Construction and Optimality of an Unoriented Variation on de Bruijn Sequences

Christie S. Burris
Colorado State University
08 · 20 · 15

Examples

00110	0100011101	22202101020002212111010120112		
00	010	222	200	110
01	100	220	000	101
11	000	202	002	100
10	001	021	022	001
	011	210	221	012
	111	101	212	120
	110	010	121	201
	101	102	221	011
		020	111	112

- Σ_k = alphabet of size k
- n = length of subwords (underlined)

EXAMPLES

$B(k,n)$	$B(2,2)$	$B(2,3)$	$B(3,3)$		
	00110	**0100011101**	**2220210102000...**		
subwords	00	010	222	200	110
	01	100	220	000	101
	11	000	202	002	100
	10	001	021	022	001
		011	210	221	012
		111	101	212	120
		110	010	121	201
		101	102	221	011
			020	111	112

EXAMPLES

$B(k,n)$	$B(2,2)$	$B(2,3)$	$B(3,3)$		
	00110	**0100011101**	**2220210102000...**		
subwords	00	010	222	200	110
	01	100	220	000	101
	11	000	202	002	100
	10	001	021	022	001
		011	210	221	012
		111	101	212	120
		110	010	121	201
		101	102	221	011
			020	111	112

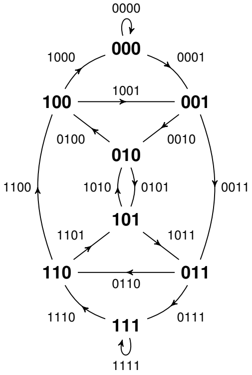
CONSTRUCTION

Forming a de Bruijn sequence

Traverse an Eulerian circuit in $Bg(2,4)$:

$$000 \xrightarrow{0001} 001 \xrightarrow{0010} 010 \xrightarrow{0100} 100 \xrightarrow{1001} 001 \xrightarrow{0011} 011 \xrightarrow{0111}$$
$$111 \xrightarrow{1111} 111 \xrightarrow{1110} 110 \xrightarrow{1101} 101 \xrightarrow{1010} 010 \xrightarrow{0101} 101 \xrightarrow{1011}$$
$$011 \xrightarrow{0110} 110 \xrightarrow{1100} 100 \xrightarrow{1000} 000$$

$B(2,4) = 0001001111010110000$



$Bg(2,4)$

DEFINITION: de Bruijn Sequence, $B(k, n)$

A finite sequence with elements drawn from an alphabet Σ_k of size k whose subwords of length n are exactly the k^n words of length n on k characters.

- $l(k, n) = k^n + n - 1$ is the length of any de Bruijn sequence
- de Bruijn proved that such sequences exist for all k and n

How did de Bruijn prove this?

Theory

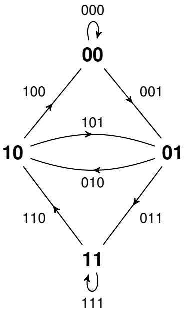
DEFINITION: Eulerian Circuit (Path)

A circuit that traverses every edge of a graph exactly once, and begins and ends on the same (different) vertex (vertices).

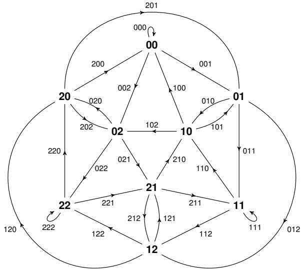
Condition imposed on graphs where an Eulerian circuit (path) exists: **The number of odd degree vertices is 0 (2)**.

When the graph has 2 odd degree vertices, the Eulerian path starts and ends on these vertices.

DE BRUIJN GRAPHS



(a) $Bg(2,3)$



(b) $Bg(3,3)$

For every $v \in Bg(k,n)$, $\deg v = 2k$. i.e. traversing an Eulerian circuit in the graph gives a de Bruijn sequence.

Unorientation

Consider **00010111**, $n = 3$.

