Random and Gray Code Integer Partitions

October 2, 2001

1 Generating Partitions in Gray Code Order

Wilf posed the problem: is it possible to list the partitions of an integer n in such a way that an element differs from its predecessor in the list in that one part has increased by 1 and one part has decreased by 1. (A part of size 1 may decrease to become a part of size 0 or an imaginary "part" of size 0 may increase to 1.) Carla Savage solved this problem in 1989 with a fairly complicated construction. We will present this construction below.

First we give some examples to help in gaining some insight into the problem. Let $P(n, k)$ denote the set of n-partitions with maximum part no greater k. In the following example, $P(5, 3)$ is generated using the *Combinatorica* function `Partitions`. `Partitions` generates partitions in reverse lexicographic order. In this example, this ordering is identical to Gray code ordering.

```
In[3]:= Partitions[5, 3]

Out[3]= {{3, 2}, {3, 1, 1}, {2, 2, 1}, {2, 1, 1, 1}, {1, 1, 1, 1, 1}}
```

$P(6, 4)$ is generated in reverse lexicographic order below and we see that this is not in Gray code order because $(3, 3)$ follows $(4, 1, 1)$.

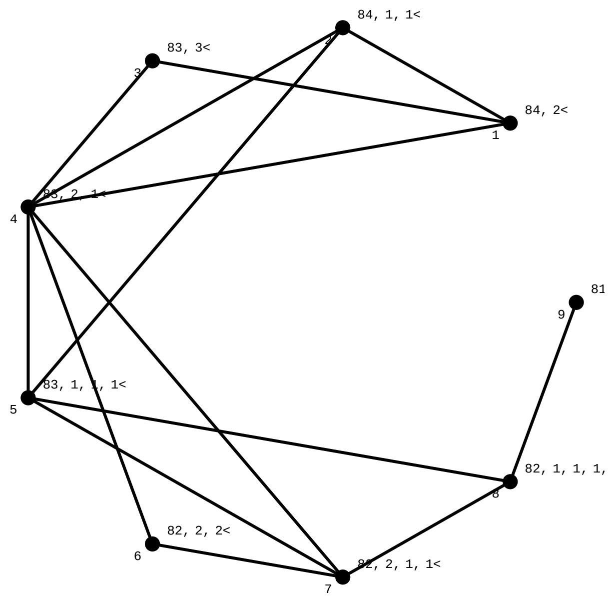
```
In[4]:= Partitions[6, 4]

Out[4]= {{4, 2}, {4, 1, 1}, {3, 3}, {3, 2, 1}, {3, 1, 1, 1}, {2, 2, 2},

>    {2, 2, 1, 1}, {2, 1, 1, 1, 1}, {1, 1, 1, 1, 1, 1}}
```

Is there a Gray code ordering of $P(6, 4)$? To answer this we first present a function `AdjacentPartitionsQ` that checks if a pair of partitions are "adjacent" in Gray code order.

```
AdjacentPartitionsQ[p_?PartitionQ, q_?PartitionQ] :=
    Module[{np = p, nq = q, diff},
            If[Length[p] < Length[q], {np, nq} = {nq, np}];
            If[Length[np] == Length[nq],
               diff = np - nq;
               Return[(Count[diff, 1] == 1 ) &&
                      (Count[diff, -1] == 1) &&
                      (Count[diff, 0] == Length[np] - 2)]
               ];
            If[Length[np] == Length[nq] + 1,
               diff = Take[np, Length[np] - 1] - nq;
               Return[(Count[diff, -1] == 1) &&
                      (Count[diff, 0] == Length[nq] - 1)]
               ];
            False
    ]
```

Using the above function, it is possible to construct a partitions graph $G_{n,k} = (V_{n,k}, E_{n,k})$, whose vertex set $V_{n,k} = P(n,k)$ and whose edge set $E_{n,k}$ consists of edges connecting vertices that are adjacent in Gray code order. Existence of Gray code ordering of $P(n,k)$ is equivalent to existence of a Hamiltonian path in $G_{n,k}$. In the example below, $G_{6,4}$ is constructed and displayed.

```
In[4]:= g = SetVertexLabels[
                MakeGraph[Partitions[6, 4], AdjacentPartitionsQ, Type -> Undirected],
                Partitions[6, 4]
            ];
In[5]:= ShowGraph[g, VertexNumber -> On];
```



