Math 530 Spring 2022, Lecture 20: Trees

website: `https://www.cip.ifi.lmu.de/~grinberg/t/22s`

1. Trees and arborescences (cont'd)

1.1. de Bruijn sequences

Here is a more intricate application of what we have learned about arborescences.

A little puzzle first: What is special about the periodic sequence

$$||:\ 0000\ 1111\ 0110\ 0101\ :|| \qquad ?$$

(This is an infinite sequence of 0's and 1's; the spaces between some of them are only for readability. The $||:$ and $:||$ symbols are "repeat signs" – they mean that everything that stands between them should be repeated over and over. So the sequence above is 0000 1111 0110 0101 0000 1111 ....)

One nice property of this sequence is that if you slide a "length-4 window" along it, you get all 16 possible bitstrings of length 4 depending on the position of the window, and they don't repeat until you move 16 steps to the right. Just see:

$$\boxed{0000}\,11110110010100001111\ldots$$
$$0\,\boxed{0001}\,1110110010100001111\ldots$$
$$00\,\boxed{0011}\,110110010100001111\ldots$$
$$000\,\boxed{0111}\,10110010100001111\ldots$$
$$0000\,\boxed{1111}\,0110010100001111\ldots$$
$$00001\,\boxed{1110}\,110010100001111\ldots$$
$$000011\,\boxed{1101}\,10010100001111\ldots$$
$$0000111\,\boxed{1011}\,0010100001111\ldots$$
$$00001111\,\boxed{0110}\,010100001111\ldots$$
$$000011110\,\boxed{1100}\,10100001111\ldots$$
$$0000111101\,\boxed{1001}\,0100001111\ldots$$
$$00001111011\,\boxed{0010}\,100001111\ldots$$
$$000011110110\,\boxed{0101}\,00001111\ldots$$
$$0000111101100\,\boxed{1010}\,0001111\ldots$$
$$00001111011001\,\boxed{0100}\,001111\ldots$$
$$000011110110010\,\boxed{1000}\,01111\ldots$$

This is nice and somewhat similar to Gray codes. Recall: in a Gray code, you run through all bitstrings of a given size in such a way that only a single bit is changed at each step. Here, instead, as you slide the window along the sequence, at each step, the first bit is removed and a new bit is inserted at the end.

Can we find such nice sequences for any window length, not just 4 ?

Here is an answer for window length 3, for instance:

$$|| : \ 000\ 111\ 01\ \ : || \ .$$

What about higher window length?

Moreover, we can try to do this with other alphabets. For instance, instead of bits, here is a similar sequence for the alphabet $\{0, 1, 2\}$ (that is, we use the numbers $0, 1, 2$ instead of 0 and 1) and window length 2:

$$|| : 00\ 11\ 22\ 02\ 1\ \ : || \ .$$

What about the general case? Let us give it a name:

Definition 1.1.1. Let n and k be two positive integers, and let K be a k-element set.

A **de Bruijn sequence** of order n on K means a k^n-tuple $(c_0, c_1, \ldots, c_{k^n-1})$ of elements of K such that

(A) for each n-tuple $(a_1, a_2, \ldots, a_n) \in K^n$ of elements of K, there is a **unique** $r \in \{0, 1, \ldots, k^n - 1\}$ such that

$$(a_1, a_2, \ldots, a_n) = (c_r, c_{r+1}, \ldots, c_{r+n-1}) \ .$$

Here, the indices under the letter "c" are understood to be periodic modulo k^n; that is, we set $c_{q+k^n} = c_q$ for each $q \in \mathbb{Z}$ (so that $c_{k^n} = c_0$ and $c_{k^n+1} = c_1$ and so on).

For example, for $n = 2$ and $k = 3$ and $K = \{0, 1, 2\}$, the 9-tuple

$$(0, 0, 1, 1, 2, 2, 0, 2, 1)$$

is a de Bruijn sequence of order n on K, because if we label the entries of this 9-tuple as $c_0, c_1, \ldots, c_8$ (and extend the indices periodically, so that $c_9 = c_0$), then we have

$$\begin{aligned}
&(0,0) = (c_0, c_1) ; \qquad &&(0,1) = (c_1, c_2) ; \qquad &&(0,2) = (c_6, c_7) ; \\
&(1,0) = (c_8, c_9) ; \qquad &&(1,1) = (c_2, c_3) ; \qquad &&(1,2) = (c_3, c_4) ; \\
&(2,0) = (c_5, c_6) ; \qquad &&(2,1) = (c_7, c_8) ; \qquad &&(2,2) = (c_4, c_5) .
\end{aligned}$$

It turns out that de Bruijn sequences always exist:

Theorem 1.1.2 (de Bruijn, Sainte-Marie)**.** Let n and k be positive integers. Let K be a k-element set. Then, a de Bruijn sequence of order n on K exists.

