

Combinatorial proofs, in my opinion, are difficult because they seem like there are no formulaic approaches to converge to the solution. Solutions to combinatorial proofs often look as though they can only be solved with the specific flash of insight that the problem creator had in mind. This is my attempt at making the process a little bit more methodical. While there are, I am sure, many ways to approach combinatorial proofs, this is just how I do it.

To do this, I want to go through an example problem, not from the perspective of one who knows the answer and is trying to explain the solution, but rather from the perspective of a student attempting to solve the problem as given. Often times, solutions start with the combinatorial story, but very rarely will I look at a problem and start with a story in mind. That being said, there are a couple identities and heuristics to learn, so we'll identify them along the way. The only prior knowledge I expect of the reader is a basic understanding of counting techniques such as $\binom{n}{k}$ and branching (e.g. 10 coin flips yields 2^{10} possible outcomes).

The problem of interest is this:

$$\sum_{k=1}^n k \cdot (n-k) \cdot \binom{n}{k}^2 = n^2 \binom{2n-2}{n-2}$$

You can find this problem at the HKN CS 70 exam repository from Sp 14, MT3.

Before going forward, it's important to know how to interpret some of the terms. Our goal is to translate everything in the equation into something that we can interpret logically as part of a counting process.

1. $\sum_{k=1}^n$ usually indicates splitting our items in two with k items on the left and $n-k$ items on the right. Each way of splitting is mutually exclusive and going through all n ways is exhaustive, so we're able to capture all possibilities simply by the summation.
2. $n \cdot \textit{whatever}$ or $k \cdot \textit{whatever}$ usually indicates selecting a single item among n or k items respectively. If there are n items, then there are n ways to select a single item. Another way to think about this is $\binom{n}{1} = n$.
3. $\binom{n}{k}$ is just "choose k items out of n items" as usual.
4. $(\cdot)^k$ means to apply the (\cdot) procedure k times. If n means pick 1 item out of n items, then n^3 might mean pick 1 item out of n items 3 times.

There are other types of terms you may see in combinatorial proof problems, but we will examine the above 4 as they are the only ones we really need to solve our problem.

To start, we pick a side and try to interpret one of them. It seems like the RHS is easier (summations are hard), so we start there.

$$n^2 \binom{2n-2}{n-2}$$

Before we examine the expression above, I want to look at a more general expression first. One thing that appears relatively frequently in combinatorial proofs is $n \cdot \binom{n-1}{k-1}$, or something similar to that form. As mentioned above, $n \cdot \textit{something}$ means selecting a single item out of n items. $\binom{n-1}{k-1}$ means we're choosing $k-1$ items out of $n-1$ items. So we can interpret this to say that we had n items to start and we selected 1 out of the n items. That means we have $n-1$ items left, and we have $k-1$ left to choose.

Does that mean that $\binom{n}{k} = n \cdot \binom{n-1}{k-1}$? Not quite. In a sense, the LHS is completely unordered while the RHS maintains some order because we're selecting 1 out of n *first*. Turns out the actual equality is $\binom{n}{k} = \frac{n}{k} \cdot \binom{n-1}{k-1}$. We'll come back to this later.

A good practice in CS 70 is to take the problem that you're given and make it look like something you've

encountered before. So we try to fit the RHS to the "out of n items, choose 1, then choose $k-1$ out of $n-1$ " method that is described by $n \cdot \binom{n-1}{k-1}$.

But we run into a problem. Since we're choosing from $n-2$ items, it should've been that we picked 2 first. Therefore, we should expect to see something like $\binom{n}{2}$ or $n \cdot (n-1)$ or something, not n^2 . In fact, n^2 suggests that we selected 1 item out of n items twice, or we had 2 groups of n items, among which we selected 1 item from each group.

Splitting our items into 2 groups is a relatively common tactic in combinatorial proofs, as mentioned above in the note about summations. If we separated our $2n$ items into two groups of n and selected 1 from each of them, then the n^2 makes sense, especially considering that there is a summation on the LHS which separates things into 2 groups anyway. We've interpreted all the terms on the RHS, so we can start working on the LHS.

Perhaps it's worthwhile to take a moment to talk about combinatorial proofs in general. There are 2 things to keep in mind about all this.

