Enumerating Consecutive and Nested Partitions for Graphs

F. K. Hwang

AT&T Bell Laboratories
Murray Hill, New Jersey 07974

*G. J. Chang**

Department of Applied Mathematics
National Chiao Tung University
Hsinchu, 30050 Taiwan, ROC

ABSTRACT

We extend the study of consecutive and nested partitions on a set of integers to the vertex-set of a graph. A subset of vertices is considered consecutive if the subgraph induced by the subset is connected. In this sense the partition problem on a set of integers can be treated as a special case when the graph is a line. In this paper we give the number of consecutive and nested partitions when the graph is a cycle. We also give a partial order on general graphs with respect to these numbers.

*Supported partially by the National Science Council of the Republic of China under grant NSC81-0208-M009-26, and partially by DIMACS, Rutgers University.

References

[1] E. R. Barnes, A. J. Hoffman and U. G. Rothblum, "Optimal partitions having disjoint convex and conic hulls," *Math. Program*, **54** (1992) 69–86.

[2] F. K. Hwang and C. L. Mallows, "Enumerating nested and consecutive partitions," to appear.

$G_n = L_n$. If $T'_n \neq L_n$, do the same. Eventually we obtain L_n. By Theorem 12, the number of nested or order-consecutive partitions is nonincreasing through these transformations. Hence Theorem 18.

For $p = 2$ we can say a little more for T_n.

Theorem 19. $N_N(T_n, 2) = N_{OC}(T_n, 2) =$ *(the number of subtrees T_n) $- n + 1$.*

Proof: For $p = 2$, a partition is nested if and only if it is order-consecutive. While every nested partition must contain a subtree as a part, every subtree also induces a partition for which vertices in and out of the subtree constitute the two parts. However, each consecutive partition yields two subtrees, i.e., both subtrees induce the same partition. We correct this overcount by subtracting $N_C(T_n, 2) = n - 1$. ∎

The number $g(T_n)$ of subtrees in T_n can be counted algorithmically. Designate a node r as the root of T_n. Let r' be a node adjacent to r. View T_n as the union of two trees $T(r)$ with root r and $T(r')$ with root r', where r and r' are linked by an edge. Define $g_1(T, r_T)$ as the number of subtrees in T with the root r_T in the subtree, and $g_2(T, r_T)$ without. Then

$$\begin{aligned} g(T_n) &= g_1(T_n, r) + g_2(T_n, r), \\ g_1(T_n, r) &= g_1(T(r), r)[1 + g_1(T(r'), r')], \\ g_2(T_n, r) &= g_2(T(r), r) + g(T(r'), r'). \end{aligned}$$

These recursive equations can be solved in linear time if adding two numbers can be done in constant time.

Finally, we comment that by Theorems 14 and 16, any G_n with a vertex of degree $n-1$ has

$$N_N(G_n, p) = N_{OC}(G_n, p) = N(n, p) .$$

disjoint. By interchanging the parts i and j over all vertices in ℓ, the nonnestedness caused by i and j is eliminated (this interchange does not preserve the size of a part).

If there are more than two parts involved in nonnestedness, then these parts can be ordered according to their nested order in ℓ, namely, the span of a later part is not covered by the span of an earlier part. Reverse their ordering in ℓ_2 will result in a nested partition on G_n.

To summarize, for every nested partition P' which induces a nonnested partition P, we find a nested partition Q which is induced by a nonnested partition Q'. Thus the number of nested partitions of G_n is at least as numerous as that of G'_n.

A similar scheme works for order-consecutiveness. Suppose that parts are labeled according to their order in the order-consecutive sequence. It is easily verified that the only time P becomes not order-consecutive is when part 1 appears both at the end of ℓ_1 and the beginning of ℓ_2. Let v_1 has part i and the end vertex of ℓ_2 has part j. Then the ordering of parts in ℓ' is $(1, 2, \ldots, i, 1, 2, \ldots, j)$. Reverse the order of ℓ_1. Then the new order of ℓ' $(i, \ldots, 2, 1, 1, 2, \ldots, j)$ is order-consecutive. Since ℓ_1 and R can have at most one part in common, i.e., part i, the reverse operation preserves the order-consecutiveness in $\ell_1 \cup R$. ∎

We now give a lower bound.