Proof. It looks reasonable to approach this using a digraph. For example, we can define a digraph whose vertices are the n-tuples in K^n, and that has an arc from one n-tuple i to another n-tuple j if j can be obtained from i by dropping the first entry and adding a new entry at the end. Then, a de Bruijn sequence (of order n on K) is the same as a Hamiltonian cycle of this digraph.

Unfortunately, we don't have any useful criteria that would show that such a cycle exists. So this idea seems to be a dead end.

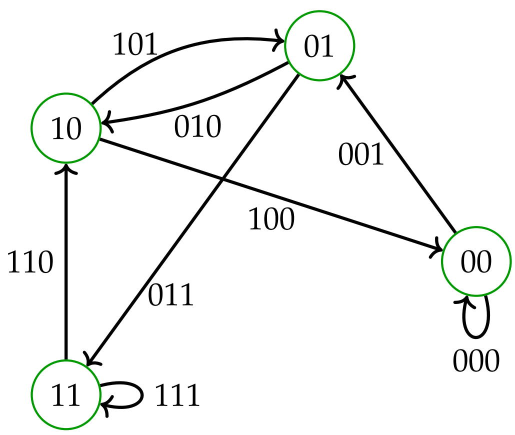
However, let us do something counterintuitive: We try to reinterpret de Bruijn sequences in terms of Eulerian circuits (rather than Hamiltonian cycles), since we have a good criterion for the existence of Eulerian circuits (unlike for that of Hamiltonian cycles)!

We need a different digraph for that. Namely, we let D be the multidigraph $\left(K^{n-1}, K^n, \psi\right)$, where the map $\psi : K^n \to K^{n-1} \times K^{n-1}$ is given by the formula

$$\psi\left(a_1, a_2, \ldots, a_n\right) = \left(\left(a_1, a_2, \ldots, a_{n-1}\right), \ \left(a_2, a_3, \ldots, a_n\right)\right).$$

Thus, the vertices of D are the $(n-1)$-tuples (not the n-tuples!) of elements of K, whereas the arcs are the n-tuples of elements of K, and each such arc $(a_1, a_2, \ldots, a_n)$ has source $(a_1, a_2, \ldots, a_{n-1})$ and target $(a_2, a_3, \ldots, a_n)$. Hence, there is an arc from each $(n-1)$-tuple $i \in K^{n-1}$ to each $(n-1)$-tuple $j \in K^{n-1}$ that is obtained by dropping the first entry of i and adding a new entry at the end. (Be careful: If $n = 1$, then D has only one vertex but n arcs. If this confuses you, just do the $n = 1$ case by hand. For any $n > 1$, there are no parallel arcs in D.)

Example 1.1.3. For example, if $n = 3$ and $k = 2$ and $K = \{0, 1\}$, then D looks as follows (we again write our tuples without commas and without parentheses):



Let us make a few observations about D:

- The multidigraph D is strongly connected.

  [*Proof:* We need to show that for any two vertices i and j of D, there is a walk from i to j. But this is easy: Just insert the entries of j into i one by one, pushing out the entries of i. In other words, using the notation k_p for the p-th entry of any tuple k, we have the walk

$$\begin{aligned} i &= (i_1, i_2, \ldots, i_{n-1}) \\ &\to (i_2, i_3, \ldots, i_{n-1}, j_1) \\ &\to (i_3, i_4, \ldots, i_{n-1}, j_1, j_2) \\ &\to \cdots \\ &\to (i_{n-1}, j_1, j_2, \ldots, j_{n-2}) \\ &\to (j_1, j_2, \ldots, j_{n-1}) = j. \end{aligned}$$

  Note that this walk has length $n-1$, and is the unique walk from i to j that has length $n-1$. Thus, the # of walks from i to j that have length $n-1$ is 1. This will come useful further below.]

- Thus, the multidigraph D is weakly connected (since any strongly connected digraph is weakly connected).

- The multidigraph D is balanced, and in fact each vertex of D has outdegree k and indegree k.

  [*Proof:* Let i be a vertex of D. The arcs with source i are the n-tuples whose first $n-1$ entries form the $(n-1)$-tuple i while the last, n-th entry is an arbitrary element of K. Thus, there are $|K|$ many such arcs. In other words, i has outdegree k. A similar argument shows that i has indegree k. This entails that $\deg^- i = \deg^+ i$. Since this holds for every vertex i, we conclude that D is balanced.]