1. This may be obvious, but we're evaluating 2 different methods of counting to solve the same problem. For example, $\binom{n}{k} = \binom{n}{n-k}$ because choosing k out of n items to take with me is the same as choosing $n-k$ out of n items to leave behind. Usually, one side is more intuitive than the other (we suspect that by the end of this, we'll confirm that the RHS of our problem is easier than the LHS), so we'll need to get creative.
2. Sometimes, what we're doing on one side can raise false suspicion. For example, we saw that $\binom{n}{k} \neq n \cdot \binom{n-1}{k-1}$. Recall that this is not an equality is because the LHS is orderless while the RHS is not. So we might be tempted to falsely interpret the RHS as orderless or second-guess our judgment because it seems like we're choosing, which usually is orderless. Don't do this! If the LHS is ordered, then RHS should also be ordered. If the LHS is orderless, only then should RHS be orderless. Don't assume one side should be orderless when the otherside is ordered. Of course, if this is not a source of confusion for you, feel free to ignore this aside.

Back to the problem of interest. This time, we'll look at the LHS.

$$\sum_{k=1}^n k \cdot (n-k) \cdot \binom{n}{k}^2$$

First, there is a term that we are now familiar with: $k \cdot \binom{n}{k}$. We know that means we chose k out of n , then we chose 1 item out of the k that we chose. Now, what remains for us to resolve are the summation and the remaining $(n-k) \cdot \binom{n}{k}$. We'll examine the latter first.

Hopefully, the next trick should be pretty easy to spot: $\binom{n}{k} = \binom{n}{n-k}$. This means we have $(n-k) \cdot \binom{n}{n-k}$, which we know to interpret as choosing $n-k$ items out of n , then selecting 1 item out of the $n-k$ that we chose.

By now, the intuition behind the proof should be coming together. We separate our $2n$ items into two groups of n . We choose k from the first n and $n-k$ from the second n , and then from each of those selections, we pick 1 item. This is similar to the story we came up with for the RHS. For the RHS, we separated our $2n$ into two groups of n , then selected 1 from each. We then took the remaining $2n-2$ and chose $n-2$ from it.

The last bit we need to interpret is the summation. If we are splitting our $2n$ items into two groups of n and n , then there are several ways that we can choose n items. Let's call the first group A and the second group B. We can choose 0 from A and n from B, or 1 from A and $n-1$ from B, or 2 from A and $n-2$ from B, or n from A and 0 from B, and so on. In fact, we can do this for $k=0$ to $k=n$. For our problem, we're forcing at least 1 to be selected from each group because we have those k and $n-k$ terms that tell us to pick one item out of the k and $n-k$ items that we choose from group A and B respectively. But since we have $n-k$ in our term, when $k=n$, the addend goes to 0, so we're fine including $k=n$. We may as well be summing from $k=1$ to $k=n-1$. It's this last portion that connects our LHS to the $\binom{2n-2}{n-2}$ on the RHS.

Putting it all together, we interpret the RHS as such: We split the $2n$ items into 2 groups of n , then pick 1 out of each. From the $2n - 2$ remaining, we choose $n - 2$. The final "story" that we could converge at might be to split $2n$ people into 2 teams, then select a team captain from each. From the remaining $2n - 2$ people, we select $n - 2$ to be some group G .

We interpret the LHS as such: For any given k between 1 and n , we split the $2n$ items into 2 groups of n and choose k from the first group and $n - k$ from the second group. We then pick 1 from the k selections from the first group and 1 from the $n - k$ selections from the second group. Thus, we have 2 picked out and $n - 2$ from the rest of $2n - 2$. We then do this for every k to account for every k and $n - k$ split between the two groups. The final "story" we could converge at for the LHS might be to split $2n$ people into 2 teams, then select k nominees from the first team to be team captain, then choose 1 out of the k nominees. Similarly, from the second team, we have $n - k$ nominees, then choose 1 out of the $n - k$ nominees. The rest of the nominees form the group G .

And that concludes our proof. Both the heuristics and common tactics are the results of having looked through a lot of combinatorial proofs. The lesson to draw here isn't which rules and identities to write on your cheat sheet, but to use this as a starting point to practice working through a lot more combinatorial proofs. Kudos to you if you actually went through the entire walkthrough and got to this point. Hopefully you'll earn a couple more points on the exam.

As a final side note, What we did with the summation is an application of Vandermonde's Identity. If the part about the summation was confusing, check out the combinatorial proof for Vandermonde's identity in the Wikipedia link. It's relatively simple and won't take more than a few minutes to understand.