Theorem 18. *$N_c(G_n, p) \geq N_c(L_n, p)$ for $c \in \{C, N, OC\}$.*

Proof: Since any G_n which is not a tree can be reduced to a tree by deleting edges,

$$N_C(G_n, p) \geq N_C(T_n, p) = N_C(L_n, p)$$

by Theorems 16 and 10. We also have

$$N_c(G_n, p) \geq N_c(T_n, p) \quad \text{for some } T_n, \quad c \in \{N, OC\}\ .$$

If $T_n \neq L_n$, then there exists a vertex v of degree at least three and with two linear branches. Combine the two linear branches to obtain a tree T'_n. Define $D(G_n)$ to be the sum of degree over all vertices with degree at least three. Then $D(T_n) > D(T'_n)$. Since $D(G_n) \geq 2n - 2$ for any graph G_n with n vertices, and the equality holds if and only if

Proof: Let P' denote a nested partition on G'_n. Suppose that ℓ consists of the sequence of vertices $v_1, \ldots, v_m$. Let P be a partition on G_n obtained from P' by distributing the parts assigned to ℓ to $\ell_1 \cup \ell_2$. We will actually give a one-to-one mapping between the vertices in ℓ and those in $\ell_1 \cup \ell_2$, with the understanding that the corresponding vertices are assigned the same part. Let R denote the rest of G_n, i.e., $R = G_n \backslash \ell = G'_n \backslash \{\ell_1 \cup \ell_2\}$. Then the mapping is shown in Fig. 1.

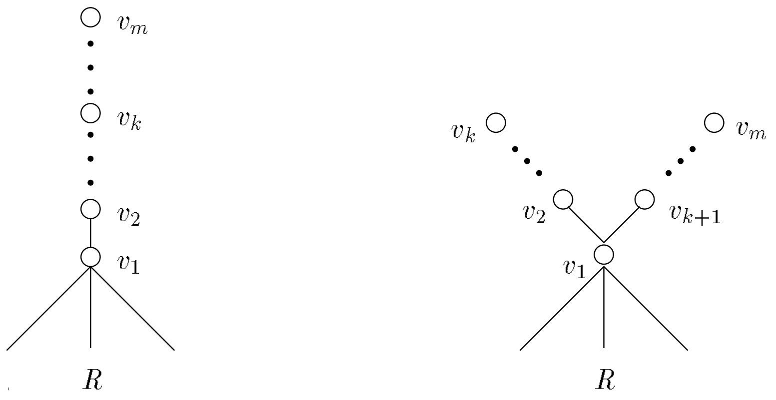

Figure 1: The Mapping between ℓ and $\ell_1 \cup \ell_2$.

Note that the ordering of vertices in $\ell_1 \cup R$ and $\ell_2 \cup R$ is the same as in $\ell \cup R$. Therefore, if P is not nested, there must exist four vertices, two belonging to part i, and two to part j, whose order in ℓ induces the order (i, j, j, i) of parts, but whose order in $\ell' = (v_k, v_{k-1}, \ldots, v_1, v_{k+1}, \ldots, v_m)$ induces the order (j, i, j, i). Counting v_1 as a vertices of ℓ_1, then two of the four vertices must be in ℓ_1 and two in ℓ_2 (or the ordering in ℓ and ℓ' would not be different). Furthermore, if v_1 is one of the four vertices (so v_1 belongs to part i), then R cannot have any vertex with part j; and if v_1 is not, then R cannot have a pair of vertices with parts (i, j). (It is easily verified that otherwise, P' would not be nested.) By interchanging the parts i and j of the two involved vertices in ℓ_2, the four vertices now have the nested order in ℓ' but not in ℓ.

After making the above interchange, ℓ' may still not be nested because there may exist other sets of four nonnested vertices. But an extension of the above scheme can handle that. First assume that each of the four involved vertices really represents a group of vertices all having the same part. Because of the nestedness of ℓ, the spans of these groups over ℓ are

Proof: Any cutting at $p-1$ edges not in the cycle or at p edges including at least two in the cycle induces a p-partition. There are $\binom{n-m}{p-1}$ of the first type and $\binom{n}{p} - \binom{n-m}{p} - m\binom{n-m}{p-1}$ of the second type. ∎

Corollary 13. $N_C\left(U_n(m)\right) = 2^n - m2^{n-m}$.

Let S_n denote a star with n vertices.

Theorem 14. $N_N(S_n, p) = N_{OC}(S_n, p) = N(n, p)$.

Proof: Any p-partition is an order-consecutive partition, hence a nested partition, by labeling the part containing the center of the star π_1. ∎

Corollary 15. $N_N(S_n) = N_{OC}(S_n) = \sum_{p=1}^{n} N(n, p)$.

We do not have explicit formulas for the nested and the order-consecutive class, which are likely to depend on some graph parameters. But we will give two partial orders on G_n and use them to obtain bounds.