- The digraph D has a Eulerian circuit.

  [*Proof:* This follows from the directed Euler–Hierholzer theorem (Theorem 1.4.2 in Lecture 10), since D is weakly connected and balanced. Alternatively, we can derive this from the BEST theorem (Theorem 1.1.1 in Lecture 17) as follows: Pick an arbitrary arc a of D, and let r be its source. Then, r is a from-root of D (since D is strongly connected), and thus D has a spanning arborescence rooted from r (by Theorem 1.2.4 in Lecture 16). In other words, using the notations of the BEST theorem (Theorem 1.1.1 in Lecture 17), we have $\tau(D, r) \neq 0$. Moreover, each vertex of D has indegree $k > 0$. Thus, the BEST theorem yields

$$\varepsilon(D, a) = \underbrace{\tau(D, r)}_{\neq 0} \cdot \underbrace{\prod_{u \in V} \left(\deg^- u - 1\right)!}_{\neq 0} \neq 0.$$

  But this shows that D has an Eulerian circuit whose last arc is a.]

So we know that D has a Eulerian circuit $\mathbf{c}$. This Eulerian circuit leads to a de Bruijn sequence as follows:

Let $p_0, p_1, \ldots, p_{k^n-1}$ be the arcs of $\mathbf{c}$ (from first to last). Extend the subscripts periodically modulo k^n (that is, set $p_{q+k^n} = p_q$ for all $q \in \mathbb{N}$). Thus, we obtain an infinite walk[1] with arcs $p_0, p_1, p_2, \ldots$ (since $\mathbf{c}$ is a circuit). In other words, for each $i \in \mathbb{N}$, the target of the arc p_i is the source of the arc p_{i+1}.

In other words, for each $i \in \mathbb{N}$, the last $n-1$ entries of p_i are the first $n-1$ entries of p_{i+1} (since the target of p_i is the tuple consisting of the last $n-1$ entries of p_i, whereas the source of p_{i+1} is the tuple consisting of the first $n-1$ entries of p_{i+1}). Therefore, for each $i \in \mathbb{N}$ and each $j \in \{2, 3, \ldots, n\}$, we have

$$\begin{aligned} &(\text{the } j\text{-th entry of } p_i) \\ &= (\text{the } (j-1)\text{-st entry of } p_{i+1}) . \end{aligned} \tag{1}$$

Now, for each $i \in \mathbb{N}$, we let x_i denote the first entry of the n-tuple p_i. Then, $x_{q+k^n} = x_q$ for all $q \in \mathbb{N}$ (since $p_{q+k^n} = p_q$ for all $q \in \mathbb{N}$). In other words, the sequence $(x_0, x_1, x_2, \ldots)$ repeats itself every k^n terms. Note that the k^n-tuple $(x_0, x_1, \ldots, x_{k^n-1})$ consists of the first entries of the arcs $p_0, p_1, \ldots, p_{k^n-1}$ of $\mathbf{c}$ (by the definition of x_i).

For each $i \in \mathbb{N}$ and each $s \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned} &(\text{the } s\text{-th entry of } p_i) \\ &= (\text{the } (s-1)\text{-st entry of } p_{i+1}) && (\text{by (1)}) \\ &= (\text{the } (s-2)\text{-nd entry of } p_{i+2}) && (\text{by (1)}) \\ &= (\text{the } (s-3)\text{-rd entry of } p_{i+3}) && (\text{by (1)}) \\ &= \cdots \\ &= (\text{the 1-st entry of } p_{i+s-1}) \\ &= x_{i+s-1} && (\text{since } x_{i+s-1} \text{ was defined as the first entry of } p_{i+s-1}) . \end{aligned}$$

In other words, for each $i \in \mathbb{N}$, the entries of p_i (from first to last) are $x_i, x_{i+1}, \ldots, x_{i+n-1}$. In other words, for each $i \in \mathbb{N}$, we have

$$p_i = (x_i, x_{i+1}, \ldots, x_{i+n-1}) . \tag{2}$$

Now, recall that $\mathbf{c}$ is a Eulerian circuit. Thus, each arc of D appears exactly once among its arcs $p_0, p_1, \ldots, p_{k^n-1}$. In other words, each n-tuple in K^n appears exactly once among $p_0, p_1, \ldots, p_{k^n-1}$ (since the arcs of D are the n-tuples in K^n). In other words, as i ranges from 0 to $k^n - 1$, the n-tuple p_i takes each possible value in K^n exactly once.

