Exercise 1 (Exercise 2)**.** Let $d > 1$ be a positive integer. A *d-ary De Bruijn sequence* of degree n is a sequence of length d^n containing every length n sequence in $([d-1]^*)^n$ exactly once as a circular factor.

i) Show that there always exists a d-ary De Bruijn sequence of degree n for any n.
ii) Find the number of d-ary De Bruijn sequences that begin with n zeroes. ⟦Hint: You may want to consult the computation for $d = 2$ given at the end of chapter 5.⟧

Remark. For this exercise, we will shift notation up by one and instead of considering the alphabet set as $[d-1]^* = \{0, 1, \dots, d-1\}$, we will consider $[d] = \{1, \dots, d\}$. We will also call $B(d, n)$ the set of De Bruijn sequences of length d^n with alphabet $[d]$.

Also, we will define the Bruijn graph $G_{d,n}$ as follows:

$$\begin{cases} V = [d]^n = \{ (s_1, \dots, s_n) : \ \forall i (s_i \in [d]) \}, \\ E = \{ ((s_1, \dots, s_n), (t_1, \dots, t_n)) : \ \forall i (1 \leqslant i \leqslant n-1 \Rightarrow t_i = s_{i+1}) \wedge t_n \in [d] \}. \end{cases}$$

That is, the edge set is formed by pairs of strings of the form $(s_1, s_2, \dots, s_n)$ and $(s_2, s_3, \dots, s_n, t)$ where $t \in [d]$. We are shifting all indices in our string and adding a new admissible character.

Answer

i) First, let us prove that $G_{d,n}$ is Eulerian. Consider any vertex $v = (s_1, \dots, s_n) \in G_{d,n}$, it holds that $d_{\text{out}}(v) = d_{\text{in}}(v) = d$.

First, consider the edges out of v: any out-neighbor of v is a vertex $(s_2, \dots, s_n, t)$ with $t \in [d]$. Since there are d options for the character t, v sends an edge to each one of them.

Similarly, every edge which connects to v comes from a vertex of the form $(t, s_1, \dots, s_{n-1})$. There are d vertices of that form in $G_{d,n}$. We conclude that $d_{\text{out}}(v) = d_{\text{in}}(v) = d$.

The De-Bruijn graph is also strongly connected: any two vertices u, v can be reached from one another after deleting and appending sufficient characters. It follows that $G_{d,n}$ is Eulerian for any d and any n.

Take an Eulerian cycle in a De Bruijn graph $G_{d,n-1}$. Such cycle traverses all the edges of our graph, and given that $d_{\text{out}}(v) = d$ for all v, it holds that

$$|E(G_{d,n-1})| = d \cdot |G_{d,n-1}| = d \cdot d^{n-1} = d^n.$$

Labeling the edges by the character it appends to each vertex, we get a string of length d^n which contains all possible substrings of length n in d characters. A d-ary de Bruijn sequence is minimal with respect to this property so it must hold that the sequence generated is a De Bruijn sequence.

Exercise 2 (Exercise 4)**.** The adjacency matrix of a directed graph D is A such that $a_{ij} = [\![(i,j) \in E]\!]$.

i) Show that the $(i,j)^{\text{th}}$ entry of A^k is the number of directed paths of length k from v_i to v_j in D.
ii) Verify that the following equality holds, where we consider both sides as formal power series in x with coefficients in the ring $M_n(\mathbb{Q})$:

$$(I - xA)^{-1} = 1 + xA + x^2A^2 + x^3A^3 + \dots$$

iii) Using the previous part along with the explicit formula for the inverse of a matrix (in terms of cofactors), show that the generating function for the number of paths $p_{i,j}(n)$ of length n from v_i to v_j is

$$\sum_{n=0}^{\infty} p_{i,j}(n)x^n = \frac{(-1)^{i+j}\det\left[(I - xA)^{(j,i)}\right]}{\det(I - xA)},$$

where $A^{(i,j)}$ is the $(i,j)^{\text{th}}$ minor of A.

This result is called the **transfer-matrix method**, as it gives a method of proving a sequence has a rational generating function, by showing that the sequence counts paths in a certain directed graph.

iv) Let b_n be the number of sequences of length $n+1$ with entries from $[3]$ that start with 1, end with 3, and do not contain the subsequences 22 or 23. Find a closed formula for the generating function of b_n using the transfer-matrix method, by constructing a directed graph in which certain paths are counted by b_n. You may use a computer to calculate the determinants, but you must write out the directed graph and the corresponding matrices.

