TwoFx's blog

[Tutorial] Burnside's lemma (with example)

By TwoFx, history, 6 years ago,

Hello Codeforces!

The recent Educational Round 52 featured problem 1065E - Side Transmutations, which is a counting problem that most people solved using some ad-hoc observations. However, like many problems of the form "count number of objects up to some symmetry/transformation", it can be solved using a neat little theorem called "Burnside's Lemma", which I would like to introduce to those who do not know it yet using the recent task as an example (even though it is completely overkill for this problem).

This is my first tutorial, so any feedback is greatly appreciated.

Prerequisites

Definition of a group

Group actions

In these kinds of counting problems, you always have some set of objects X you are working with and some set of operations G you can use to transform one object into another. Formally, we have a function $\bullet : G \times X \to X$ so that $g \bullet x$ is the object we get when applying operation g to object x.

If the operations form a group, and the operations fulfill two rules, $e_G \bullet x = x$ and $g_1 \bullet (g_2 \bullet x)) = (g_1 \cdot g_2) \bullet x$, then we call $\bullet$ a "group action". The rules basically say that the action is compatible with the group structure and are usually easy to verify.

Orbits and fixed points

For some $x \in X$, we can look at its orbit $G \bullet x := \{g \bullet x \mid g \in G\}$: The set of all objects which can be attained by doing some transformation on x. We immediately see why this is a useful notion for us: Orbits are exactly the classes of objects which we consider equal, because they only differ by some symmetry or other transformation. Our goal is to count the number of orbits, which is sometimes denoted by $|X / G|$.

For some $g \in G$, we may also look at its fixed points $X^g := \{x \in X \mid g \bullet x = x\}$: The set of objects which are not changed by the operation g. We'll see in a minute why we need this notation.

Our task

In our task, the objects are all strings of length n over the alphabet A. The operations are slightly more difficult. We cannot just take all single moves as described in the task as our operations, because they do not form a group: It is not possible to combine two moves into another (single) move.

However, note that the order in which moves are made is not significant. Also, doing a move twice is exactly the same as not doing it at all. Therefore, an operation can be seen as a subset of all allowed moves, applied in some arbitrary order. So we choose G as the set of subsets (the powerset!) of $B = \{b_1, ..., b_m\}$. We also need a group structure on G, and the correct way to get it is this: $P \cdot Q := (P \cup Q) \setminus (P \cap Q)$. Think of this as the element-wise XOR of the two sets (remember, P and Q are subsets of B). We need XOR here because of our observation that doing a move twice is the same as not doing it at all.

Let's have a look at the orbits and fixed points of this group action. The orbits are exactly the sets of strings which can be transformed into each other by some sequence of moves. Perfect, that is exactly what we wanted.

Now consider some element $g = \{a_1 > a_2 > \ldots > a_k\} \in G$. How many letters are swapped by this transformation? Move a_1 swaps a_1 pairs of letters. Then, move a_2 swaps a_2 of these pairs back, so after a_1 and a_2, $a_1 - a_2$ pairs of letters have been swapped. Then, move a_3 swaps a_3 more pairs, which have all been swapped back in the last step, so $a_1 - a_2 + a_3$ pairs have moved, and so on. So after all moves,

the number of pairs of letters which have (effectively) been swapped is $l := \sum_{i=1}^{k} (-1)^{i-1} a_i$. So how does a string look that is a fixed point of this transformation? For each pair of swapped letters, we can only choose one, because the other has to be the same for the operation not to change the string. So we end up with $|A|^{n-l}$ choices for the string and conclude $|X^g| = |A|^{n-l}$.

Burnside's lemma

Now that all preparations are done, Burnside's lemma gives a straight-up formula for the answer of the problem:

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|$$

That's it! Once you understand how many fixed points there are for each operation, you can use the formula to get the number of orbits.

This is nice in multiple ways. First, since you only look at the fixed points for a single operation at a time, finding the fixed points is often not very hard, as we saw above. Second, applying Burnside's lemma generally requires almost no ad-hoc thinking: You simply define the group and the action, look at the fixed points, and arrive at the solution without the need for any clever insight, but just by following the plan. Because of these reasons, many counting problems become significantly easier when applying Burnside's lemma.

Solving our problem

Since we have already looked at the fixed points of our action, we can just plug it into the formula to get the solution:

$$|X/G| = \frac{1}{2^m} \sum_{\{a_1 > \ldots > a_k\} \subseteq B} |A|^{n - \sum_{i=1}^{k} (-1)^{i-1} a_i}$$

Implementing the formula in the straightforward way using bitmasks to enumerate subsets gives correct answers. Most problems where Burnside's lemma is needed are solved at this point.

Making it fast

However in this problem, we have $m = 10^5$, but the bitmask implementation takes time exponential in m, which is clearly too slow. We need to find some way to evaluate the formula faster. First move $|A|^n$ out of the sum to get

$$|X/G| = \frac{|A|^n}{2^m} \sum_{\{a_1 > \ldots > a_k\} \subseteq B} |A|^{-\sum_{i=1}^{k} (-1)^{i-1} a_i}$$

Let's have a look at how this sum looks for some small values of m:

$$m = 1 : |X/G| = \frac{|A|^n}{2^m}(1 + |A|^{-b_1})$$
$$m = 2 : |X/G| = \frac{|A|^n}{2^m}((1 + |A|^{-b_1}) + |A|^{-b_2}(1 + |A|^{b_1}))$$
$$m = 3 : |X/G| = \frac{|A|^n}{2^m}(((1 + |A|^{-b_1}) + |A|^{-b_2}(1 + |A|^{b_1})) + |A|^{-b_3}((1 + |A|^{b_1}) + |A|^{b_2}(1 + |A|^{-b_1})))$$

We can use this pattern to evaluate the sum in linear time. Let $p_0 = r_0 = 1$. p_i contains the above sum without the leading factor considering only $\{b_1, ..., b_i\}$. r_i is the same, but for each summand we use its multiplicative inverse. From the above pattern we deduce:

$$p_i = p_{i-1} + |A|^{-b_i} r_{i-1}$$
$$r_i = r_{i-1} + |A|^{b_i} p_{i-1}$$

and finally, $|X/G| = \frac{|A|^n}{2^m} p_m$, which we can easily calculate in linear time to get AC.

Practice

You can test your newly acquired skills by solving 101873B - Buildings or any of the problems listed here.

math, tutorial, burnside lemma, group theory, combinatorics