De Bruijn Sequences

Drew Armstrong
armstron@math.umn.edu

September 28, 2006

What is special about the following cyclic binary word (cycle of 0's and 1's)?

$$\begin{array}{cccc} & 0 & 0 & \\ 1 & & & 0 \\ 0 & & & 1 \\ & 1 & 1 & \end{array}$$

As we travel around the cycle (either clockwise or counterclockwise), we will encounter each of the $2^3 = 8$ three-digit patterns 000, 001, 010, 011, 100, 101, 110, 111 *exactly once*. In a sense, this cycle is a very efficient encoding of 24 digits of information into only 8 digits. The formal name for this kind of pattern is a *De Bruijn sequence*.

Definition 1. A De Bruijn sequence of rank n on an alphabet of size k is a cyclic word in which each of the k^n words of length n appears exactly once as we travel around the cycle.

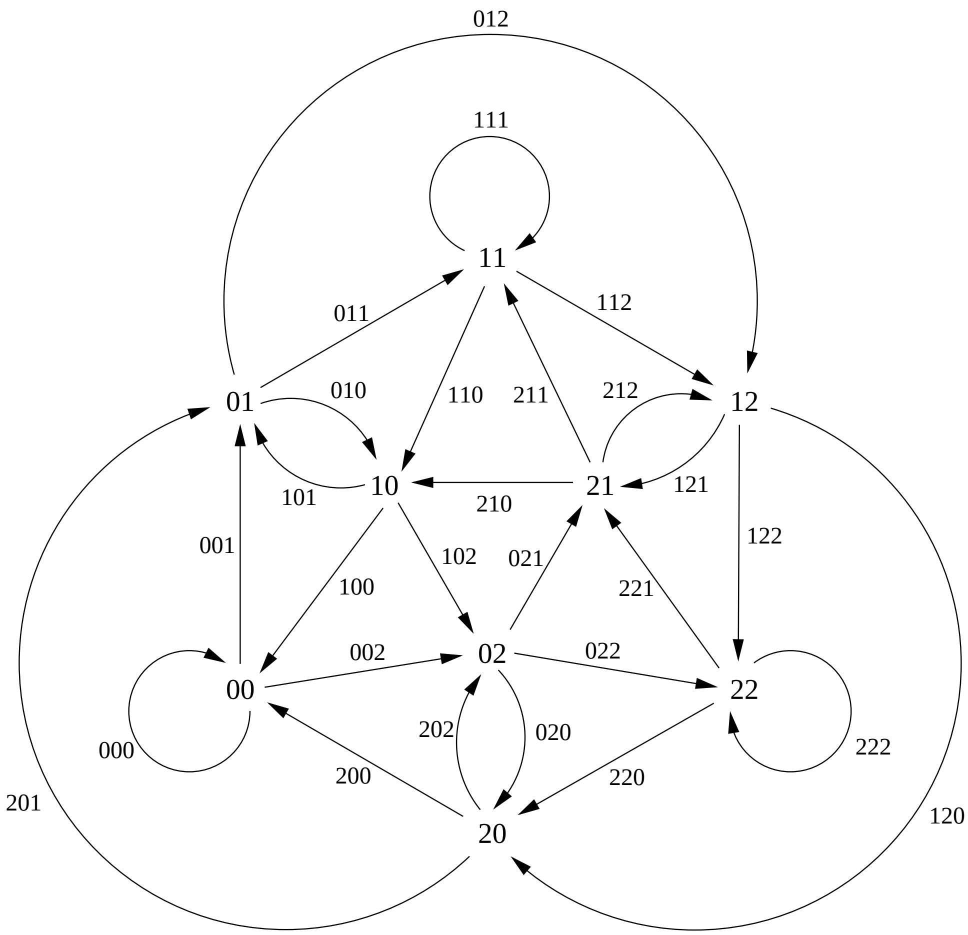
For example, the above cycle is a De Bruijn sequence with $n = 3$ and $k = 2$ (the alphabet is $\{0, 1\}$). De Bruijn sequences have applications to computers (electronic memory and coding theory), crime (efficiently cracking a combination lock), and of course magic tricks (as we saw in class). One thing that is not immediately obvious, however, is whether these things actually *exist*. We have seen one example, but how do we know that there are others, and how do we find them?

Theorem 2. *De Bruijn sequences exist for all n and k.*

To prove this, we will apply the theory of directed graphs. There is a special directed graph we want to consider.

Definition 3. The De Bruijn graph for n and k has one vertex for each of the k^{n-1} words of length $n-1$ from an alphabet of size k. We put a directed edge $w_1 \to w_2$ from word w_1 to word w_2 if the last $n-2$ digits of w_1 agree with the first $n-2$ digits of w_2.

For example, here is the De Bruijn graph with $n = 3$ and $k = 3$.



Notice that the vertices correspond to the $3^2 = 9$ different words of length 2 on the alphabet $\{0, 1, 2\}$ of size 3. Interestingly, notice that each edge can be identified with one of the $3^3 = 27$ words of length 3. Notice, further, that if we travel over one edge, then another, the corresponding 3 digit words overlap nicely. In fact, this is exactly what we want to happen in a De Bruijn sequence! Thus if we can find a circuit in the De Bruijn graph that crosses each edge exactly once, this will construct for us a De Bruijn sequence. We conclude that:

$$\text{De Bruijn sequence} \quad \equiv \quad \text{directed Euler circuit in the De Bruijn graph}$$

You might want to check that the sequence

$$000111222012022110021210102$$

is a De Bruijn sequence (where we wrap the word into a cycle). Indeed, each of the $3^3 = 27$ three digit patterns appears exactly once as we travel around. Can you find the directed Euler circuit in the above graph corresponding to this sequence?

Since we already know a lot about directed Euler circuits, this now allows us to prove that De Bruijn sequences exist.

Proof. Consider the De Bruijn graph on n and k. Every edge coming out of a word corresponds to "what digit comes next?". Since there k digits available, every vertex will have out-degree k. Similarly, every edge coming in to a vertex corresponds to "what digit came before?", and again, there are k digits to choose from, so every vertex has in-degree k. Finally, it is not hard to see that by successively adding digits, we can move from any word to any other word. Thus the De Bruijn graph is strongly connected.

Since the De Bruijn graph is strongly connected and every vertex has in-degree equal to its out-degree, Theorem 6.54 tells us that the graph contains a directed Euler circuit. Thus we have found a De Bruijn sequence. $\square$