, order-consecutive partitions, and nested partitions of N_n, each with p parts, be $\mathrm{OCPS}(n, p)$, $\mathrm{OCP}(n, p)$, and $N(n, p)$, respectively. Clearly

$$C(n, p) \leqslant \mathrm{OCP}(n, p) \leqslant \mathrm{OCPS}(n, p), \qquad \mathrm{OCP}(n, p) \leqslant N(n, p).$$

We shall determine $\mathrm{OCP}(n, p)$, $\mathrm{OCPS}(n, p)$, and $N(n, p)$. Some numerical values appear at the end of this paper.

2. The Number of Nested Partitions

Suppose $\{S_1, S_2, ..., S_p\}$ is a nested partition of N_n. Since the partition is nested, we can represent it by placing pairs of parentheses suitably around and between the ordered elements of N_n; e.g., with $n=9$, $p=4$,

$$(1(23)\,45(6)(7)\,8)(9) \tag{4}$$

represents the nested (but not order-consecutive) partition

$$S_1=\{1, 4, 5, 8\}, S_2=\{2, 3\}, S_3=\{6\}, S_4=\{7\}, S_5=\{9\}. \tag{5}$$

Note that we place the parentheses in a pair as close together as possible; thus we do not allow the representation

$$(1(23)\,45((6)\,7)\,8)(9)$$

because here the parentheses defining the subset $\{7\}$ are not as close as possible. Note that in a representation such as (4) of a nested partition we cannot have two adjacent left parentheses or two adjacent right ones; if this happens, the outer parenthesis can be moved to be closer to its mate. The only way two adjacent parentheses can occur (between two integers) is as a ')('-pair. We call such a configuration "N-proper." An N-proper representation of a nested partition is clearly unique. Note that if we remove all)(-pairs from such a representation, the resulting configuration of parentheses will continue to satisfy the usual constraint that the number of left parentheses, counting from left to right, is never less than the number of right parentheses.

It is well known that the number of ways k pairs of parentheses can be arranged, satisfying the usual constraint that the number of left parentheses, counting from left to right, is never less than the number of right parentheses, is the Catalan number

$$C_k = \frac{1}{k+1}\binom{2k}{k}.$$

We need a simple lemma.

LEMMA 1. *Let the number of ways k pairs of parentheses can be arranged, subject to the usual condition and such that the mate of the first (left) parenthesis is the last (right) parenthesis, be* C'_k. *Then* $C'_k = C_{k-1}$.

Proof. We have the generating function

$$\begin{aligned} C(x) &= \sum_{k=0}^{\infty} C_k x^k \\ &= \frac{1}{2x}\,(1-\sqrt{1-4x}). \end{aligned}$$

Enumerating the partitions that are counted by C_k according to the position of the mate of the first (left) parenthesis shows that

$$\sum_{k=0}^{\infty} C'_k(x) = C(x)^2 = (C(x)-1)/x,$$

which proves the lemma. ∎

THEOREM 2. *The number of nested partitions of N_n with p parts and having j)(-pairs is*

$$\frac{n-1!}{j!\,(p-j)!\,(p-j-1)!\,(n-2p+j+1)!}. \tag{6}$$

Proof. We show that the number we require is

$$\binom{n-1}{2p-j-2}\binom{2p-j-2}{j} C_{p-j-1} \tag{7}$$

which reduces to (6). We have to place p pairs of parentheses in and around N_n so as to define a proper nested partition, and with exactly j)(-pairs occurring. Suppose we remove the j)(-pairs. Then it is necessary that the mate of the first (left) parenthesis is the last (right) parenthesis, since otherwise some element of N_n would not be included in any part of the partition. The)(-pairs can occur anywhere between these extreme parentheses. We have $2p-2j$ single parentheses, which according to Lemma 1 can be arranged in C_{p-j-1} ways. There are $2p-2j-2$ parentheses between the extreme pair; we can insert the j)(-pairs in $\binom{2p-j-2}{j}$ ways. Now we have $2p-j-1$ gaps between the single parentheses and the)(-pairs, each of which must be assigned at least one element of N_n. We can do this by first placing one element in each gap and then permuting the remaining $n-(2p-j-1)$ elements with $2p-j-2$ separators, in $\binom{n-1}{2p-j-2}$ ways. This proves (7). ∎

COROLLARY 3.

$$N(n, p) = \frac{1}{n}\binom{n}{p}\binom{n}{p-1}. \tag{8}$$

Proof. Using (6), a simple (Vandermonde) summation gives

$$\sum_{j=0}^{p-1} \frac{n-1!}{j!\,(p-j)!\,(p-j-1)!\,(n-2p+j+1)!}$$
$$= \frac{1}{n}\binom{n}{p} \sum_{j=0}^{j=p-1} \binom{p}{j}\binom{n-p}{p-j-1} = (8). \quad ∎$$

COROLLARY 4. *The total number of nested partitions of N_n is*

$$\sum_{p=1}^{n} N(n, p) = C_n. \tag{9}$$

Proof. Vandermonde is used again.