It is clear that this graph does not have a Hamiltonian cycle. Does it have a Hamiltonian path? In the following we add an edge between vertex 1 (the lexicographically largest partition) and vertex 9 (the lexicographically smallest partition). The resulting graph still has not Hamiltonian cycle, implying that there is no Gray code order on $P(6,4)$ that starts with the lexicographically smallest *and* ends with the lexicographicxally largest.

```
In[9]:= HamiltonianQ[AddEdge[g, {1, 9}]]

Out[9]= False
```

However, adding an edge between vertices 2 and 9 makes the graph Hamiltonian! Below, we show the Hamiltonian path from vertex 9 which corresponds to $(1,1,1,1,1,1)$ and vertex 2 which corresponds to $(4,1,1)$. While the sequence starts with the lexicographically smallest vertex, it does not end with the lexicographically largest.

```
In[10]:= HamiltonianQ[h = AddEdge[g, {2, 9}]]

Out[10]= True

In[11]:= c = HamiltonianCycle[h]
```

```
Out[11]= {1, 2, 9, 8, 5, 7, 6, 4, 3, 1}

In[13]:= Partitions[6, 4][[ RotateLeft[Rest[c]] ]]

Out[13]= {{1, 1, 1, 1, 1, 1}, {2, 1, 1, 1, 1}, {3, 1, 1, 1}, {2, 2, 1, 1},

>     {2, 2, 2}, {3, 2, 1}, {3, 3}, {4, 2}, {4, 1, 1}}
```

Savage's goal is to construct a list $L(n,k)$ of elements in $P(n,k)$ in Gray code order. Let $\min P(n,k)$ and $\max P(n,k)$ denote the lexicographically smallest and lexicographically largest elements in $P(n,k)$ respectively. Savage would like to additionally claim that $L(n,k)$ starts with $\min P(n,k)$ and ends with $\max P(n,k)$. However, as we have seen already, this is false for $(6,4)$. So she makes the "next best" claim: for all integers $(n,k) \neq (6,4)$, it is possible to construct a list $L(n,k)$ of elements in $P(n,k)$ in Gray code order such that the list starts with $\min P(n,k)$ and ends with $\max P(n,k)$. For $(6,4)$ we know that there is a list $L(6,4)$ containing the elements in $P(6,4)$ in Gray code order starting with $\min P(6,4)$.

Ideally, we would like to prove the existence of $L(n,k)$ by induction, expressing $L(n,k)$ in terms of $L(n',k')$ for smaller n' and k'. However, matters are not as simple as that and we need to start with a stronger claim for the induction to go through. First we introduce some notation. It will be convenient to think of partitions as strings of integers. So we will write $(4,2,1,1)$ as $4 \cdot 2 \cdot 1 \cdot 1$ and further more we will use the shorthard a^b to denote b copies of a. So $(4,2,1,1)$ will be written as $4 \cdot 2 \cdot 1^2$. For any set of strings S and an element x, i we use $x \oplus S$ to denote the set $\{x \cdot w \mid w \in S\}$. So $x \oplus S$ denotes the set obtained from S by prepending x to each element in S. $S \oplus x$ is similarly defined. We will use $S \oplus x$ and $x \oplus S$ even if S is a sequence instead of a set. Finally, for any sequence S, $\overline{S}$ will denote its reverse.

The strengthened claim is as follows: for all non-negative integers n and k,

- there exists a list $L(n,k)$ of elements in $P(n,k)$ in Gray code order which, unless $(n,k) = (6,4)$, begins with $\min P(n,k)$ and ends with $\max P(n,k)$.
- and if $k \geq 1$ and $n \geq 2k+1$, there exists a list $M(n,k)$ of elements in

  $$P(n,k) \cup (k+1 \oplus P(n-k-1,k-1))$$

  in Gray code order, beginning with $\min P(n,k)$ and ending with $\max P(n,k)$.

We prove the existence of both $L(n,k)$ and $M(n,k)$ by induction. Temporarily we will behave as though $L(6,4)$ is not a problem and fix this problem later. We start with $L(n,k)$ and consider the cases $n < 2k-2$, $n \geq 2k$, $n = 2k-2$, and $n = 2k-1$ separately.