In view of (2), we can rewrite this as follows: As i ranges from 0 to $k^n - 1$, the n-tuple $(x_i, x_{i+1}, \ldots, x_{i+n-1})$ takes each possible value in K^n exactly once (since this n-tuple is precisely p_i, as we have shown in the previous paragraph). In other words, for each $(a_1, a_2, \ldots, a_n) \in K^n$, there is a **unique** $r \in \{0, 1, \ldots, k^n - 1\}$ such that $(a_1, a_2, \ldots, a_n) = (x_r, x_{r+1}, \ldots, x_{r+n-1})$.

[1]We have never formally defined infinite walks, but it should be fairly clear what they are.

Hence, the k^n-tuple $(x_0, x_1, \ldots, x_{k^n-1})$ is a de Bruijn sequence of order n on K. This shows that a de Bruijn sequence exists. Theorem 1.1.2 is thus proven.

Example 1.1.4. For $n = 3$ and $k = 2$ and $K = \{0, 1\}$, one possible Eulerian circuit $\mathbf{c}$ of D is

(00, **001**, 01, **010**, 10, **101**, 01, **011**, 11, **111**, 11, **110**, 10, **100**, 00)

(where we have written the arcs in bold for readability). The first entries of the arcs of this circuit form the sequence 0010111, which is indeed a de Bruijn sequence of order 3 on $\{0, 1\}$. Any 3 consecutive entries of this sequence (extended periodically to the infinite sequence $|| : 0010111 : ||$) form the respective arc of $\mathbf{c}$.

□

Theorem 1.1.2 is merely the starting point of a theory. Several specific de Bruijn sequences are known, many of them having peculiar properties. See [Freder82] for a survey of various such sequences[2] (note that they are called "full length nonlinear shift register sequences" in this survey). (My favorite is the one obtained by concatenating all Lyndon words whose length divides K in lexicographically increasing order. See [Moreno04] for the details of that construction.)

There are also several variations on de Bruijn sequences. For some of them, see [ChDiGr92]. (Note that some of the open questions in that paper are still unsolved.) A variation that recently became quite popular is the notion of a "universal cycle for permutations" – a string that contains all "permutations" (more precisely, n-tuples of distinct elements of K) as factors. See [EngVat18] for some recent progress on minimizing the length of such a string, including a cameo by a notorious hacker known as 4chan. (This is no longer really about Eulerian circuits, since some amount of duplication cannot be avoided in these strings.)

Let us move in a different direction. Having proved the existence of de Bruijn sequences in Theorem 1.1.2, let us try to count them!

Question. Let n and k be two positive integers. Let K be a k-element set. How many de Bruijn sequences of order n on K are there?

To solve this, it makes sense to apply the BEST theorem to the digraph D we have constructed above. Alas, D is not of the form G^{bidir} for some undirected graph G, so we cannot apply the undirected MTT (Matrix-Tree Theorem). However, D is a balanced multidigraph, and for such digraphs, a version of the undirected MTT still holds:

[2]Some of these sequences (the "prefer-one" and "prefer-opposite" generators) are just disguised implementations of the algorithm for finding a Eulerian circuit implicit in our proof of the BEST theorem.

Theorem 1.1.5 (balanced Matrix-Tree Theorem)**.** Let $D = (V, A, \psi)$ be a balanced multidigraph. Assume that $V = \{1, 2, \ldots, n\}$ for some positive integer n.

Let L be the Laplacian of D. Then:

(a) For any vertex r of D, we have

$$(\text{\# of spanning arborescences of } D \text{ rooted to } r) = \det\left(L_{\sim r, \sim r}\right).$$

Moreover, this number does not depend on r.

(b) Let t be an indeterminate. Expand the determinant $\det(tI_n + L)$ (here, I_n denotes the $n \times n$ identity matrix) as a polynomial in t:

$$\det(tI_n + L) = c_n t^n + c_{n-1} t^{n-1} + \cdots + c_1 t^1 + c_0 t^0,$$

where $c_0, c_1, \ldots, c_n$ are numbers. (Note that this is the characteristic polynomial of L up to substituting $-t$ for t and multiplying by a power of -1. Some of its coefficients are $c_n = 1$ and $c_{n-1} = \operatorname{Tr} L$ and $c_0 = \det L$.) Then, for any vertex r of D, we have

$$(\text{\# of spanning arborescences of } D \text{ rooted to } r) = \frac{1}{n} c_1.$$

(c) Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of L, listed in such a way that $\lambda_n = 0$. Then, for any vertex r of D, we have

$$(\text{\# of spanning arborescences of } D \text{ rooted to } r) = \frac{1}{n} \cdot \lambda_1 \lambda_2 \cdots \lambda_{n-1}.$$

(d) Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of L, listed in such a way that $\lambda_n = 0$. If all vertices of D have outdegree > 0, then

$$(\text{\# of Eulerian circuits of } D) = |A| \cdot \frac{1}{n} \cdot \lambda_1 \lambda_2 \cdots \lambda_{n-1} \cdot \prod_{u \in V} \left(\deg^+ u - 1\right)!.$$

(If you identify an Eulerian circuit with its cyclic rotations, then you should drop the $|A|$ factor on the right hand side.)