Traverse forward: $\xrightarrow{00010111}$, and backward: $\xleftarrow{00010111}$

$$\begin{array}{ccc}
000 & \quad & 111 \\
001 & & 110 \\
010 & & 101 \\
101 & & 010 \\
011 & & 100 \\
111 & & 000
\end{array}$$

We say $\{001, 100\}$ and $\{011, 110\}$ are *reflected pairs*.

Unoriented de Bruijn Sequences

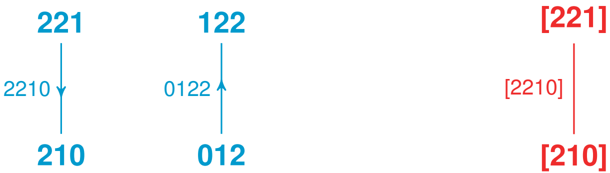
Definition: Unoriented de Bruijn Sequence $uB(k,n)$

A sequence of characters drawn from an alphabet Σ_k of k characters that contains as subwords one member from each of the length-n reflections on k symbols exactly once.

- By combinatorial argument, the length of any sequence $uB(k,n)$ is

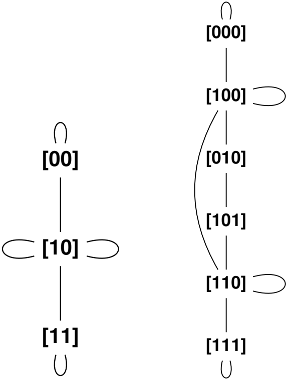
$$l(k,n) = \frac{k^n + k^{\lceil n/2 \rceil} + 2n - 2}{2}.$$

Why unoriented?



This is the case in general. The undirected edges $[v] \leftrightarrow [w]$, may be thought of as being formed from the two directed edges $v \to w$ and $w' \to v'$.

Unoriented de Bruijn Graphs



(c) $uBg(2,3)$ (d) $uBg(2,4)$

Vertices (edges) represent length-$(n-1)$ (length-n) reflected pairs.

Nonexistence

Theorem: Unoriented de Bruijn sequences $uB(k, n)$ do not exist if either k or n is greater than 3.

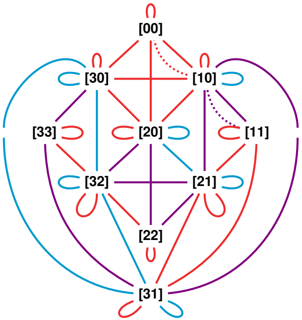


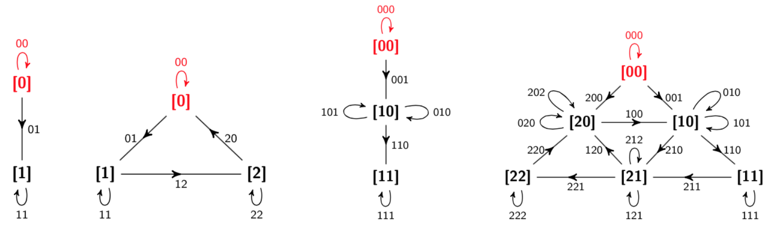
Figure : $uBg(4, 3)$

Example: An Eulerian path doesn't exist because 4 vertices have odd degree.

Existing Unoriented de Bruijn Sequences

	$k = 2$	$k = 3$
$n = 2$	0011	0011220
$n = 3$	00010111	00010111212020122200

Existing Unoriented de Bruijn Sequences

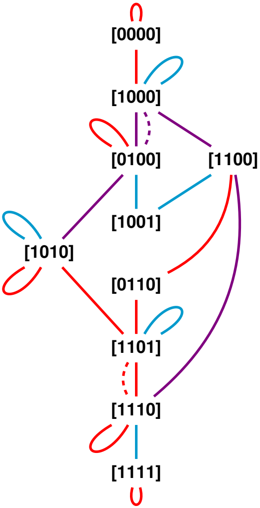


Arrows indicate path taken on the graph (edges remain undirected). Red edges and vertices represent the start of the path.

Future Work

- Eulerizations
- Can we create an algorithm to find suboptimal sequences?

Dotted vertices represent an *Eulerization* of the graph. One can form a sequences *almost* as optimal as an unoriented de Bruijn sequence.



Questions, Comments?

Thank you!