Theorem 16. *If G'_n is obtained from G_n by adding an edge, then $N_c(G'_n, p) \geq N_c(G_n, p)$ for $c \in \{C, N, OC\}$.*

Proof: Clearly, any c partitions on G_n is also a c partition on G'_n. ∎

Theorem 17. *Let G_n be a graph which contains a vertex v_1 of degree at least three and with two linear branches, say, ℓ_1 and ℓ_2 (both including v_1). Let G'_n be a graph obtained from G_n by combining ℓ_1 and ℓ_2 into one linear branch, say, ℓ, at v_1. Then*

$$N_N(G_n, p) \geq N_N(G'_n, p)$$

and

$$N_{OC}(G_n, p) \geq N_{OC}(G'_n, p) \ .$$

Corollary 9. $N_{OC}(C_n) = 1 + \sum_{p=2}^{n} \frac{n}{p(p-1)} \sum_{i \geq 0} \binom{n-1}{p+i-1} \binom{p}{i} \binom{p-1}{i+1}$.

From Theorems 4, 6 and 8, $N_C(C_n, p)$, $N_N(C_n, p)$ and $N_{OC}(C_n, p)$ are all polynomial in n for fixed p, while without these restrictions, $N(C_n, p) = N(n, p)$ is exponential in n (the Bell number). Therefore, if we know that there exists an optimal consecutive (or nested, or order-consecutive) partition under certain objective function, then there exists an efficient algorithm to search for an optimal partition.

A partition problem involving C_n comes up in the following context. Barnes, Hoffman and Rothblum [1] considered a partition problem of points in a d-dimensional space. They showed that if for certain objective functions, there exist an optimal partition such that the conic hulls (issued from the origin) of the points in a part are all disjoint. For $d = 2$, the points (as vectors) can be cyclically ordered by the angles of the vectors. Thus there exists an optimal consecutive partitions of vertices on C_n.

4. Other graphs

Let T_n denote a tree with n vertices. Then

Theorem 10. $N_C(T_n, p) = \binom{n-1}{p-1}$.

Proof: There exists a one-to-one mapping between the set of p-partitions on T_n and the set of choices of $p-1$ edges.

Corollary 11. $N_C(T_n) = 2^{n-1}$.

Let $U_n(m)$ denote a connected graph with n vertices and a unique cycle of size m, i.e., U_n is a tree plus an additional edge.

Theorem 12. $N_C\left(U_n(m), p\right) = \binom{n}{p} - (m-1)\binom{n-m}{p-1} - \binom{n-m}{p}$ *for* $p \geq 2$.

Theorem 6. $N_N(C_n, p) = N_N(L_n, p) = \dfrac{1}{n}\dbinom{n}{p-1}\dbinom{n}{p}$.

Proof: There exists a one-to-one mapping between the set of nested p-partitions of L_n and the corresponding set of C_n (the mapping is by bending the line into a cycle). ∎

Corollary 7. $N_N(C_n) = N_N(L_n) = \dfrac{1}{n+1}\dbinom{2n}{n}$.

Theorem 8. $N_{OC}(C_n, p) = \dfrac{n}{p(p-1)}\sum_{i\geq 0}\dbinom{n-1}{p+i-1}\dbinom{p}{i}\dbinom{p-1}{i+1}$ *for* $p \geq 2$.

Proof: An order-consecutive p-partition can be obtained by first partitioning the vertices into $p+i$, $i \geq 0$, consecutive parts, and then choosing i pairs from the $p+i$ parts and combining each such pair into one final part. These pairs must satisfy the following conditions:

(i) The two parts in a pair are nonadjacent.

(ii) Consider any two pairs with parts (a, b) and (c, d). Then the relative positions of a, b, c, d on the cycle cannot be a, c, b, d or a, d, b, c (it doesn't matter clockwise or counter-clockwise).

The reason for condition (i) is we want to add up the p-partitions generated from an initial choice of $p+i$ parts over i. Combining two adjacent parts into a pair reduces an initial choice of $p+i$ parts to a choice of $p+i-1$ parts. The reason of condition (ii) is to preserve the order-consecutiveness.