Proof. **(a)** The equality comes from the MTT (Theorem 1.1.2 in Lecture 19). It remains to prove that the # of spanning arborescences of D rooted to r does not depend on r. But this is Corollary 1.1.6 in Lecture 17.

(b) follows from **(a)** as in the undirected graph case (Lecture 19, proof of Theorem 1.2.1 **(b)**).[3]

[3] In more detail: Just as we proved in Lecture 19 (for the undirected case), we have $c_1 =$

(c) follows from **(b)** as in the undirected graph case (Lecture 19, proof of Theorem 1.2.1 **(c)**).

(d) Assume that all vertices of D have outdegree > 0. Then,

$$\begin{aligned}&(\text{\# of Eulerian circuits of } D)\\&=\sum_{a\in A}(\text{\# of Eulerian circuits of } D \text{ whose first arc is } a)\,.\end{aligned}$$

However, if $a \in A$ is any arc, and if r is the source of a, then

$$\begin{aligned}&(\text{\# of Eulerian circuits of } D \text{ whose first arc is } a)\\&=(\text{\# of spanning arborescences of } D \text{ rooted to } r)\cdot\prod_{u\in V}\left(\deg^+ u-1\right)!\\&\qquad\qquad(\text{by the BEST' theorem (Theorem 1.3.5 in Lecture 16)})\\&=\frac{1}{n}\cdot\lambda_1\lambda_2\cdots\lambda_{n-1}\cdot\prod_{u\in V}\left(\deg^+ u-1\right)!\qquad(\text{by part } \textbf{(c)})\,.\end{aligned}$$

Hence,

$$\begin{aligned}&(\text{\# of Eulerian circuits of } D)\\&=\sum_{a\in A}\underbrace{(\text{\# of Eulerian circuits of } D \text{ whose first arc is } a)}_{=\frac{1}{n}\cdot\lambda_1\lambda_2\cdots\lambda_{n-1}\cdot\prod_{u\in V}\left(\deg^+ u-1\right)!}\\&=\sum_{a\in A}\frac{1}{n}\cdot\lambda_1\lambda_2\cdots\lambda_{n-1}\cdot\prod_{u\in V}\left(\deg^+ u-1\right)!\\&=|A|\cdot\frac{1}{n}\cdot\lambda_1\lambda_2\cdots\lambda_{n-1}\cdot\prod_{u\in V}\left(\deg^+ u-1\right)!.\end{aligned}$$

This proves part **(d)**. □

$\sum_{r=1}^{n}\det\left(L_{\sim r,\sim r}\right)$. However, part **(a)** shows that the number $\det\left(L_{\sim r,\sim r}\right)$ does not depend on r. Thus, the sum $\sum_{r=1}^{n}\det\left(L_{\sim r,\sim r}\right)$ consists of n equal addends, which can be written as $\det\left(L_{\sim r,\sim r}\right)$ for any vertex r of D. Therefore, this sum can be rewritten as $n\cdot\det\left(L_{\sim r,\sim r}\right)$ for any vertex r of D. Hence, the equality $c_1=\sum_{r=1}^{n}\det\left(L_{\sim r,\sim r}\right)$ can be rewritten as $c_1 = n\cdot\det\left(L_{\sim r,\sim r}\right)$ for any vertex r of D. Therefore, $\det\left(L_{\sim r,\sim r}\right)=\frac{1}{n}c_1$ for any vertex r of D. Since part **(a)** yields

$$(\text{\# of spanning arborescences of } D \text{ rooted to } r)=\det\left(L_{\sim r,\sim r}\right),$$

we can rewrite this equality as

$$(\text{\# of spanning arborescences of } D \text{ rooted to } r)=\frac{1}{n}c_1.$$

Now, let's try to solve our question – i.e., let's count the de Bruijn sequences of order n on K.

