Existence of Balanced Generalized deBruijn Sequences

Bhumika Mittal
(Co-authored by: Matthew Baker, Haran Mouli, Eric Tang)

Ashoka University

June 19, 2025

The Introductory Trick

Performing the Trick

- Take a normal deck of playing cards with 52 cards plus two Jokers.
- Deal the top five cards into a pile on the table.
- The player should have more cards of one color than the other. The player discards all cards of the **minority** color.
- Choose **one** of the suits and discard everything else.
- Choose any one card from those remaining in your hand and discard the rest. The magician can predict the chosen card.

The Actual Deck

The deck was pre-determined. As it so happens, the information given by the order and position of the discards uniquely determines every five-hand substring of this 54-card deck. This particular deck is known as the *Sledgehammer Stack*.

Sledgehammer Stack

SCCSCHCHSCSSHHCSHDSHCHDCDDJCSCDDDDCDSSHCSHHDSJSHCHHDDD

DeBruijn Sequences

deBruijn Sequences

Binary deBruijn Sequence

A **binary deBruijn sequence** of order l, is a length 2^l binary string, where every binary sub-string of length l occurs exactly once.

Balanced Cyclic de Bruijn Sequences

Balanced Cyclic de Bruijn Sequence

A **balanced cyclic de Bruijn sequence** with parameters (n, l, k) is an n-bit binary sequence that satisfies the following:

- The sequence contains an equal number of 0's and 1's.
- Each substring of length l occurs at most k times.

Theorem

For any positive integers n, l, k, a balanced cyclic de Bruijn sequence with parameters (n, l, k) exists if and only if n is even and $k \geqslant \frac{n}{2^l}$.

Some important lemmas

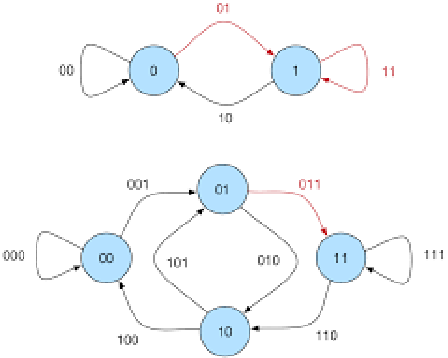
De Bruijn Graphs

Definition

For $l \geqslant 2$, the **de Bruijn Graph** of order l, denoted by G_l, is a directed graph on 2^{l-1} vertices labelled $\{0,1\}^{l-1}$ and 2^l edges labelled $\{0,1\}^l$. For any l bits $b_0, b_1, \cdots, b_{l-1}$, the edge $e = b_0 b_1 \cdots b_{l-1}$ connects $v_1 = b_0 b_1 \cdots b_{l-2}$ to $v_2 = b_1 b_2 \cdots b_{l-1}$.

Lemma

Every vertex in the deBruijn graph has indegree and outdegree 2.



Lemma

The deBruijn graph is Eulerian.

Proof

We already know for all vertex in G_l, in-degree = out-degree.
It is sufficient to show that G_l is connected. Let $v = b_0 b_1 \cdots b_{l-2}$ and $v' = b'_0 b'_1 \cdots b'_{l-2}$. $\exists$ walk from v to v':

$$b_0 b_1 \cdots b_{l-2} \to b_1 \cdots b_{l-2} b'_0 \to \cdots \to b'_0 b'_1 \cdots b'_{l-2}.$$

Definition

We define an edge in G_l to be **red** if its last bit is 0 and **blue** if its last bit is 1. A graph or subgraph is **balanced** if it contains equal number of red and blue edges.

Lemma

For any vertex $v \in V(G_l)$, v has one red and one blue outedge, and the inedges of v are of the same color.

Proof

Let $v = b_0 b_1 \cdots b_{l-2}$ be a vertex in G_l. Then by definition, the out-edges $b_0 b_1 \cdots b_{l-2} 0$ and $b_0 b_1 \cdots b_{l-2} 1$ are red and blue, respectively. The in-edges $0 b_0 b_1 \cdots b_{l-2}$ and $1 b_0 b_1 \cdots b_{l-2}$ must end with the same bit, and are thus of the same color.

Lemma

A balanced circuit of length n in G_l corresponds to a balanced cyclic de Bruijn sequence with parameters $(n, l, 1)$.

Proof

Let the edges of the circuit of length n be denoted as:

$$b_0 b_1 \cdots b_{l-1} \rightarrow b_1 b_2 \cdots b_l \rightarrow \cdots \rightarrow b_{n-1} b_0 \cdots b_{l-2}.$$

The l-bit substrings of the n-bit sequence $b_0 b_1 \cdots b_{n-1}$ precisely correspond to the edges of the circuit of length n. Thus, as the edges on the circuit are distinct, so are the l-bit substrings. Moreover, the number of 0s and 1s in the string are equal as well.

Lemma

If the graph G_l has a balanced circuit of length n, then the graph G_{l+1} has a balanced cycle of length n.

Lemma

Let $l \geqslant 2$, and let n be an even positive integer at most 2^l. Then, the graph G_l contains a balanced circuit of length n.

Theorem

For any positive integers n, l, k, a balanced cyclic de Bruijn sequence with parameters (n, l, k) exists if and only if n is even and $k \geqslant \frac{n}{2^l}$.

