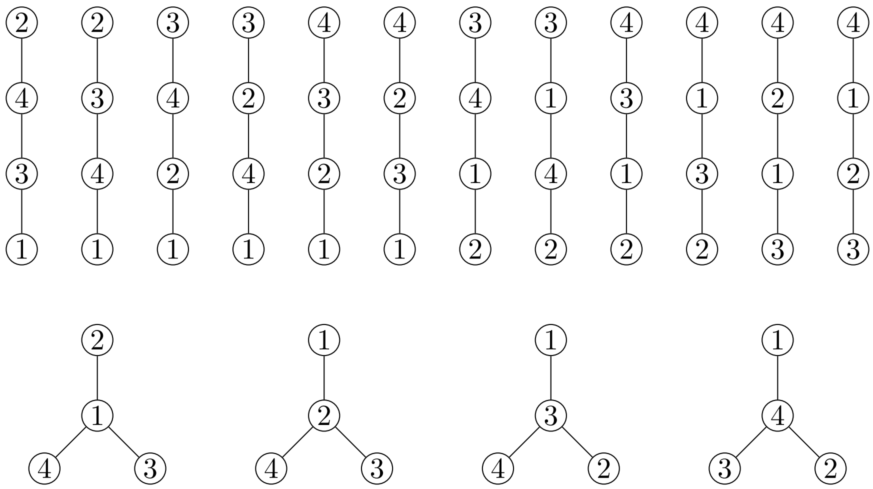
Joyal's proof of Cayley's formula

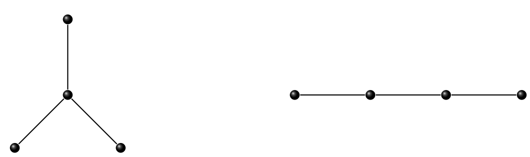
There are 16 trees with vertex set $\underline{4} = \{1, 2, 3, 4\}$:



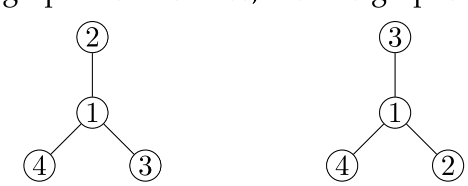
An *isomorphism* between two graphs $G = (V, E)$ and $G' = (V', E')$ is a bijection of their vertices that induces a bijection of their edges. In detail, this means there is a bijection $f\colon V \to V'$ such that for each $\{u, v\} \in E$, we have $\{f(u), f(v)\} \in E'$ and the resulting function

$$\begin{aligned} E &\mapsto E' \\ \{u, v\} &\mapsto \{f(u), f(v)\} \end{aligned}$$

is bijective. If we say $G \sim G'$ whenever there is an isomorphism between a pair of graphs G and G', it is easy to check that $\sim$ is an equivalence relation. An *unlabeled* graph is an equivalence class under this relation. For instance, there are only two equivalence classes of trees on 4 vertices. These appear below:



Suppose you did not have the list of 16 trees with vertex set $\underline{4}$. How could you go about finding them? One way to proceed is to first create the list of unlabeled graphs on 4 vertices. Since 4 is a small number, it is easy to see there are only two: the star graph and the path graph pictured above. The task then is to find the different ways of labeling the vertices of these two graphs with the numbers $1, 2, 3, 4$. For the star graph, there is one vertex with degree 3 (i.e., with 3 adjacent edges). The other vertices are "symmetric" in the sense that permuting labels on these vertices does not change the (labeled) graph. For instance, the two graphs below are the same:



We know the above two graphs are equal (not just isomorphic) since they have the same vertices and the same edges. Thus, there are 4 labeled star graphs with vertex set $\underline{4}$, each arising from a choice of a central vertex. Now consider the path graph. At first, you might think there are $4! = 24$ ways of labeling the vertices of the unlabeled path graph—one for each permutation of $\underline{4}$. However, note that the following two graphs are the same:

$$1 - 2 - 3 - 4 \qquad\qquad 4 - 3 - 2 - 1$$

Why? Again: these two graphs have the same set of vertices and the same set of edges.

PROBLEM 1. Arguing from unlabeled trees as above, determine the number of (labeled) trees on five vertices.

THEOREM 2 (Cayley's formula). *The number of trees on n vertices is n^{n-2}.*

We will give Joyal's proof of this theorem. Let T_n denote the number of trees on n vertices. Then Cayley's formula can be restated as

$$n^2 T_n = n^n.$$

To prove Cayley's formula, Joyal creates a bijection between two sets, one of size $n^2 T_n$, and the other of size n^n. The latter sets is easy to describe: it is $\underline{n}^{\underline{n}}$, the set of all functions of $\{1, 2, \ldots, n\}$ to itself. (Recall that $\underline{n}^{\underline{n}}$ has n^n elements since for each of the n elements in its domain, there are n choices for an assigned value in the codomain.) The former set—the one with $n^2 T_n$ elements—consists of what Joyal called *vertebrates*.

