

## Lecture 21

*Lecture date: April 28, 2011**Notes by: Jacob Bower*

## 1 De Bruijn Sequences

Today's lecture begins with a discussion of De Bruijn sequences. We have the following definition:

**Definition 1** *A binary De Bruijn Sequence of order  $n$  is a string of bits  $b_i \in \{0, 1\}$ ,  $b = \{b_1, \dots, b_{2^n}\}$  such that every string of length  $n$ ,  $\{a_1, \dots, a_n\} \in \{0, 1\}^n$ , occurs exactly once consecutively in  $b$ . (In other words  $\exists! k$  such that  $b_k = a_1, b_{k+1} = a_2, \dots, b_{k+n-1} = a_n$  where we take indices mod  $2^n$ .)*

This definition can perhaps best be illustrated with the example:

**Example 2** *A De Bruijn sequence of order 3 is 00010111.*

One can check by examination that all possible bitstrings of length three occur once and only once in this sequence. Remarkably, it turns out that De Bruijn sequences of order  $n$  exist for all  $n$ . We will shortly see a theorem that enumerates them, but before we look at this we first examine the definition of a De Bruijn graph, an object that will play an essential role in our proof of our enumeration theorem.

**Definition 3** *A De Bruijn graph is a directed graph*

$$DB_n = (\{0, 1\}^{n-1}, \{0, 1\}^n),$$

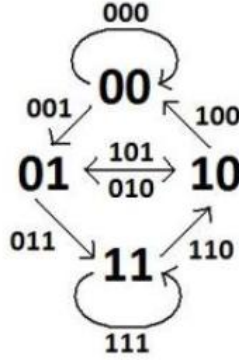
*that is, a graph with vertex set  $\{0, 1\}^{n-1}$  and edge set  $\{0, 1\}^n$  where there is an edge from  $a = \{a_1, \dots, a_{n-1}\}$  to  $a' = \{a'_1, \dots, a'_{n-1}\}$  iff:*

$$a_2 = a'_1, a_3 = a'_2, \dots, a_{n-1} = a'_{n-2}.$$

*This is equivalent to saying there is an edge from  $a$  to  $a'$  iff  $a$  is a prefix of some string  $b_1 \dots b_n$  and  $a'$  is a suffix of the same string.*

We can again look at an example to illustrate this definition.

**Example 4** The graph of  $DB_3$  is:



With this definition we are ready to state and prove our theorem about the enumeration of De Bruijn sequences.

**Theorem 5** Binary De Bruijn sequences of order  $n$  exist for all  $n$ , and there are exactly  $2^{2^n - 1}$  of them.

**Proof:** To begin our proof we note that a De Bruijn sequence of order  $n$  is just an Eulerian tour of  $DB_n$ . We first wish to prove that an Eulerian Tour exists. To do this we note that  $DB_n$  is balanced. This is because we can take any sequence of length  $n - 1$  and delete the last digit, then add a 0 or 1 in front, giving us that  $indeg = 2$  for each vertex. Similarly we can take any sequence of length  $n - 1$  and delete the first digit, then add a 0 or a 1 to the end, giving us that  $outdeg = 2$  for each vertex. Because  $indeg = outdeg = 2$  our graph is balanced. As we have seen in previously an Eulerian tour exists as long as a graph is balanced, so we know that that for all  $n$  there exists a binary De Bruijn sequence of order  $n$ .

Now that we have proved the existence of De Bruijn sequences we want to prove their enumeration. From the equivalence of Eulerian Tours in  $DB_n$  and De Bruijn sequences combined with results we have found from graph theory we have:

$$\begin{aligned}
 \#\{DB \text{ sequences of order } n\} &= \#\{\text{Eulerian Tours of } DB_n\} \\
 &= \sum_{e \in E} \tau(DB_n, e) \\
 &= \sum_{e \in E} \kappa(DB_n, e) \prod_{v \in V} (outdeg(v) - 1)! \\
 &= \sum_{e \in E} \kappa(DB_n, e)
 \end{aligned}$$

The last equation here comes from the fact that every vertex of the graph has  $outdeg = 2$  as we discussed before. We thus want to count the number of spanning trees of  $DB_n$  in order to count the number De Bruijn sequences. To do this we use a special trick. Consider any two vertices  $v$  and  $w$  in  $DB_n$  and count the number of paths of length  $n - 1$  between  $v$  and  $w$ . In particular let:

$$v = (v_1, \dots, v_{n-1})$$

$$w = (w_1, \dots, w_{n-1})$$

Looking at these there is only 1 way to append our two sequences. In particular we must do it in the fashion:

$$(v_1 \dots v_{n-1})w_1 \dots w_{n-1}$$

$$(v_1(v_2 \dots w_1)w_2 \dots w_{n-1})$$

$$v_1 v_2 (v_3 \dots w_2)w_3 \dots w_{n-1}$$

$$\vdots$$

$$v_1 \dots v_{n-1} (w_1 \dots w_{n-1}).$$

Because we have exactly one path from  $v$  to  $w$  of length  $n - 1$  for all vertices  $v$  and  $w$  we must have:

$$A^{n-1} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$$

where  $A$  is the  $2^{n-1} \times 2^{n-1}$  adjacency matrix of  $DB_n$ .

