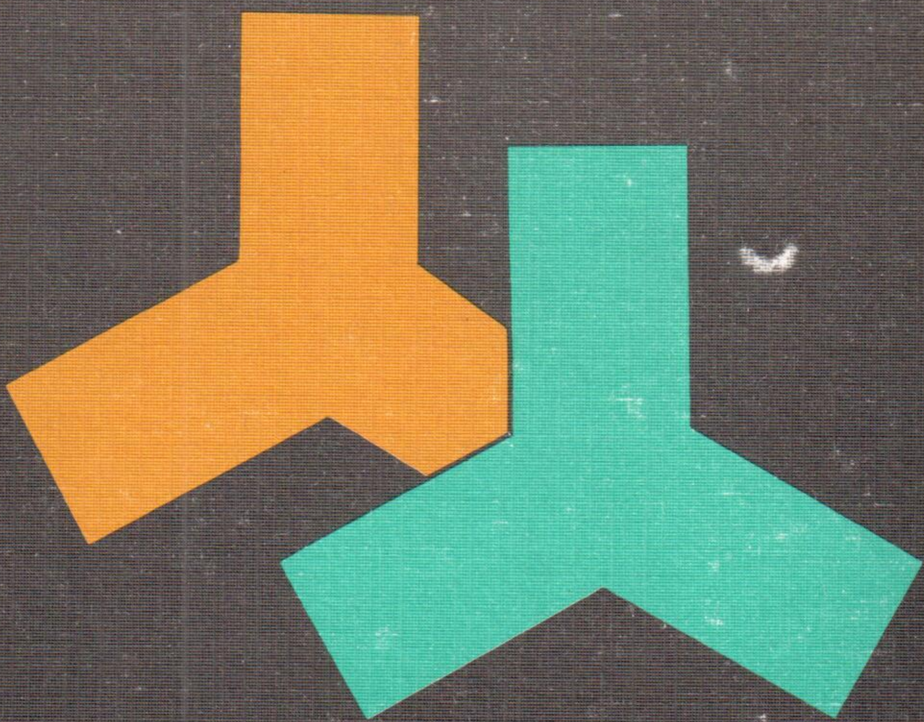
Combinatorial Problems and Exercises

by László Lovász

2nd edition



Akadémiai Kiadó · Budapest

Combi… Problems and Exercises

5. Let T be a tree on points $v_1, \ldots, v_n$. Delete the endpoint having the least index and write down the index of its neighbor. Repeat this procedure with the resulting tree, until a tree with only one point remains. This associates a sequence of $n-1$ numbers with T, called the *Prüfer code* of T. Prove that

(a) the Prüfer code of T uniquely characterizes T,

(b) given any sequence $(a_1, \ldots, a_{n-1})$ such that $1 \le a_i \le n$, $a_{n-1} = n$, there is a (unique) tree with this Prüfer code.

(c) Deduce the Cayley formula.

5. (a) Let $b_1, \ldots, b_{n-1}$ be the indices of removed points. Let us see how to determine b_i, if we know the Prüfer code.

b_i is obviously different from $b_1, \ldots, b_{i-1}, a_i$. Also, $b_i \ne a_j$ for $j > i$: for b_i is removed and it cannot be the neighbor of an endpoint at a later step. Conversely, if $k \notin \{b_1, \ldots, b_{i-1}, a_i, \ldots, a_{n-1}\}$, then v_k is an endpoint of $T - \{v_{b_1}, \ldots, v_{b_{i-1}}\}$: otherwise it would be a neighbor of a point removed at a later step. Thus,

$$b_i = \min\{k : k \notin \{b_1, \ldots, b_{i-1}, a_i, \ldots, a_{n-1}\}\}. \tag{1}$$

Thus, the Prüfer code uniquely determines the numbers b_i. Since (v_{a_i}, v_{b_i}) are the edges of T, the Prüfer code uniquely determines T.

(b) Let $(a_1, \ldots, a_{n-1})$ be any sequence of integers with $1 \le a_i \le n$, $a_{n-1} = n$. Define b_i recursively by (1) and join v_{a_i} to v_{b_i} for $i = 1, \ldots, n-1$. We claim that the resulting graph T is a tree with Prüfer code $(a_1, \ldots, a_{n-1})$. Both assertions will follow, if we show that v_{b_i} is an endpoint of the graph $T_i = T - \{v_{b_1}, \ldots, v_{b_{i-1}}\}$ and no point with smaller index is endpoint. We have

$$v_{a_i} \in V(T_i)$$

because $a_i \ne b_1, \ldots, b_{i-1}$ by (1). Thus, v_{b_i} has a neighbor in T_i. v_{b_i} cannot be adjacent to any other point of T_i: for suppose that (v_{a_j}, v_{b_j}) is another edge of T_i adjacent to b_i, then $j > i$ as $v_{b_j} \in V(T_i)$ and either $b_i = b_j$ or $b_i = a_j$, which both contradict (1). Hence v_{b_i} is an endpoint of T_i. This proves that T and all T_i's are trees.

Now suppose that T_i has an endpoint v_k with $k < b_i$. Since k did not come into consideration when defining b_i by (1), it follows that either $k = b_\nu$, $\nu < i$ or $k = a_j$, $j \ge i$. Now the first possibility does not occur because $v_{b_\nu} \in V(T_i)$, so $k = a_j$, $j \ge i$. Since $a_{n-1} = n \ge b_i > k$, we have $j \le n-2$. By the argument above, v_{b_j} is an endpoint of T_j and its neighbor is v_{a_j}. But $v_{a_j} = v_k$ is an endpoint of T_i, and therefore it must be an endpoint of T_j too. Hence $V(T_j) = \{v_{a_j}, a_{b_j}\}$ which is impossible as T_j has $n - j + 1 > 3$ points.

The Cayley formula follows immediately: The number of sequences $(a_1, \ldots, a_{n-1})$ with $1 \le a_i \le n$, $a_{n-1} = n$ is, obviously, n^{n-2}. [A. Prüfer, *Archiv f. Math. u. Phys.* **27** (1918) 142–144.]

[From L. Lovasz
Combinatorial Problems
and Exercises,
Akademiai Kiadó and
Elsevier, 1993]