PROFESSOR: We're going to spend a couple of minutes talking about the binomial theorem, which is probably familiar to you from high school, and is a nice first illustration of the connection between algebra and computation. So the idea that underlies the connection is illustrated by the distributive law. And I'm purposely writing it in this wacky way-- it's purely symbolic-- where I'm going to multiply three beanies by two neckties. What do I get if I multiply three beanies by two neckties?

Well, applying the distributive law says that we get every possible pairing of a beanie and a necktie multiplied together and you add them all up. So there are three terms here and two terms there. I'm going to wind up with six terms, which you see laid out.

The basic rule is that a product of sums by the distributive law becomes a sum of products. And the sums that you get involves products of all of the terms from each of the components in every possible way.

Let's look at that as it applies to the binomial theorem. So the binomial theorem is interested in the question of let's look at the expression 1 plus x raised to the NTH power. And we know that this will be a polynomial of degree n, so it can be written in the form a constant, c0 plus c1 times x to 1, c2 x to the 2, cn x to the n. And we like to ask, what are the expressions for ck?

So for example, here's a layout of the first four powers of x. Let's say 1 plus x to the fourth is 1 plus 4x plus 6x squared 4x cubed 1x fourth. What's the pattern that underlies those coefficients, 1, 4, 6, 4, 1?

Well, one way to think about it is that if I wrote out 1 plus x to the n fully it is, of course, a product of n occurrences of 1 plus x times 1 plus x. And applying the distributive law, the product of sums equals sum of products rule, but this time there are n products. And what I wind up with is 2 to the n terms that I'm adding up. Each of them is a product of a selection of n items, one from each of the factors. So a term among these 2 to the n terms corresponds to selecting a 1 or an x from each of the n factors.

So if I started off, for example, by selecting a 1, a 1, a 1 from each of the n factors, I'd get the term 1 to the n. If I selected an x, an x, an x, an x from each of the terms, I'd get this last term x to the n. And if I made some arbitrary selection, like I selected an x from the first one, and a 1 from the second one, and a 1 from the third one-- I'm reading this term-- 1 from the fourth,

and so on, and x from the next to the last, and a 1 from the last, I would get this particular term, which is not the next one that would occur in alphabetical order, but it's just an example.

So that's the simple idea. If you multiply out n terms, each of which is a sum of two things, you're going to wind up with 2 to the n terms corresponding to every possible way of selecting one or the other of the components from each of the n products.

So what's the coefficient of x to the k? Well, the coefficient of x to the k is the number of those terms among the 2 to the n in which the power of x is k. That is, in which I selected k x's and willy nilly n minus k ones.

Well, how many ways are there to select k x's among these n terms? And we know the answer to that. It's n choose k. It's all the ways of choosing a subset of k items out of n items. And that's the answer. ck is simply n choose k.

And this is called the binomial formula. So now we know that 1 plus x to the n is n choose 0 plus n choose 1x, n choose 2x squared. n choose kx to the k is the general term. Ending with n choose n to the x to the n.

So this expression, 1 plus x, is called a binomial expression. And the choose numbers, which we've seen previously, is the number of ways to choose, in this case, k out of n elements, are called binomial coefficients. And this is why they're called binomial coefficients.

So if I was going to express it more generally, I didn't need it to be 1. I used 1 plus x just because it was easier to follow the structure of the formula. But if I look at x plus y to the NTH power, again, the coefficients are the same. It's n choose 0, but this time y to the n, n choose 1, xy to the n minus 1. What's happening now is that I'm choosing an x or a y instead of an x or a 1 from each of the terms, so that the xy terms are always going to have a sum of degrees that is equal to n. If I choose k x's, I must have chosen n minus k y's.

And there it is expressed in more concise form using sigma notation. That is the binomial formula.