We should then see the Laplacian Matrix having eigenvalues of 0 with multiplicity 1 and 2 with multiplicity  $2^{n-1}$ .

Finally we apply matrix-tree-theory to arrive at the result:

$$\#\{DB \text{ sequences of order } n\} = \sum_{e \in E} \kappa(DB_n, e)$$

$$= \sum_{e \in E} \frac{\lambda_1 \dots \lambda_{(2^{n-1}-1)}}{2^{n-1}}$$

$$= \sum_{e \in E} \frac{2^{2^{n-1}-1}}{2^{n-1}}$$

$$= 2^n \frac{2^{2^{n-1}-1}}{2^{n-1}}$$

$$= 2^{2^{n-1}}.$$

□

**Remark 6** One may note that the number of De Bruijn sequences is:

$$2^{2^{n-1}} = \sqrt{2^{2^n}}$$

where  $2^{2^n}$  counts the number of all possible bitstrings of length  $2^n$ . This leads to the question if there is a bijection:

$$\{\text{pairs of DB sequences of order } n\} \leftrightarrow \{\text{all binary bitstrings of length } 2^n\}.$$

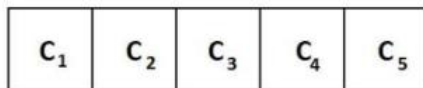
It turns out there is, and this fact was proved by two MIT students, Kishore and Bidkhor, illustrating the many open and accessible problems that exist in the exciting field of combinatorics!

This concludes our discussion of De Bruijn sequences.

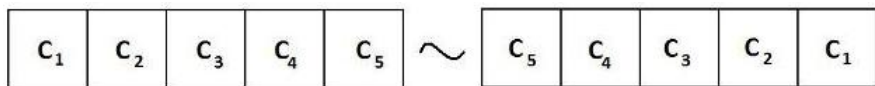
## 2 Pólya Theory

The general idea of Pólya theory is counting things up to certain equivalence classes. Let us look at a very simple example to start.

**Example 7** Let us consider the number of colorings of 5 squares in a line, one with each of  $n$  colors. That is the number of colorings of:



It is straight forward to see that under no restrictions the number of colorings is  $n^5$ , because we can choose any of the  $n$  colors separately for each of the 5 squares. As a less trivial example let us impose the equivalence that two colorings are the same if they are mirror images of each other. That is, we have equivalence:



Under this equivalence first let us consider the number of colorings that are equivalent to themselves. This is all colorings of the form:

$c_1$	$c_2$	$c_3$	$c_2$	$c_1$
-------	-------	-------	-------	-------

Because we have three relevant colors to choose and  $n$  choices for each of them, there are  $n^3$  such colorings. All other colorings have exactly one unique equivalent coloring, so we take the remaining  $n^5 - n^3$  colorings and divide them by 2 to account for the fact each one of these is equivalent to exactly one other. We then conclude:

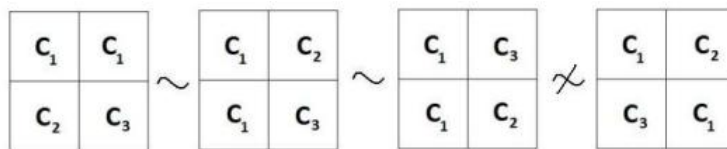
$$\begin{aligned} & \#\{\text{equivalence classes of colorings}\} \\ &= n^3(\text{symmetric colorings}) + \frac{n^5 - n^3}{2}(\text{non-symmetric colorings}) \\ &= \frac{1}{2}(n^5 + n^3). \end{aligned}$$

In general we have some finite set  $X$  where  $|X| = k$ , a group  $S \subseteq S_k$  (a subgroup of the group of all permutations of  $X$ ), and a finite set of colors  $C$ . Then the set of colorings is  $Y = C^X = \{f : X \rightarrow C\}$ . We then consider some action  $G$  that acts on  $Y$ , so that  $G \times Y \rightarrow Y$ . We then wish to count the orbits  $|Y/G|$  (i.e. equivalence of colorings). As a more complex example consider the following problem.