A *vertebrate* on n vertices is a tree T with vertex set $\underline{n}$ and a choice of an ordered pair (t, h) consisting of vertices t and h of T (where $t = h$ is allowed). The vertex t is called the *tail* of the vertebrate, and h is the *head*. The number of vertebrates on n vertices is $n^2 T_n$ since there are T_n choices for T, and for each each of these, there are n possibilities for each of t and h. Let $\mathcal{V}_n$ denote the set of all vertebrates on n vertices. To prove the validity of Cayley's formula, it suffices to create a bijection:

$$J \colon \mathcal{V}_n \to \underline{n}^{\underline{n}}.$$

Why the word *vertebrate*? Given the tree T with tail t and head h, there is a unique path from t to h, which we imagine to be the *spine* of some creature. The edges not on this path are the creature's *appendages*. When drawing a vertebrate, we will highlight its spine, which will be used in the construction of our bijection. See Figure 1.

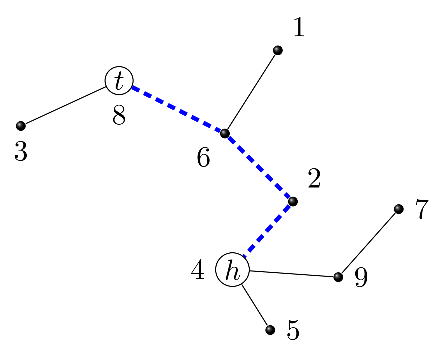


FIGURE 1. Vertebrate on 9 vertices with tail vertex 8 and head vertex 4.

Bijection. We first describe the mapping $J \colon \mathcal{V}_n \to \underline{n}^{\underline{n}}$ by example using the vertebrate T of Figure 1 with tail $t = 8$ and head $h = 4$. To find the corresponding mapping $f \colon \underline{9} \to \underline{9}$, start with the values of f along the spine. The spine vertices, in their tail-to-head order along the spine, are $8, 6, 2, 4$. List these numbers in two rows. The top row is the natural ordering of these numbers, and the bottom is their "spine-ordering":

(∗)

i	2	4	6	8
$f(i)$	8	6	2	4

Then start to define f be sending each number in the top row to its corresponding number below it, as shown in the table.

It remains to assign values to the vertices along the appendages. To do this, direct the edges incident on appendage vertices so that they point towards the spine, as shown in Figure 2. If the

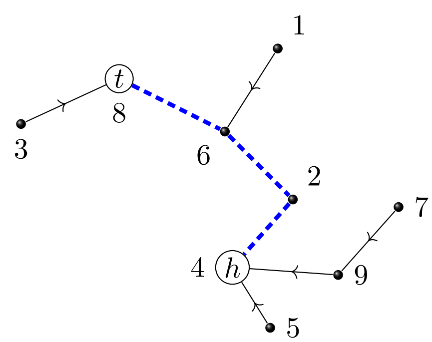


FIGURE 2. Directing addendage edges towards the spine.

integer i is an appendage vertex, let $f(i)$ be vertex adjacent to i on the path leading to the spine. Thus, for instance, $f(7) = 9$ and $f(9) = 4$. Filling in these values defines f on the rest of its domain:

i	1	2	3	4	5	6	7	8	9
$f(i)$	6	8	8	6	4	2	9	4	4

FIGURE 3. Function corresponding to the vertebrate in Figure 1. Spinal vertices are in blue.

We now let $J(T, (t, h)) = f \in \underline{9}^{\underline{9}}$.

Definition of the mapping $J \colon \mathcal{V}_n \to \underline{n}^{\underline{n}}$

Let T, (t, h) be a vertebrate. Our task is to define $f := J(T, (t, h)) \in \underline{n}^{\underline{n}}$.

(1) First define f for the vertices along the spine. Say the spinal vertices are $v_1, \ldots, v_k$, in order along the spine from tail to head. Let $a_1 < \cdots < a_k$ be the permutation of these spinal vertices into their natural ordering as integers. Then define $f(a_i) = v_i$ for $i = 1, \ldots, k$. Thus, f permutes the spinal vertices.
(2) Next, direct all edges incident on appendage (non-spinal) vertices so that they point towards the spine. If i is an appendage vertex, define $f(i) = j$ if j is the vertex adjacent to i along the directed path from i to the spine.

i	1	2	3	4	5	6	7	8	9
$f(i)$	3	5	8	7	5	1	4	1	2

FIGURE 4. A function $f\colon \underline{9} \to \underline{9}$.