3. The Number of Order-Consecutive Partitions

Order-consecutive partition sequences can also be represented by placing parentheses around and between the elements of N_n. For example, with $n=9$, $p=4$,

$$(1(2(3)(45)\,6)(78)\,9)$$

represents the partition

$$\{S_1=\{1,9\}, S_2=\{2,6\}, S_3=\{3\}, S_4=\{4,5\}, S_5=\{7,8\}\},$$

which is order-consecutive since the ordered partition

$$(S_3, S_4, S_2, S_5, S_1)$$

is order-consecutive. (Note that S_3 and S_4 could be taken in the reverse order.) We call such a representation, with all pairs of parentheses as close together as possible, OCP-proper. There is now an additional constraint.

LEMMA 5. *If all)(-pairs in an OCP-proper configuration are deleted, then the remaining pairs of parentheses are completely nested.*

Proof. Suppose to the contrary that we have a right parenthesis $)_1$ to the left of a left parenthesis $(_2$ not as a)(-pair, so that there is a non-empty set x of integers between them. Without loss of generality, we may assume that no other parentheses, except possibly some that form)(-pairs, lie between $)_1$ and $(_2$, since otherwise we could replace $\{)_1, (_2\}$ by a closer pair of parentheses. If there exists no)(-pair between $)_1$ and $(_2$, consider

TABLE I

Values of OCP(n, p), the Number of p-Part Order-Consecutive Partitions of N_n

$n \backslash p$	1	2	3	4	5	6	7	8
1	1							
2	1	1						
3	1	3	1					
4	1	6	6	1				
5	1	10	19	10	1			
6	1	15	45	45	15	1		
7	1	21	90	141	90	21	1	
8	1	28	161	357	357	161	28	1

TABLE II

Values of $N(n, p)$, the Number of p-Part Nested Partitions of N_n

$n \backslash p$	1	2	3	4	5	6	7	8
1	1							
2	1	1						
3	1	3	1					
4	1	6	6	1				
5	1	10	20	10	1			
6	1	15	50	50	15	1		
7	1	21	105	175	105	21	1	
8	1	28	196	490	490	196	28	1

the left parenthesis $(_1$ that is the mate of $)_1$ and the right parenthesis $)_2$ that is the mate of $(_2$. Then the configuration must be

$$\cdots (_1\, w\,)_1\, x\, (_2\, y\,)_2 \cdots,$$

where each of w, x, y is non-empty (and might contain more parentheses). Now consider the pair of parentheses $(_0$, $)_0$ that define the subset containing x. These lie beyond $(_1$ and $)_2$, so we must have

$$\cdots (_0\, v\, (_1\, w\,)_1\, x\, (_2\, y\,)_2\, z\,)_0 \cdots,$$

with v and z non-empty. But this configuration does not define a order-consecutive partition.

If there are one or more)(-pairs between $)_1$ and $(_2$, let $)'$ be the left-most right parenthesis of the)(-pairs, and let $('' $ be the right-most left parenthesis of the)(-pairs. Let $(_1$ and $('$ be the mates of $)_1$ and $)'$, respectively, defining subsets S_1 and S' respectively. Similarly, let $)_2$ and $)''$ be the mates of $(_2$ and $(''$, defining subsets S_2 and S''. Then we have

$$('\, S'\, (_1\, S_1\,)_1\, S'\,)'\, S_0\, (''\, S''\, (_2\, S_2\,)_2\, S''\,)''$$

where S_0 may be empty. Since there is no left parenthesis between $)_1$ and $)'$, $('$ must be to the left of $)_1$, and because of the order-consecutive property is also to the left of $(_1$. This implies that S' has a non-zero number of elements between $('$ and $(_1$, and also between $)_1$ and $)'$. Therefore, S' must appear after S_1 in any ordering of the subsets of the partition that has the order-consecutive property. Similarly, we conclude that S'' must appear after S_2. But S_1 is not consecutive to either S_2 or S'', and neither is S_2 consecutive to S_1 or S'. Hence there is no ordering of the four subsets S_1, S_2, S', S'' that preserves the order-consecutive property, in contradiction to our assumption. ∎

THEOREM 6.