Recall the digraph D from our above proof of Theorem 1.1.2. We constructed a de Bruijn sequence of order n on K by finding an Eulerian circuit of D. This actually works both ways: The map

$$\{\text{Eulerian circuits of } D\} \to \{\text{de Bruijn sequences of order } n \text{ on } K\}, \\ \mathbf{c} \mapsto (\text{the sequence of first entries of the arcs of } \mathbf{c})$$

is a bijection (make sure you understand why!). Hence, by the bijection principle, we have

$$(\text{\# of de Bruijn sequences of order } n \text{ on } K) \\ = (\text{\# of Eulerian circuits of } D). \qquad (3)$$

By Theorem 1.1.5 **(d)**, however, we have

$$(\text{\# of Eulerian circuits of } D) \\ = |K^n| \cdot \frac{1}{k^{n-1}} \cdot \lambda_1 \lambda_2 \cdots \lambda_{k^{n-1}-1} \cdot \prod_{u \in K^{n-1}} \left(\deg^+ u - 1\right)!, \qquad (4)$$

where $\lambda_1, \lambda_2, \ldots, \lambda_{k^{n-1}}$ are the eigenvalues of the Laplacian L of D, indexed in such a way that $\lambda_{k^{n-1}} = 0$. (Note that the digraph $D = \left(K^{n-1}, K^n, \psi\right)$ has k^{n-1} vertices, not n vertices, so the "n" in Theorem 1.1.5 is k^{n-1} here.)

As we know, each vertex of D has outdegree k. That is, we have $\deg^+ u = k$ for each $u \in K^{n-1}$. Thus,

$$\prod_{u \in K^{n-1}} \left(\deg^+ u - 1\right)! = \prod_{u \in K^{n-1}} (k-1)! = ((k-1)!)^{k^{n-1}}.$$

Also,

$$|K^n| \cdot \frac{1}{k^{n-1}} = k^n \cdot \frac{1}{k^{n-1}} = k.$$

It remains to find $\lambda_1 \lambda_2 \cdots \lambda_{k^{n-1}-1}$. What are the eigenvalues of L ?

The Laplacian L of our digraph D is a $k^{n-1} \times k^{n-1}$-matrix whose rows and columns are indexed by $(n-1)$-tuples in K^{n-1}. Strictly speaking, we should relabel the vertices of D as $1, 2, \ldots, k^{n-1}$ here, in order to have a "proper matrix" with a well-defined order on its rows and columns. But let's not do this; instead, I trust you can do the relabeling yourself, or just use the more general notion of matrices that allows for the rows and the columns to be indexed by arbitrary things (see `https://mathoverflow.net/questions/317105` for details).

Let C be the adjacency matrix of the digraph D; this is the $k^{n-1} \times k^{n-1}$-matrix (again with rows and columns indexed by $(n-1)$-tuples in K^{n-1}) whose (i,j)-th entry is the # of arcs with source i and target j. In particular, the trace of C is thus the # of loops of D. It is easy to see that the loops of D are precisely the arcs of the form $(x, x, \ldots, x) \in K^n$ for $x \in K$; thus, D has exactly k loops. Hence, the trace of C is k.

Recall the definition of the Laplacian matrix L. We can restate it as follows:

$$L = \Delta - C, \tag{5}$$

where Δ is the diagonal matrix whose diagonal entries are the outdegrees of the vertices of D. Since each vertex of D has outdegree k, the latter diagonal matrix Δ is simply $k \cdot I$, where I is the identity matrix (of the appropriate size). Hence, (5) can be rewritten as

$$L = k \cdot I - C.$$

Thus, if $\gamma_1, \gamma_2, \ldots, \gamma_{k^{n-1}}$ are the eigenvalues of C, then $k - \gamma_1, k - \gamma_2, \ldots, k - \gamma_{k^{n-1}}$ are the eigenvalues of L. Computing the former will thus help us find the latter.

Furthermore, let J be the $k^{n-1} \times k^{n-1}$-matrix (again with rows and columns indexed by $(n-1)$-tuples in K^{n-1}) whose all entries are 1. It is easy to see that the eigenvalues of J are

$$\underbrace{0, 0, \ldots, 0}_{k^{n-1}-1 \text{ many zeroes}}, k^{n-1}.$$

(The easiest way to see this is by noticing that J has rank 1 and trace k^{n-1}. [4])

Now, here is something really underhanded: We observe that

$$C^{n-1} = J.$$

[*Proof*: We need to show that all entries of the matrix C^{n-1} are 1. So let i and j be two vertices of D. We must then show that the (i, j)-th entry of C^{n-1} is 1.