Proof

($\Rightarrow$) Since a balanced generalized de Bruijn sequence has an equal number of 0s and 1s, n must be even. Moreover, by the Pigeonhole Principle, we must have $k \geqslant \frac{n}{2^l}$, since there are only 2^l possible l-bit substrings and n total substrings.

Now, let's show that these constraints are sufficient for a balanced cyclic DeBruijn sequence to exist. Assume that $k = \lceil \frac{n}{2^l} \rceil$.

Let $l \geqslant 2$. We'll use the induction on k to prove our claim. From the above lemmas, the base case $k = 1$ follows.

Let $b_0 b_1 \cdots b_{n-1}$ be a balanced cyclic DeBruijn sequence with parameters (n, l, k), and let $b'_0 b'_1 \cdots b'_{2^l-1}$ be one with parameters $(2^l, l, 1)$. The latter contains every l-bit string as a substring exactly once. We may hence assume that $b'_0 b'_1 \cdots b'_{l-1} = b_0 b_1 \cdots b_{l-1}$. Hence:

$$b_0 b_1 \cdots b_{2^l-1} b'_0 b'_1 \cdots b'_{l-1}$$

must be a balanced sequence, where each substring of length l appears at most k times in the first part of the sequence and exactly once in the second part.

Thus, this must be a balanced cyclic DeBruijn sequence with parameters $(n + 2^l, l, k + 1)$. Hence, our theorem follows by induction. □

Let's go back to the trick

The trick is based on a balanced generalized de Bruijn sequence with parameters $(52, 5, 2)$, for example the following one:

0000011101010010001011001101111100000101101111101001

From the following cheatsheet, the magician knows which substring is in use and can determine the card using the discard order.

$A\heartsuit$ $7\heartsuit$ $3\diamondsuit$ $Q\diamondsuit$ $2\diamondsuit$ $K\spadesuit$ $8\spadesuit$ $10\spadesuit$ $2\heartsuit$ $7\clubsuit$

$K\diamondsuit$ $3\clubsuit$ $5\diamondsuit$ $10\diamondsuit$ $6\clubsuit$ $6\heartsuit$ $8\diamondsuit$ $9\heartsuit$ $Q\clubsuit$ $J\heartsuit$

$J\spadesuit$ $5\clubsuit$ $3\heartsuit$ $Q\heartsuit$ $A\clubsuit$ $2\clubsuit 4\diamondsuit$ $5\spadesuit$ $9\spadesuit$ $10\clubsuit$ $7\spadesuit$

$4\clubsuit$ $A\diamondsuit$ $7\diamondsuit$ $6\diamondsuit$ $8\heartsuit$ $9\diamondsuit$ $Q\spadesuit$ $4\heartsuit$ $2\spadesuit$ $6\spadesuit$ $J\diamondsuit$

$9\clubsuit$ $K\clubsuit$ $8\clubsuit$ $J\clubsuit$ $3\spadesuit$ $5\heartsuit$ $A\spadesuit$ $10\heartsuit$ $K\heartsuit$ $4\spadesuit$


Thank You!

Proof

By induction on l:

Basis: For $l = 2$, we can either have $n = 2$ or $n = 4$. We may take the circuits $0 \rightarrow 1$ and $0 \rightarrow 1 \rightarrow 1 \rightarrow 0$, respectively.

Hypothesis: Assume that the claim holds true for l. We need to prove it holds true for $l + 1$.

Induction Step: For even $n \leqslant 2^l$, we may use the induction hypothesis to find a balanced circuit of length n in G_l. From the above lemma, there exists a balanced cycle (also a circuit) of length n in G_{l+1}.
We now show the claim for $2^l < n < 2^{l+1}$ where n is even. Since $2 \leqslant 2^{l+1} - n < 2^l$, there exists a balanced n-cycle in G_{l+1}, say H. Then the graph $G_{l+1} - H$ contains n edges. Assume that $G_{l+1} - H = H_1 \cup H_2 \cup \cdots \cup H_t$, where each H_i for $1 \leqslant i \leqslant t$ is a component of $G_{l+1} - H$. Since both G_{l+1} and H are balanced and Eulerian, $G_{l+1} - H$ must be balanced and each H_i must be Eulerian.

If $G_{l+1} - H$ is connected, we are done, since it is a circuit in G_{l+1} of length n. Otherwise, since G_{l+1} is connected, there exists an edge e in H which connects vertices in two different components of $G_{l+1} - H$, say H_1 and H_2. Let the edge e go from $v_1 \in V(H_1)$ to $v_2 \in V(H_2)$. Without loss of generality, assume e is red. Now, let e_1 be an edge from v_1 to u_1 in H_1, and let e_2 be an edge from u_2 to v_2 in H_2. Then the last $l-1$ bits of v_1 and u_2 are the first $l-1$ bits of v_2 and u_1. Hence, there must be an edge e' from u_2 to u_1 in H. We deduce from the above lemma that since e is red, e_1 must be blue, e_2 must be red, and e' must be blue.

Consider the new subgraph obtained by replacing e_1 and e_2 with e and e'. One checks easily that the degrees of all vertices, as well as the number of red and blue edges, are preserved. Moreover, the number of connected components has been reduced. Thus, we end up with a balanced subgraph with n edges and fewer components, all of which are Eulerian. By repeating this process, we will eventually end up with a balanced subgraph of n edges which is a circuit. Finally, for $n = 2^{l+1}$, all of the edges of the graph form a circuit since G_{l+1} is Eulerian. The result follows. □