$$\mathrm{OCP}(n, p)=\sum_{j=0}^{p-1}\binom{n-1}{2p-j-2}\binom{2p-j-2}{j}. \tag{10}$$

Proof. We count the number of OCP-proper configurations. There must be a single left parenthesis in the space before the first element and a single right parenthesis after the last element. Suppose the remaining parentheses contain j)(-pairs. Then there are $2p-2j-2$ parentheses not involved in)(-pairs. Each of these $j+(2p-2j-2)$ objects (i.e., single parentheses and)(-pairs) must fill a different space chosen from the $n-1$ spaces between the elements. The first factor in (9) counts the number of ways these spaces can be chosen. The second factor represents the number of ways the j)(-pairs can be inserted into a sequence of $p-j-1$ left parentheses followed by $p-j-1$ right parentheses. ▮

4. THE NUMBER OF ORDER-CONSECUTIVE PARTITION SEQUENCES

Suppose $S=\{S_1, ..., S_p\}$ is an order-consecutive partition sequence of N_n. We can represent S (uniquely) by inserting $2p-2$ characters, alternately $p-1$ commas and slashes, (i.e., ,/,/ $\cdots$,/ in that order) into the n spaces between the elements of N_n (the space after the last element is allowed to contain a slash), subject to the constraint that if we ignore the slashes, the commas divide N_n into a proper p-part partition, i.e., there must be at least one element of N_n between each pair of commas. In this representation, the slashes indicate how each successive part of the partition relates to the previously accumulated parts: elements between the jth comma and the jth slash correspond to elements of S_{j+1} that lie to the left of $\bigcup_{k=1}^{j} S_k$, and elements between the jth slash and the $(j+1)$th comma correspond to elements of S_{j+1} that lie to the right of this union. For example, the order-consecutive sequence in (3) above is here represented as

1,/23,4/5,/67,8/9

which shows that S_1 has just one element; S_2 has two elements, both of which lie to the right of S_1; S_3 has two elements, one of which lies to the left of $S_1 \cup S_2$ and the other to the right; S_4 has two elements, both to the right of $S_1 \cup S_2 \cup S_3$; and S_5 has two elements, one to the left and one to the right of $\bigcup_{k=1}^{4} S_k$.

THEOREM 7.

$$\mathrm{OCPS}(n, p)=\sum_{k=0}^{p-1}(-1)^{p-1-k}\binom{p-1}{k}\binom{n+2k-1}{2k}. \tag{11}$$

TABLE III

Values of OCPS(n, p), the Number of p-Part Order-Consecutive Partition Sequences of N_n

$n \backslash p$	1	2	3	4	5	6	7	8
1	1							
2	1	2						
3	1	5	4					
4	1	9	16	8				
5	1	14	41	44	16			
6	1	20	85	146	112	32		
7	1	27	155	377	456	272	64	
8	1	35	259	833	1408	1312	640	128

Proof. We count the number of representations of the form described above. If we were to ignore the requirement that the commas define a proper (consecutive) p-part partition, the number of ways of inserting the commas and slashes would be

$$\binom{n+2p-3}{2p-2}.$$

But this counts many arrangements with parts of size zero. Using inclusion–exclusion, we first subtract the number of arrangements in which for some j, $0 \leqslant j \leqslant p-1$, the jth part is empty; in these arrangements the jth comma is immediately followed by a slash and another comma. Deleting this slash and the second comma, we have one of the arrangements counted by

$$\binom{n+2p-5}{2p-4}.$$

Continuing, we arrive at (11) (with $k = p-1-j$). ∎

5. Some Related Topics

Let $N^*(n, p)$ denote the number of nested partitions such that the mate of the first left parenthesis is the last right parenthesis, i.e., both the first and the last element of N_n belong to the same subset. We call this the *-property.

Lemma 8.

$$N^*(n, p) = N(n-1, p).$$

Proof. We establish a one–one mapping between the members of the two sets that are enumerated by $N^*(n, p)$ and $N(n-1, p)$. Given any (n, p) nested partition satisfying the * property, simply delete the element n; given any $(n-1, p)$ nested partition, ad the element n to the subset containing the first element. ▮

We now show that the number $N(n, p)$ arises in some other contexts.

THEOREM 9. *The number of ways n pairs of parentheses can be arranged (subject to the usual constraint) such that exactly $p-1$)(-pairs occur is $N(n, p)$.*

Proof. We will call the parentheses to be arranged in this theorem "brackets" to avoid confusion with the parenthesis that are counted by $N(n, p)$. We will define a one–one mapping between arrangements of n pairs of brackets containing exactly $p-1$][-pairs and nested partitions of N_n into p parts. Note that this mapping is not the same as the one discussed in the proof of Theorem 2. Some examples of the correspondence we shall set up, with $n=4$, $p=3$, are

(1)(2)(34) [ ][ ][[ ]]
(1)(23)(4) [ ][[ ]][ ]
(12)(3)(4) [[ ]][ ][ ]
(1)(2(3)4) [ ][[ ][ ]]
(1(2)3)(4) [[ ][ ]][ ]
(1(2)(3)4) [[ ][ ][ ]].