Answer

i) Suppose D is a directed graph with n vertices, we will proceed by induction and use A^2 as a base case. The $(i,j)^{\text{th}}$ entry of A^2 is given by

$$\sum_{k=1}^{n} a_{ik}a_{kj} = a_{i1}a_{1j} + a_{i2}a_{2j} + \cdots + a_{in}a_{nj},$$

and every term $a_{ik}a_{kj}$ counts the number of edges from v_i to v_k times the number of edges from v_k to v_j. But a length 2 path from v_i to v_j can go through any other vertex $u \in D$. Since each path must be different, we sum each possibility to get the complete number.

Suppose now that the $(i,j)^{\text{th}}$ entry of A^m is the number of paths of length m from v_i to v_j for $m \leqslant k-1$. Now

$$A_k = (A^{k-1})A \Rightarrow (A^k)_{ij} = \sum_{\ell=1}^{n}(A^{k-1})_{il}A_{\ell j}.$$

We can thus decompose the $(i,j)^{\text{th}}$ entry of A^k into a sum of terms of the form $(A^{k-1})_{il}A_{\ell j}$. By induction hypothesis $(A^{k-1})_{il}$ is the number of paths from v_i to v_ℓ and adding the edge which could go from v_ℓ to v_j we get a new path. However v_ℓ can be any vertex in D, so summing the possibilities we get the complete number of length k walks.

Thus we conclude that the $(i,j)^{\text{th}}$ entry of A^k is the number of paths of length k from v_i to v_j in D.

ii) We need to prove that the inverse of $I - xA$ is the matrix $1 + xA + x^2A^2 + x^3A^3 + \ldots$ and to do that we will multiply them:

$$\begin{aligned}&(I - xA)(1 + xA + x^2A^2 + x^3A^3 + \ldots)\\ =&(1 + xA + x^2A^2 + x^3A^3 + \ldots) - (xA - x^2A^2 - x^3A^3 - x^4A^4 - \ldots) = I.\end{aligned}$$

We do not worry about convergence issues because this is a formal power series.

iii) The $(i,j)^{\text{th}}$ entry of $\sum_{n=0}^{\infty} A^n x^n$ is $\sum_{n=0}^{\infty} p_{i,j}(n)x^n$ and recalling that this matrix is actually $(I - xA)^{-1}$, we must find the entries of this matrix. By the inverse formula for a matrix we get

$$(I - xA)^{-1} = \frac{1}{\det(I - xA)}\operatorname{adj}(I - xA) = \frac{1}{\det(I - xA)}\operatorname{cof}(I - xA)^{\mathsf{T}}.$$

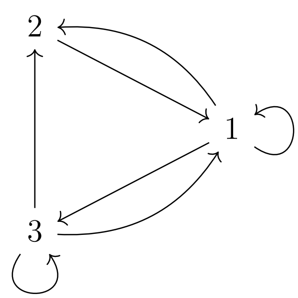
The $(i,j)^{\text{th}}$ entry of the cofactors matrix is $(-1)^{i+j}\det\big[(I - xA)^{(i,j)}\big]$, the determinant of the minor matrix obtained by deleting the i^{th} row and j^{th}

column. Thus the $(i, j)^{\text{th}}$ of the transpose is obtained by switching the indices.

We conclude that the $(i, j)^{\text{th}}$ entry of $\sum_{n=0}^{\infty} A^n x^n$ is

$$\left(\frac{1}{\det(I - xA)}\operatorname{cof}(I - xA)^{\mathsf{T}}\right)_{i,j} = \frac{1}{\det(I - xA)}(-1)^{i+j}\det\big[(I - xA)^{(j,i)}\big].$$

iv) Consider the following graph which encodes the construction of our string:



Finding the amount of strings of length $n + 1$ beginning with 1, ending in 3 is the same as counting walks from vertex 1 to 3 of length n. This means that $b_n = p_{1,3}(n)$ and we know that that sequence's generating function is

$$\frac{1}{\det(I - xA)}(-1)^{1+3}\det\big[(I - xA)^{(3,1)}\big].$$

To find a closed form we start by noting that the adjacency matrix of our graph is

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \Rightarrow I - xA = \begin{pmatrix} 1 - x & -x & -x \\ -x & 1 & 0 \\ -x & -x & 1 - x \end{pmatrix}.$$

We expand by minors using the third row so that we can calculate the minor determinant at once:

$$\begin{aligned}
&\det(I - xA) \\
=&(-x)\det\begin{pmatrix} -x & -x \\ 1 & 0 \end{pmatrix} - (-x)\det\begin{pmatrix} 1 - x & -x \\ -x & 0 \end{pmatrix} + (1 - x)\det\begin{pmatrix} 1 - x & -x \\ -x & 1 \end{pmatrix} \\
=&(-x)(0 - (-x)) + (x)(0 - x^2) + (1 - x)[(1 - x) - x^2] \\
=&1 - 2x - x^2
\end{aligned}$$

We conclude that the generating function is $\frac{-x}{1-2x-x^2}$.