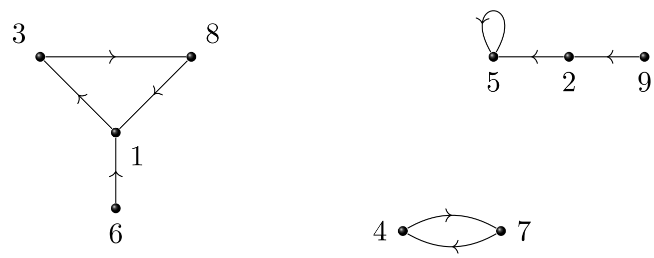


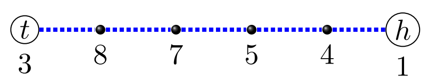
FIGURE 5. Directed graph associated with the function in Figure 4.

The inverse mapping. We now describe the inverse of the mapping $\mathcal{V}_n \to \underline{n}^{\underline{n}}$, starting with an example. Consider the function given by the table in Figure 4. We are hunting for a corresponding vertebrate. To begin, associate a directed graph to f with vertex set $\underline{9}$ and with edges $(i, f(i))$ for $i \in \underline{9}$. This graph is pictured in Figure 5. Each of the components of the resulting graph has a unique cycle.[1]

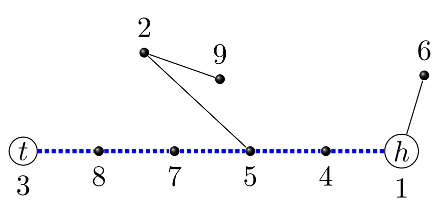
The cycles are $1 \to 3 \to 8 \to 1$, and $5 \to 5$, and $4 \to 7 \to 4$. Consider the function restricted to the vertices in these cycles:

i	1	3	4	5	7	8
$f(i)$	3	8	7	5	4	1

The list of vertices in the bottom row of the table defines the spine, from tail to head, of the vertebrate we are seeking:



Finally, for each appendage vertex i, we attach the edge $\{i, f(i)\}$. These are undirected versions of the edges appearing in Figure 5:



PROBLEM 3. Apply the mapping $J\colon \mathcal{V}_n \to \underline{n}^{\underline{n}}$ to the above vertebrate to see that you recover the original function f.

[1]It is generally true that each component of the directed graph associated to a function $f\colon \underline{n} \to \underline{n}$ will have a unique cycle. To see this, consider a component H of the graph. Each vertex of H has a single out-going edge, and thus, the number of vertices and edges of H are equal. One characterization of a tree is a connected graph with one fewer edge than vertex. Thus, H is connected but not a tree. So H must have a cycle. Removing one edge from the cycle leaves a tree, and it is a general fact that adding an edge to a tree produces a unique cycle.

The inverse mapping $J^{-1}\colon \underline{n}^{\underline{n}} \to \mathcal{V}_n$

Let $f\colon \underline{n} \to \underline{n}$. Our task is to find a vertebrate $T, (t, h)$ such that $J(T, (t, h)) = f$.

(1) Create a directed graph G with vertex set $\underline{n}$ and directed edges $(i, f(i))$ for $i \in \underline{n}$.
(2) Let $i_1 < i_2 < \cdots < i_k$ (with the natural ordering as integers) be the vertices appearing in cycles in G. Define the spine of the vertebrate $T, (t, h)$ we are constructing to be the path graph with vertices $f(i_1), \ldots, f(i_k)$. Thus, $t := f(i_1)$ and $h := f(i_k)$.
(3) Finally, for each vertex i of G that is not in a cycle, add the (undirected) edge $\{i, f(i)\}$ to T.

Example. Here is a final example illustrating the special case where $t = h$. Start with the vertebrate $T, (t, h)$ in Figure 6

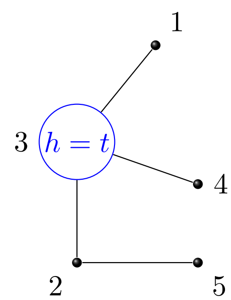


FIGURE 6. Vertebrate in for which $t = h$.

To define the corresponding function $f := J(T, (t, h))$, we first define f along the spine as in the table ($*$). This tells us that $f(3) = 3$. We then direct the appendage edges (in this case, all of the edges) towards the spine and read off the rest of the function:

i	1	2	3	4	5
$f(i)$	3	3	3	3	2

To reverse the process, first draw the directed graph G corresponding to f as in Figure 7.

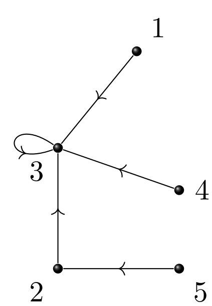


FIGURE 7. Graph for the function corresponding to the vertebrate in Figure 6.

There is only one connected component to G, and it has a single cycle: a loop at 3. This means that the corresponding vertebrate has a spine with $t = h = 3$. Adding the appendage edges $\{i, f(i)\}$ for $i \neq 3$ then recovers the original vertebrate.

PROBLEM 4. Choose a vertebrate with vertex set $\underline{n}$ for some n, and then determine its corresponding function f under our bijection. Next choose some function $f\colon \underline{n} \to \underline{n}$, and determine its corresponding vertebrate.