Case 1: $n < 2k-2$ This is the easy case. In this case, we claim that $L(n,k)$ can be constructed as follows

$$\begin{aligned} L(n,k) &= L(n,k-2), \\ &\quad (k-1) \oplus \overline{L(n-k+1,k-1)}, \\ &\quad k \oplus L(n-k,k). \end{aligned}$$

To see the correctness of this, observe that $L(n,k-2)$ starts with 1^n and ends with

(i) $(k-2)^2 \cdot 1$ if $n = 2k-3$.

(ii) $k-2 \cdot (n-k+2)$ if $n < 2k-3$.

Since $n < 2k-2$, $n-k+1 < k-1$ and so $L(n-k+1,k-1) = L(n-k+1,n-k+1)$. The first element of $L(n-k+1,n-k+1)$ is 1^{n-k+1} and the last element is $n-k+1$.

Hence, the first element of $(k-1) \oplus \overline{L(n-k+1,k-1)}$ is $k-1 \cdot (n-k+1)$ and the last element is $k-1 \cdot 1^{n-k+1}$. Similarly, since $n < 2k-2$, we have that $n-k < k-2$ and so $L(n-k,k) = L(n-k,n-k)$. The first element of $L(n-k,n-k)$ is 1^{n-k} and the last element is $n-k$. This implies that the first element of $k \oplus L(n-k,k)$ is $k \cdot 1^{n-k}$ and the last element is $k \cdot (n-k)$.

The induction hypothesis tells us that within each list the Gray code property holds. We only need to make sure that this is true at the "boundary" between then lists. At the boundary between the first and the second lists we have:

- $(k-2)^2 \cdot 1$ followed by $(k-1) \cdot (k-2)$ if $n = 2n-3$.
- $(k-2) \cdot (n-k+2)$ followed by $k-1 \cdot (n-k+1)$ if $n < 2k-3$.

In each case, the second partition can be obtained from the first by incrementing a single part and decrementing a single part. At the boundary between the second and third lists we have $(k-1) \cdot 1^{n-k+1}$ followed by $k \cdot 1^{n-k}$ which clearly satisfies the Gray code property. Finally, note that the first element in the entire sequence is 1^n and the last element in the entire sequence is $k \cdot (n-k)$. These are, as predicted, $\min P(n,k)$ and $\max P(n,k)$ respectively.

Case 2: $n = 2k-2$ In this case, $L(n,k)$ is constructed as follows:

$$\begin{aligned} L(n,k) &= L(n,k-2), \\ &\quad (k-1) \oplus \overline{L(n-k+1,k-3)}, \\ &\quad k \oplus L(n-k,k-3), \\ &\quad (k-1) \cdot (k-2) \cdot 1, \\ &\quad (k-1) \cdot (k-1), \\ &\quad k \cdot (k-2). \end{aligned}$$

So $L(n,k)$ is the concatenation of three lists, to which three partitions are appended. First convince yourself that all elements in $P(n,k)$ appear in the right hand side above. The first list on the right hand side $L(n,k-2)$ starts with 1^n and ends with $(k-2)^2 \cdot 2$. The list $\underline{L(n-k+1,k-3)}$ starts with 1^{n-k+1} and ends with $(k-3) \cdot 2$. Therefore, the list $(k-1) \oplus \overline{L(n-k+1,k-3)}$ starts with $(k-1) \cdot (k-3) \cdot 2$ and ends with $(k-1) \cdot 1^{n-k+1} = (k-1) \cdot 1^{k-1}$. Therefore at the boundary between the first and second list we have $(k-2)^2 \cdot 2$ followed by $(k-1) \cdot (k-3) \cdot 2$ and these are Gray code adjacent. $L(n-k,k-3)$ starts with 1^{n-k} and ends with $(k-3) \cdot 1$. Therefore, $k \oplus L(n-k,k-3)$ starts with $k \cdot 1^{n-k} = k \cdot 1^{k-2}$ and ends with $k \cdot (k-3) \cdot 1$. The boundary between the second and third list contains the partition $(k-1) \cdot 1^{k-1}$ followed by $k \cdot 1^{k-2}$ and these are Gray code adjacent. The third list ends with $k \cdot (k-3) \cdot 1$ and immediately following that we have $(k-1) \cdot (k-2) \cdot 1, (k-1) \cdot (k-1), k \cdot (k-2)$. Clearly, these four partitions are Gray code adjacent. Finally, note that the list on the right hand side starts with 1^n and ends with $k \cdot (k-2)$ thereby also satisfying the property that the list begin with $\min P(n,k)$ and end with $\max P(n,k)$.