**Example 8** Let us consider the case where  $X = [2] \times [2]$  where we color with the set  $C = [n]$ , so that we are looking at a square of the form:

$c_1$	$c_2$
$c_3$	$c_4$

We wish to count the number of colorings under the action  $G = D_8$  which is equivalent to saying any two colorings are equivalent if any reflection or rotation of one produces the other. We can consider the following as examples of equivalence (where the last block is NOT equivalent to the other three):



We can then count the number of equivalence classes directly. We have the following classes occurring with the following multiplicities.

 $\sim$ 

$c_1$	$c_1$
$c_1$	$c_1$

$n$  - because we have  $n$  choices for  $c_1$ .

 $\sim$ 

$c_1$	$c_2$
$c_1$	$c_1$

$n(n-1)$  - because we have  $n$  choices for  $c_1$  then  $n-1$  choices for  $c_2$ .

 $\sim$ 

$c_1$	$c_2$
$c_1$	$c_2$

$\binom{n}{2}$  - because we must choose  $c_1$  and  $c_2$  but we can restrict to  $c_1 > c_2$  due to symmetry.

 $\sim$ 

$c_1$	$c_2$
$c_2$	$c_1$

$\binom{n}{2}$  - because we must choose  $c_1$  and  $c_2$  but we can restrict to  $c_1 > c_2$  due to symmetry.

 $\sim$ 

$c_1$	$c_1$
$c_2$	$c_3$

$n\binom{n-1}{2}$  - because we must choose  $c_1$  from  $n$  colors, then choose  $c_2$  and  $c_3$  but we can restrict  $c_2 > c_3$  by symmetry.

 $\sim$ 

$c_1$	$c_2$
$c_3$	$c_1$

$n\binom{n-1}{2}$  - because we must choose  $c_1$  from  $n$  colors, then choose  $c_2$  and  $c_3$  but we can restrict  $c_2 > c_3$  by symmetry.

 $\sim$ 

$c_1$	$c_2$
$c_3$	$c_4$

$3\binom{n}{4}$  - because we choose 4 colors then have to choose 1 of  $\{c_2, c_3, c_4\}$  to be opposite  $c_1$ .

To count the total number of colorings that are not equivalent under rotation or reflection we sum up the multiplicities of the equivalences. Doing this we conclude:

$$\begin{aligned} & \#\{\text{non-equivalent colorings}\} \\ &= n + n(n-1) + \binom{n}{2} + \binom{n}{2} + n\binom{n-1}{2} + n\binom{n-1}{2} + 3\binom{n}{4} \\ &= \frac{1}{8}(n^4 + 2n^3 + 3n^2 + 2n). \end{aligned}$$

**Theorem 9** The number of equivalence classes of colorings with  $n$  colors equals:

$$\frac{1}{|G|} \sum_{\pi \in G} n^{C(\pi)}$$

where

$$C(\pi) = \# \text{ of cycles of permutation } \pi.$$

**Proof:** This proof makes use of Burnside's Lemma.  $\square$

We may wish to generalize this theorem. Consider the vector:

$$I = (i_1, \dots, i_n),$$

where

$$\sum_{j=1}^n i_j = k.$$

We may then ask how many equivalence class of colorings exist such that color  $j$  occurs  $i_j$  times. We will call this number  $K(I)$ . To study this we first present the definition of the cycle indicator.

**Definition 10** The cycle indicator of the pair  $(X, G)$  is the polynomial:

$$Z_g(z_1, \dots, z_k) = \frac{1}{|G|} \sum_{\pi \in G} z_1^{C_1(\pi)} \dots z_k^{C_k(\pi)}$$

where  $c_j(\pi) = \# \{ \text{cycles of length } j \text{ in } \pi \}$ .

With this definition we then have then general, weighted version of Pólya's Theorem.

**Theorem 11** Suppose we are given  $I = (i_1, \dots, i_n)$  where

$$\sum_{j=1}^n i_j = k.$$

Call the number of equivalence classes of colorings such that color  $j$  occurs  $i_j$  times  $K(I)$ .

Let:

$$F_G(r_1, \dots, r_n) = \sum_I K(I) r^I$$

where  $r^I = r_1^{i_1} \dots r_n^{i_n}$ . Then:

$$F_g(r_1, \dots, r_n) = Z_G(p_1, \dots, p_k)$$

where:

$$p_1 = r_1 + \dots + r_n$$

$$p_2 = r_1^2 + \dots + r_n^2$$

$$\vdots$$

$$p_k = r_1^k + \dots + r_n^k.$$