PROFESSOR: We're going to look at the most fundamental of all mathematical data types, namely sets, and let's begin with the definitions. So informally, a set is a collection of mathematical objects, and the idea is that you treat the collection of objects as one new object. And that's a working definition. But of course, it might help a little bit, but it's a circular definition. This is not math yet because I haven't defined what a collection is, and a collection is no clearer or easier to define than a set is.

So let's try to work up the idea of sets by looking at some examples. So we've already talked about some familiar ones. There's the real numbers for which we had this symbol R in a special font, and the complex numbers C, and the integers Z. And we might have mentioned, and you might have seen already, the idea of the empty set for which we use this symbol that looks like a zero with a line through it.

Let's look at an example to pin things down. Let's look at this set of four things. Namely, it's got two numbers, pi/2 and 7, a character string in quotes, "Albert R," and the Boolean value true. So those are the four different things in it. They're of mixed type. And you might not like to have a mixed type like this in a programming language. But mathematicians don't worry about such things very much, rarely.

Anyway, the first observation is that the order in which these elements are listed doesn't matter. This set-- the braces indicates that it's the set of these things-- is the same if I listed T first, then the string, and the two numbers last. There is no notion of order in a set.

Now, to a computer scientist this is a little unnatural. The most natural thing to be would be to define a sequence of things, like the sequence that began with 7, then had the character string, then had the number, then had the Boolean. And you could get by with working with lists of things as long as they're finite.

But they very quickly get out of hand when you have to talk about, say, a set of lists. Then it's not clear how to make a list out of those, and you wind up making sets again. So sets, in fact, are an unavoidable kind of idea.

So another basic thing to understand about the notion of a set is that an element is either in a set or not in a set. So if I write down 7, pi/2, 7, this is the same description of the same set 7, pi/2. I'm just telling you the same thing twice here. That 7 is in the set, and the 7 is in the set

again. So no notion of being in the set more than once.

Now, sometimes, technically you want to add a notion of so-called multisets in which elements can be in a set a certain number of times, an integer number of times. But there's no real need for that. It's a secondary idea. And from our point of view, you're in or out of a set. If you repeat elements, it's the same as mentioning them once.

So the most fundamental feature of a set is what's in it. And for that, there's a special notation. So we'll say that x is a member of A, where A is a set, and use this epsilon symbol to indicate membership. It's read x is a member of A. So for example, pi/2 is a member of that set that we saw before that had pi/2 in it.

14/2 is also a member of that set because 14/2 is just another description of 7. When I write 7 here, I don't mean the character 7. I mean the number 7. And so 14/2 is the description of the same number. It's in that set.

On the other hand, pi/3 is a number that's simply not in that set. So I'm using the epsilon with a vertical bar through it, or some kind of a line through it, to mean not a member of. Membership is so basic that there's a lot of different ways to say it. Besides using the membership symbol x is a member of A, you can sometimes say x is an element of A, or x is in A, as well as x is a member of A. They're all synonyms.

So for example, 7 is a member of the integers. Z is our symbol for the integers. 2/3 is not a member of the integers because it's a fraction that's not an integer. And on the other hand, the set Z of integers itself is a member of this three-element set consisting of the truth value T, the set of all integers, and the element 7. So here's an example where a set can contain sets, quite big ones even, and that's fine. That's not any problem mathematically.

Related to membership is another fundamental notion of subset. So A is a subset of B, it's pronounced. So that horizontal U with a line under it is meant to resemble a less than or equal to symbol. So you can think of it as being A is less than or equal to B. But don't overload the symbols. Less than or equal to is used on numbers and other things that we know how to order.

And this is a relation that's only allowed between sets. So A is a subset of B-- A synonym is that A is contained in B-- simply means that every element of A is also an element of B. If I wrote that out in predicate logic notation as a predicate formula, I'd say for every x, x is in A

implies x is in B. If it's in A, then it's in B. Everything in A is in B.

So some examples of the subset relation are that the integers are a kind of-- an integer is a special case of a real number. So the set of integers is a subset of the real numbers. A real number is a special case of a complex number, so the real numbers are a subset of the complex numbers.

And here's a concrete example, where I have a set of three things, 5, 7, and 3, and this is the set with just the element 3 in it. Now, we sometimes are sloppy about distinguishing the element 3 from the set that's consisting of just 3 as its only element.

But in fact, it's a pretty important distinction to keep track of. In this case, 3 is not a subset of this set on the right. But the set consisting of 3 is a subset of the set on the right because, after all, the only member of this set is 3, and that is a member of this set.

A consequence of this general definition is that every set is a subset of itself because everything in A is in A . That's not really very interesting. Another important general observation is that the empty set is a subset of everything. The empty set is a subset of every set.

Let's look at why that is in more detail. So the claim is that the empty set is a subset of everything. Let B be any old set, then the empty set is a subset of B. What exactly does that mean according to the definition of subset? Well, it says that everything that's in the empty set, if it's in the empty set, then it implies that it's in B. For every element, if it's in the empty set, then it's in B.

Well, what do we know about this? The assertion that x is in the empty set is false. No matter what x is, there's nothing in the empty set. And now I have an implication that implies where the left-hand side, the hypothesis, is false. That means that the whole implication is true, and it doesn't depend on what B is. I'm not even going to look at B. I can see that x is in empty set is false, so the whole implication is true.

And so what I'm saying is that for every x, something that's true has to be true. Well, it is. And that's why the empty set is a subset of B satisfies this definition in a formal way. And this is an example of why that convention that false implies anything is convenient and is made use here.

So when you're defining sets, if they're small, you can just list the elements, as we did with that

set with 7 and pi/2 and "Albert R." Sometimes we can even describe infinite sets as some kind of a list. Like I might describe the set of integers as saying, well, it's 0, 1, minus 1, 2, minus 2, and so on, and you would understand that.

But in general, if I'm describing a set that is not so easy to list, say the real numbers, then what I'm going to do is define a set by a defining property of in the set. So I'm interested in a property P of elements, and I'm going to look at the set of elements x that are in some set A such that P of x is true, and that's the notation we use.

So this would be read as the set of x in A such that P of x holds, that x has property P. So notice this vertical bar is read as "such that." It's just a mathematical abbreviation. This is those elements in A that have property P that P of x holds for, and that defines a set of those elements.

Let's look at a simple example. The set E of even integers is simply the set of numbers n that are integers such that n is even. So in this case, the property P of n means that n is even.

One last concept is the concept of the power set. So the power set of a set A is all of the subsets of A. So we could define it using set notation as it's the set of B such that B is a subset of A. An example would be-- let's take the power set of the two Boolean values true and false.

So the power set of true and false, of that set consisting of two elements, is-- well, what are some of its subsets? The set consisting of just true is a subset of true, false. So is the set consisting of false, and so is the whole thing. It's a subset of itself. And one final element, the empty set, is a subset of the set of Boolean values true and false.

So the power set of this two-element set is a set that has four things in it-- two elements of size 1, one element of size 2, one element of size 0. And that's going to be a general phenomenon that we'll examine more later. How big is the power set of a set?

The even numbers, E that we just defined on the previous slide, is a member of the power set of Z because it's a subset of integers. Even integers are a special case of integers. And the integers are a member of the power set of R. That's just a synonym for saying that integers are a subset of reals. Every integer is a real, so the integers are a subset of reals, which means they're a member of the power set of reals.

So the general property is that a set B is a member of the power set of A if and only if B is a

subset of A. That was the defining condition for power set. And that's a fact to remember, and it may potentially confuse you. But it's a good exercise in keeping track of the difference between is a member of and is a subset of.