Recall the combinatorial interpretation of the powers of an adjacency matrix (homework set #4, exercise 4 **(a)**): For any $\ell \in \mathbb{N}$, the (i, j)-th entry of C^ℓ is the # of walks from i to j (in D) that have length ℓ. Thus, in particular, the (i, j)-th entry of C^{n-1} is the # of walks from i to j (in D) that have length $n-1$. But this number is actually 1, as we have already shown in our above proof of Theorem 1.1.2. This completes the proof of $C^{n-1} = J$.]

How does this help us compute the eigenvalues of C ? Well, let $\gamma_1, \gamma_2, \ldots, \gamma_{k^{n-1}}$ be the eigenvalues of C. Then, for any $\ell \in \mathbb{N}$, the eigenvalues of C^ℓ are $\gamma_1^\ell, \gamma_2^\ell, \ldots, \gamma_{k^{n-1}}^\ell$ (this is a fact that holds for any square matrix, and is probably easiest to prove using the Jordan canonical form or triangularization). Hence, in

[4] Here are the details: The matrix J has rank 1 (since all its rows are the same); thus, all but one of its eigenvalues are 0. It remains to show that the remaining eigenvalue is k^{n-1}. However, it is known that the sum of the eigenvalues of a square matrix equals its trace. Thus, if all but one of the eigenvalues of a square matrix are 0, then the remaining eigenvalue equals its trace. Applying this to our matrix J, we see that its remaining eigenvalue equals its trace, which is k^{n-1}.

particular, $\gamma_1^{n-1}, \gamma_2^{n-1}, \ldots, \gamma_{k^{n-1}}^{n-1}$ are the eigenvalues of $C^{n-1} = J$; but we know that the latter eigenvalues are $\underbrace{0, 0, \ldots, 0}_{k^{n-1}-1 \text{ many zeroes}}, k^{n-1}$. Hence, all but one of the k^{n-1} numbers $\gamma_1^{n-1}, \gamma_2^{n-1}, \ldots, \gamma_{k^{n-1}}^{n-1}$ equal 0. Thus, all but one of the k^{n-1} numbers $\gamma_1, \gamma_2, \ldots, \gamma_{k^{n-1}}$ equal 0 (we don't know what the remaining number is, since $(n-1)$-st roots are not uniquely determined in $\mathbb{C}$). In other words, all but one of the eigenvalues of C equal 0. The remaining eigenvalue must thus be the trace of C (because the sum of the eigenvalues of a square matrix is known to be the trace of that matrix), and therefore equal k (since we know that the trace of C is k).

So we have shown that the eigenvalues of C are $\underbrace{0, 0, \ldots, 0}_{k^{n-1}-1 \text{ many zeroes}}, k$. Thus, the eigenvalues of L are

$$\underbrace{k-0, \ k-0, \ \ldots, \ k-0}_{k^{n-1}-1 \text{ many } (k-0)\text{'s}}, \ k-k$$

(because if $\gamma_1, \gamma_2, \ldots, \gamma_{k^{n-1}}$ are the eigenvalues of C, then $k-\gamma_1, k-\gamma_2, \ldots, k-\gamma_{k^{n-1}}$ are the eigenvalues of L). In other words, the eigenvalues of L are

$$\underbrace{k, k, \ldots, k}_{k^{n-1}-1 \text{ many } k\text{'s}}, 0.$$

Hence, the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{k^{n-1}-1}$ in (4) all equal k. Thus, (4) simplifies to

$$\begin{aligned}
&(\text{\# of Eulerian circuits of } D) \\
&= \underbrace{|K^n| \cdot \frac{1}{k^{n-1}}}_{= k^n \cdot \frac{1}{k^{n-1}} = k} \cdot \underbrace{kk \cdots k}_{k^{n-1}-1 \text{ factors}} \cdot \underbrace{\prod_{u \in K^{n-1}} \left(\deg^+ u - 1\right)!}_{= ((k-1)!)^{k^{n-1}}} \\
&= \underbrace{k \cdot \underbrace{kk \cdots k}_{k^{n-1}-1 \text{ factors}}}_{= k^{k^{n-1}}} \cdot ((k-1)!)^{k^{n-1}} = k^{k^{n-1}} \cdot ((k-1)!)^{k^{n-1}} \\
&= \left(\underbrace{k \cdot (k-1)!}_{= k!}\right)^{k^{n-1}} = k!^{k^{n-1}}.
\end{aligned}$$

In view of this, we can rewrite (3) as

$$(\text{\# of de Bruijn sequences of order } n \text{ on } K) = k!^{k^{n-1}}.$$

Thus, we have proved the following:

Theorem 1.1.6. Let n and k be positive integers. Let K be a k-element set. Then,

$$(\text{\# of de Bruijn sequences of order } n \text{ on } K) = k!^{k^{n-1}}.$$

What a nice (and huge) answer!

