

PROFESSOR: Understanding proofs includes the ability to spot mistakes in them. And as a test of that skill and also your understanding of induction, let me see if I can put one over on you. I'm going to show you a bogus proof by induction. And I'm going to prove something that's patently absurd. Namely, that all horses have the same color. Say they're all black. So, there's a bunch of black horses with maybe some highlighted brown manes.

I'm going to prove this by induction. And for a start, there's no n mentioned in the theorem, so that's common for various kinds of things that you need to prove by induction. So I have to rephrase it in terms of n . It's going to be by induction on n . The induction hypothesis is going to be that any set consisting of exactly n horses will all have the same color.

Let's proceed to prove this. Now, I'm going to use the base case n equals 1, just because I don't want to distract you with suspicions that the base case n equals 0, that is no horses, is cheating somehow. It would be, in fact, be perfectly legitimate to start with n equals 0 and argue that all the horses in the empty set have the same color, because there's nothing in the empty set. However, let's not get into that.

We'll start with n equals one. And indeed, if you look at any set of one horse, it is the same color of it as itself. And in fact, I've proved the base case n equals 1. Let's keep going.

Now, in the inductive step, I'm allowed to assume that n horses have the same color, where n is any number greater than or equal to 0. Now right here, students complain that that's not fair, because you're already assuming something false and that's absurd. Well, yeah, it is absurd. But that can't be the problem. I'm just allowed to assume an induction hypothesis. All I have to do is prove that n implies n plus 1. Since it's absurd, there ought to be some problem with the proof. Well, let's watch and see if there's a problem with the proof.

So again, I can assume that any set of n horses have the same color. I have to prove that any set of n plus 1 horses have the same color. How will I do that? Well, there's a set of n plus 1 horses, and let's consider the first n of those horses. Now by induction hypothesis, the first n of them have the same color. Black, maybe.

Also by induction hypothesis, the second set of n horses-- that is, all but the first horse-- have the same color. And what that tells me is that the first and the last horse have the same color as all of the horses in the middle. And therefore, in fact, they all have the same color. End of

proof, QED.

So, there had better be something wrong. And what's wrong? Well, what's wrong is that the proof that P of n implies P of n plus 1 is wrong. It looked good, but the proof that P of n implies P of n plus 1 has to work for all n greater than or equal to the base case. And if you look at this proof, it doesn't work in the base case. When n is 1 and you're trying to go from the base case to two and so on by implication, the proof breaks down.

Because what happens with our argument in the case that we're trying to prove that P of n implies P of n plus 1 in the case that n equals 1, well in that case, there aren't any middle horses. n plus 1 is 2, so there's a first horse. And that's the first n horses. And then the second half of set of n horses is the other horse, and there are no middle horses that they both have a color in common with. So, the proof just broke there.

But, you might not have noticed because that was the only place it was broken. This is a case where we were misled by ellipsis, by the way. Where because I was drawing n plus 1 horses with-- showing a lot of horses with dots in the middle, it looked like there were some. But they weren't. And again, as I said, the point though is that the only fallacy in this proof was that it didn't work when n was one. But it certainly worked for implying that if all sets of two horses are the same, that does imply that all sets of three horses are the same color.

And again, it's a false, we'll imply anything, kind of example. But even here, the proof was logically OK. But if it breaks in one place, if there's one domino that's missing from the line when the one before it falls, the rest of them stop falling and the proof breaks down.