We proceed by induction on n. For $n=1$ (and $p=1$) simply transform the brackets into parentheses (and delete the integer 1). For general n, consider a p-part nested partition of N_n in its parenthesis representation. Suppose that the mate of the first parenthesis lies in the ith space, which is just to the right of the element i of N_n. If $1 \leqslant i \leqslant n-1$ (as in the first five examples above), then this ith space must contain a)(-pair and the partition can be decomposed into two subpartitions, one of N_i and the other of $\{i+1, ..., n\}$. Suppose that the first subpartition contains q pairs of parentheses (including the original pair). Then the second subpartition contains $p-q$ pairs of parentheses. By the inductive hypothesis, each of these subpartitions corresponds uniquely to an arrangement of brackets, the first having i pairs of brackets and $q-1$][-pairs, and the second having $n-i$ pairs of brackets and $p-q-1$][-pairs. Concatenating these two subpartitions introduces one more][-pair, giving altogether an arrangement with n pairs of brackets containing $p-1$][-pairs.

If $i=n$ (as in the last of the six examples above), the p-part partition of N_n has the *-property of Lemma 8, and we can use the one–one mapping in the proof of that lemma to replace the (n, p) partition by a nested $(n-1, p)$ partition, simply by deleting the last element of N_n. Now take the parenthesis representation of that partition, replace it (by the inductive hypothesis) by its corresponding bracket configuration, and place an extra pair of brackets around it. This gives an arrangement of $1+(n-1)=n$ pairs of brackets having $p-1$][-pairs, as required. All the steps in these constructions are reversible. ∎

The correspondence in Theorem 9 provides an alternative proof of (9).

The number $N(n, p)$ arises also in the following context: it is the number of weak-lead lattice paths from $(0, 0)$ to (n, n) that have exactly p horizontal (and p vertical) segments; i.e., it is the number of arrangements of n votes for each of two candidates A and B such that in the counting, A never trails B and the votes arrive in exactly $2p$ blocks, alternately for A and B. Such vote sequences are in 1–1 correspondence with $(n+1)$-node rooted trees, in two distinct ways: in the first, one circumnavigates the tree, going up a new branch for an A-vote and down the other side of an old branch for a B-vote. For the second correspondence, color the nodes of the tree black and white alternately. Circumnavigate the tree, starting at the (black) root, assigning labels $0, 1, 2, ..., n-1$ to the black nodes as they are encountered. (We do not assign a label n to the root node on completing the circuit.) Then, for each black node, each of its labels is replaced by a copy of its lowest numerical label. Now put the set of assigned labels into increasing order, giving $c_1, c_2, ..., c_n$. Then this represents (uniquely) a vote sequence in which c_i is the number of votes that B has obtained when A receives his ith vote. (See [7].)

Finally, we present an identity for which we have only an algebraic proof.

THEOREM 10.

$$N(n, p)=\sum_{j=1}^{p-1}\binom{n+j-1}{2p-2} N(p-1, j).$$

REFERENCES

1. S. ANILY AND A. FEDERGRUEN, Structured partitioning problems, *Oper. Res.* **39** (1991), 130–149.
2. E. BOROS AND P. L. HAMMER, On clustering problems with connected optima in Euclidean spaces, *Discuss. Math.* **75** (1989), 81–88.

3. A. K. Chakravarty, J. B. Orlin, and U. G. Rothblum, A partitioning problem with additive objective with an application to optimal inventory grouping for joint replenishment, *Oper. Res.* **30** (1982), 1018–1022.
4. A. K. Chakravarty, J. B. Orlin, and U. G. Rothblum, Consecutive optimizers for a partition problem with applications to optimal inventory groups for joint replenishment, *Oper. Res.* **33** (1985), 820–834.
5. F. K. Hwang, Optimal partitions, *J. Optim. Theory Appl.* **34** (1981), 1–10.
6. F. K. Hwang, J. Sun, and E. Y. Yao, Optimal set partitioning, *SIAM J. Algebraic Discrete Methods* **6** (1985), 163–170.
7. C. L. Mallows and K. W. Wachter, Valency enumeration of rooted plane trees, *J. Austral. Math. Soc.* **13** (1972), 472–476.