Our above proof of Theorem 1.1.6 is essentially taken from [Stanle18, Chapter 10].

We note that a combinatorial proof of Theorem 1.1.6 (avoiding any use of linear algebra) has been recently given in [BidKis02].

1.2. On the left nullspace of the Laplacian

Much more can be said about the Laplacian of a digraph. The study of matrices associated to a graph or digraph is known as **spectral graph theory**; I'd say the Laplacian is probably the most prominent of these matrices (even though the adjacency matrix is somewhat easier to define). The original form of the matrix-tree theorem (undirected, I believe) was found by Gustav Kirchhoff in his study of electricity; the effective resistance between two nodes of an electrical network is a ratio of spanning-tree counts and thus can be computed using the Laplacian (see, e.g., [Vos16, §2 and §3]). To be more precise, this relies on a "weighted count" of spanning trees, which is more general than the counting we have done so far; we will learn about it next time.

Another application of Laplacians is to drawing graphs: see "spectral layout".

Let me mention one more result about Laplacians of digraphs that answers a rather natural question you might already have asked yourself. Recall that the Laplacian L of a digraph D always satisfies $Le = 0$, where $e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$. Thus, the vector e belongs to the right nullspace (= right kernel) of L. It is not hard to see that if D has a to-root and we are working over a characteristic-0 field, then e spans this nullspace, i.e., there are no vectors in that nullspace other than scalar multiples of e. (This is actually an "if and only if".) What about the left nullspace of L ? Can we explicitly find a nonzero vector f with $fL = 0$? The answer is positive:

Theorem 1.2.1 (harmonic vector theorem for Laplacians)**.** Let $D = (V, A, \psi)$ be a multidigraph, where $V = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. For each $r \in V$, let $\tau(D, r)$ be the # of spanning arborescences of D rooted to r. Let f be the row vector $(\tau(D,1), \tau(D,2), \ldots, \tau(D,n))$. Then, $fL = 0$.

Proof. See homework set #7 exercise 1 **(b)**, or [Sahi14, Theorem 1]. □

Theorem 1.2.1 (or, more precisely, its weighted version, which we will see in the next lecture) can be used to explicitly compute the steady state of a Markov chain (see [KrGrWi10]); a similar interpretation, but in economical terms (emergence of money in a barter economy), appears in [Sahi14, §1].

References

[BidKis02] Hoda Bidkhori, Shaunak Kishore, *Counting the spanning trees of a directed line graph*, arXiv:0910.3442v1.
Later published under the title *A Bijective Proof of a Theorem of Knuth*, Combinatorics, Probability & Computing **20**(1), pp. 11–25, 2011.

[ChDiGr92] Fan Chung, Persi Diaconis, Ron Graham, *Universal cycles for combinatorial structures*, Discrete Mathematics **110** (1992), pp. 43–59, `http://www.math.ucsd.edu/~fan/wp/universalcycle.pdf`.

[EngVat18] Michael Engen, Vincent Vatter, *Containing all permutations*, arXiv:1810.08252v4, Amer. Math. Monthly **128** (2021), pp. 4–24. `https://arxiv.org/abs/1810.08252v4`

[Freder82] Harold Fredricksen, *A Survey of Full Length Nonlinear Shift Register Cycle Algorithms*, SIAM Review **24**, No. 2, April 1982, pp. 195–221. `https://doi.org/10.1137/1024041`

[KrGrWi10] Alex Kruckman, Amy Greenwald, John Wicks, *An elementary proof of the Markov chain tree theorem*, 6 August 2010. `https://math.berkeley.edu/~kruckman/MCTT.pdf`

[Moreno04] Eduardo Moreno, *On the theorem of Fredricksen and Maiorana about de Bruijn sequences*, Advances in Applied Mathematics **33** (2004), pp. 413–415. `https://doi.org/10.1016/j.aam.2003.10.002`

[Sahi14] Siddhartha Sahi, *Harmonic vectors and matrix tree theorems*, Journal of Combinatorics **5**, Number 2, pp. 195–202, 2014.

[Stanle18] Richard P. Stanley, *Algebraic Combinatorics: Walks, Trees, Tableaux, and More*, 2nd edition, Springer 2018.
See `https://math.mit.edu/~rstan/algcomb/errata2.pdf` for errata.

[Vos16] Vaya Sapobi Samui Vos, *Methods for determining the effective resistance*, Master's thesis, 20 December 2016. `https://www.universiteitleiden.nl/binaries/content/assets/science/mi/scripties/master/vos_vaya_master.pdf`