Viewing the $p+i$ parts on the cycle as the vertices of a $(p+i)$-gon preserving their adjacency relation on the cycle, then the number of ways of choosing i pairs satisfying conditions (i) and (ii) is simply $f(p+i, i)$. Therefore

$$N_{OC}(C_n, p) = \sum_{i\geq 0}\binom{n}{p+i} f(p+i, i)$$

$$= \sum_{i\geq 0}\binom{n}{p+i}\binom{p}{i}\binom{p-2}{i}\frac{p+i}{(i+1)p}$$

$$= \frac{n}{p(p-1)}\sum_{i\geq 0}\binom{n-1}{p+i-1}\binom{p}{i}\binom{p-1}{i+1}\,. \qquad \blacksquare$$

Proof: While it is difficult to obtain a closed-form solution of $f(g,d)$ from the recursive equations of Lemma 2, once a solution is available, it can be inserted into the recursive equations for a straightforward verification. The formula is easily checked to be correct for $f(g,0)=1$ and $f(g,g-1)=0$. For $g-1>d\geq 1$,

$$\begin{aligned}
f(g,d) &= \frac{g}{g-2d}\, f(g-1,d)+\frac{g}{g-2}\, f(g-2,d-1)\\
&= \frac{g}{g-2d}\binom{g-d-1}{d}\binom{g-d-3}{d}\frac{g-1}{(d+1)(g-d-1)}\\
&\quad + \frac{g}{g-2}\binom{g-d-1}{d-1}\binom{g-d-3}{d-1}\frac{g-2}{d(g-d-1)}\\
&= \frac{g(g-2d-2)}{(g-d)(g-d-2)}\binom{g-d}{d}\binom{g-d-2}{d}\frac{g-1}{(d+1)(g-d-1)}\\
&\quad + \frac{gd}{(g-d)(g-d-2)}\binom{g-d}{d}\binom{g-d-2}{d}\frac{1}{g-d-1}\\
&= \frac{g}{(g-d)(g-d-1)(g-d-2)}\binom{g-d}{d}\binom{g-d-2}{d}\left[\frac{(g-2d-2)(g-1)}{d+1}+d\right]\\
&= \frac{g}{(g-d)(g-d-1)(g-d-2)}\binom{g-d}{d}\binom{g-d-2}{d}\frac{g^2-(2d+3)g+(d+1)(d+2)}{d+1}\\
&= \binom{g-d}{d}\binom{g-d-2}{d}\frac{g}{(d+1)(g-d)}\,. \qquad \blacksquare
\end{aligned}$$

3. The cycle numbers

The numbers of consecutive partitions and nested partitions are rather easy to obtain.

Theorem 4. $N_C(C_n,p)=\binom{n}{p}$ *for* $p\geq 2$.

Proof: There exists a one-to-one mapping between the set of consecutive p-partitions and the set of choices of p edges (deleting the edges chosen partitions the vertices into p connected components). $\blacksquare$

Corollary 5. $N_C(C_n)=2^n-n$

Proof: Let v be a vertex of the g-gon. Suppose that there are x sets of d nonintersecting diagonals involving v. Then the g vertices are involved with gx sets of d nonintersecting diagonals. But each such set involves $2d$ vertices, hence is counted $2d$ times. Thus

$$f(g,d) = gx/2d \ .$$

It follows

$$\begin{aligned} f^*(g,d) &= f(g,d) - x = f(g,x) - 2d\ f(g,x)/g \\ &= (g-2d)\ f(g,d)/g \ . \qquad \blacksquare \end{aligned}$$

We now give the recursive equations of $f(g,d)$.

Lemma 2. *$f(g,d) = 0$ for $d < 0$ or $g < 2d+2$,*

$$f(g,d) = \frac{g}{g-2d}\ f(g-1,d) + \frac{g}{g-2}\ f(g-2,d-1) \quad \textit{for} \quad g \geq 2d+2 \geq 2 \ .$$

Proof: Let v be the vertex avoided in counting $f^*(g,d)$, and let u and w be the two vertices adjacent to v on the g-gon, $f^*(g,d)$ is the sum of two types of choices: those not using the diagonal (u,w) and those using it. The number of the first type is $f(g-1,d)$, since all diagonals must come from the $(g-1)$-gon which is obtained from the g-gon by deleting v. For any choice of the second type, again we can consider that $(g-1)$-gon except only $d-1$ more diagonals are needed since (u,w) is chosen. Furthermore, no diagonal involving u or w can be chosen to avoid intersection with (u,w). Consider the $(g-2)$-gon obtained from this $(g-1)$-gon by shrinking the side (u,w). Let z denote the new vertex born from the merging of u and w. Then the requirement that no diagonal can involve either u or w is transformed into the requirement that no diagonal can involve z. Therefore $f^*(g-2,d-1)$ is the number of such choices. To summarize, we have

$$f^*(g,d) = f(g-1,d) + f^*(g-2,d-1) \ .$$

Lemma 2 now follows immediately from Lemma 1. $\blacksquare$

Theorem 3.

$$f(g,d) = \binom{g-d}{d}\binom{g-d-2}{d}\ \frac{g}{(d+1)(g-d)} \quad \textit{for} \quad g > d \geq 0 \ .$$

parts π_i, π_j such that every connected subgraph of G_n containing π_i contains a vertex of π_j, and vice versa.