Case 3: $n = 2k-1$ In this case $L(n,k)$ is constructed as follows:

$$L(n,k) = M(n,k-1), k \cdot (k-1).$$

Recall that $M(n,k-1)$ is a list of elements in $P(n,k-1) \cup k \oplus P(n-k,k-2)$ that starts with $\min P(n,k-1) = 1^n$ and ends with $\max P(n,k-1) = (k-1)^2 \cdot 1$. First note that the right hand side of the above equation contains all elements in $P(n,k)$. Then observe that since the last element of $M(n,k-1)$ is $(k-1)^2 \cdot 1$ and this is followed by $k \cdot (k-1)$, the Gray code property is satisfied. Also, the first element of the right hand side is 1^n and the last element is $k \cdot (k-1)$ and these are the lexicographically smallest and largest elements, respectively, in $P(n,k)$.

Case 4: $n \geq 2k$ First partition $P(n,k)$ into three groups of partitions

Group 1 $P(n,k-1) \cup k \oplus P(n-k,k-2)$

Group 2 $k \oplus (k-1) \oplus P(n-2k+1,k-1)$

Group 3 $k \oplus k \oplus P(n-2k,k)$

Partitions whose maximum part is at most $k-1$ are in Group 1. Partitions with exactly one part of size k appears partly in Group 1 and partly in Group 2. Partitions with more than one part of size k appear in Group 3. We use $M(n,k-1)$ as a listing of the Group 1 partitions, $k \oplus (k-1) \oplus \overline{L(n-2k+1,k-1)}$ as a listing of Group 2 partitions, and $k \oplus k \oplus L(n-2k,k)$ as a listing of Group 3 partitions. Thus, $L(n,k)$ is constructed as

$$\begin{aligned} L(n,k) \quad = \quad & M(n,k-1), \\ & k \oplus (k-1) \oplus \overline{L(n-2k+1,k-1)}, \\ & k \oplus k \oplus L(n-2k,k). \end{aligned}$$

The list $M(n,k-1)$ starts with $\min P(n,k-1) = 1^n$ and ends with $\max P(n,k-1) = (k-1)^p \cdot q$ where p is the quotient obtained from dividing n by $(k-1)$ and q is remainder. Note that if $q=0$ it does not appear in the partition. The list $L(n-2k+1,k-1)$ starts with 1^{n-2k+1} and ends with $(k-1)^r \cdot s$ where r is the quotient obtained from dividing $(n-2k+1)$ by $(k-1)$ and s is the remainder. Therefore, the list $k \oplus (k-1) \oplus \overline{L(n-2k+1,k-1)}$ starts with $k \cdot (k-1) \cdot (k-1)^r \cdot s$ and ends with $k \cdot (k-1) \cdot 1^{n-2k+1}$. So at the boundary between the first and second lists we have $(k-1)^p \cdot q$ followed by $k \cdot (k-1) \cdot (k-1)^r \cdot s = k \cdot (k-1)^{r+1} \cdot s$. It is clear that $p = r+2$ and $q = s+1$ and that these two partitions satisfy the Gray code property. The list $L(n-2k,k)$ starts with 1^{n-2k} and ends with $k^u \cdot v$, where u is the quotient obtained from dividing n by k and v is the remainder. Therefore, the list $k \oplus k \oplus L(n-2k,k)$ starts with $k^2 \cdot 1^{n-2k}$ and ends with $k^2 \cdot k^u \cdot v = k^{2+u} \cdot v$. The boundary between the second and third lists consists of the partition $k \cdot (k-1) \cdot 1^{n-2k+1}$ followed by $k^2 \cdot 1^{n-2k}$. Clearly, these two satisfy the Gray code property. Finally, note that the list on the right hand side starts with 1^n and ends with $k^{2+u} \cdot v$ and these are the lexicographically first and last partitions of $P(n,k)$.