Let $N_c(G_n, p)$ and $N_c(G_n)$ denote the number of p-partitions and open-partitions in the class c, where c is consecutive (C), order-consecutive (OC), or nested (N). When G_n is the complete graph K_n, then the three classes are all equal to the set of all partitions of n elements. Namely,

$$N_C(K_n, p) = N_{OC}(K_n, p) = N_N(K_n, p) = N(n, p)$$

$$\equiv \frac{1}{p!}\sum_{k=0}^{p-1}(-1)^k\binom{p}{k}(p-k)^n \quad \text{(the Bell number)},$$

and

$$N_C(K_n) = N_{OC}(K_n) = N_N(K_n) = \sum_{p=1}^{n} N(n,p) \ .$$

When G_n is the line L_n, then the problem is equivalent to the S_n problem. Hwang and Mallow gave

$$N_C(L_n, p) = \binom{n-1}{p-1}, \quad N_C(L_n) = 2^{n-1},$$

$$N_N(L_n, p) = \frac{1}{n}\binom{n}{p-1}\binom{n}{p}, \quad N_N(L_n) = \frac{1}{n+1}\binom{2n}{n} \quad \text{(the Catalan number)} \ ,$$

$$N_{OC}(L_n, p) = \sum_{j=1}^{p-1}\binom{n-1}{2p-j-2}\binom{2p-j-2}{j}, \quad N_{OC}(L_n, p) = \sum_{p=1}^{n} N_{OC}(L_n, p) \ .$$

In this paper we determine these numbers when G_n is the cycle C_n. We also study these numbers for other graphs. In particular, we give a partial order on general graphs with respect to these numbers.

2. The number of nonintersecting diagonals

To determine $N_{OC}(C_n)$, we need first to solve an auxiliary problem. Let $f(g, d)$ denote the number of ways to choose d nonintersecting (not even sharing a vertex) diagonals in a g-gon. Let $f^*(g, d)$ denote the same except that a fixed vertex of the g-gon is avoided.

Lemma 1.

$$\frac{f^*(g,d)}{f(g,d)} = \frac{g-2d}{g} \ .$$

Enumerating Consecutive and Nested Partitions for Graphs

F. K. Hwang

AT&T Bell Laboratories
Murray Hill, New Jersey 07974

*G. J. Chang**

Department of Applied Mathematics
National Chiao Tung University
Hsinchu, 30050 Taiwan, ROC

1. Introduction

Hwang and Mallows [2] studied the numbers of nested and consecutive partitions on the set $S_n = \{1, \ldots, n\}$. Consider a partition of S_n into p unordered parts. If p is pre-specified, then the partition is known as a *p-partition*; otherwise, it is an *open-partition*. A partition is *consecutive* if each part consists of numbers consecutive in S_n. A partition is *order-consecutive* if the parts can be labelled into $\pi_1, \ldots, \pi_p$ such that $\bigcup_{i=1}^{k} \pi_i$ is a set of consecutive integers for each $k = 1, \ldots, p$. A partition is *nested* if there do not exist four integers $a < b < c < d$ and two parts π_i, π_j such that a and c are in π_i while b and d in π_j. It is easily seen that a consecutive partition is also order-consecutive, and an order-consecutive partition is nested. But the following examples show the converse is not true.

Ex. 1. $\pi_1 = \{2\}$, $\pi_2 = \{1, 3\}$ is order-consecutive but not consecutive.
Ex. 2. $\pi_1 = \{2\}$, $\pi_2 = \{4\}$, $\pi_3 = \{1, 3, 5\}$ is nested but not order-consecutive.

Hwang and Mallows gave the numbers of consecutive partitions, order consecutive partitions and nested partitions for S_n.

In this paper we generalize the partition problem from the set S_n to a connected graph G_n with n vertices. A set V of vertices is called *consecutive* if the subgraph induced by V is connected. Thus the definitions of consecutiveness and order-consecutiveness easily extend from S_n to G_n. A partition of the vertices of G_n is called "nested" if there do not exist two

*Supported partially by the National Science Council of the Republic of China under grant NSC81-0208-M009-26, and partially by DIMACS, Rutgers University.