Second Order Logic

Review

We saw in chapter 10 that the system of quantificational logic that we are studying is called “first-order logic” because of a restriction in what we can “quantify over.” In FOL, we quantify over individuals, but not over properties.

That is, in FOL we can take an atomic sentence like Cube(b) and obtain a quantified sentence by replacing the name with a variable and attaching a quantifier:

∃x Cube(x)

But we cannot do the same with the predicate. That is, the following expression:

∃P P(b)

is not a sentence of FOL. But this is a legitimate sentence of second-order logic.

More expressive power

As a result, second-order logic has much more “expressive power” than FOL does. For example, there is no way in FOL to say that *a* and *b* have some property in common; but in second-order logic this would be expressed as ∃P (P(a) ∧ P(b)).

Similarly, second-order logic recognizes as formally valid certain inferences that are not FO-valid. For example, the seemingly valid argument:

| *a* is a cube and *b* is a cube
|—
| There is a property that *a* and *b* both have.

is not FO-valid, but it is a valid argument in second-order logic. It would be formalized as follows:

| Cube(a) ∧ Cube(b)
|—
| ∃P (P(a) ∧ P(b))

And this is a valid application of **∃-Intro** in second-order logic.

Here is another way in which second-order logic simplifies the expression of things that can be said only in a roundabout way in FOL. Suppose we would like to say in the blocks language that *a* and *b* have the same shape. Since we have the predicate SameShape, we would just write:

SameShape(a, b)

But suppose we deleted this predicate from the blocks language. Then the best we could do is something like this:

(Cube(a) ∧ Cube(b)) ∨ (Tet(a) ∧ Tet(b)) ∨ (Dodec(a) ∧ Doced(b))

Since in the blocks world the only shapes are *cube*, *tetrahedron*, and *dodecahedron*, for *a* and *b* to have the same shape is for them either to be both cubes, both tetrahedra, or both dodecahedra. But this FOL sentence doesn’t seem to mean quite the same thing as the English sentence it is translating—for example, it doesn’t say anything about the fact that it is *shape* that *a* and *b* have in common.

In second-order logic, by contrast, we could add to the blocks language a predicate Shape that is true of precisely the properties corresponding to the predicates Cube, Tet, and Dodec . That is,

$$Shape(Cube) \land Shape(Tet) \land Shape(Dodec)$$

and there are no other shapes in a blocks world. So we could write:

$$\exists P\ (Shape(P) \land P(a) \land P(b))$$

And this will come out true in exactly those worlds where *a* and *b* are either both cubes, both tetrahedra, or both dodecahedra. So in second-order logic we can express the idea of *same shape* using identity and the second-order predicate Shape; we can do without the special predicate SameShape.

Similarly, we can express the claim that *no object has every shape* in a way that brings out the quantifier in *every shape*:

$$\neg\exists x\ \forall P(Shape(P) \rightarrow P(x))$$

In FOL, the best we can do is to say that no block is both a cube, a tetrahedron, and a dodecahedron:

$$\neg\exists x\ (Cube(x) \land Tet(x) \land Dodec(x))$$

In this last FOL sentence we used all the shape predicates in the blocks language, but we did not say anything that means the same as *every shape*.

Properties of properties

Unfortunately, along with the greater expressive power of second-order logic come some very serious problems. These problems arise because of the fact we just observed above, namely, that the properties of blocks can themselves have properties. For example, we noted above that just as block *b* can have the property of being a cube (expressed in FOL as Cube(b)), so the property of being a cube can have the property of being a shape (not expressible in FOL, but expressed in second-order logic as Shape(Cube)). Note carefully that it is not the cube, *b*, that is said to have the property of being a shape, but the (first-order) property of being a cube that has the (second-order) property of being a shape.

This does not at first seem to be a problem. There seems nothing wrong, for example, in saying that some properties are common (possessed by many things) and some are uncommon (possessed by very few things). *Commonness*, then is a property of some properties. We might wish to define commonness as follows: we'll call a property *common* iff at least two things have it. Thus, the property of being a cube is common (in most worlds, at any rate), since there are many cubes, but the property of being president of the U.S. in 2001 is not common, since it is possessed only by George W. Bush. Formally speaking, our definition of *common* looks like this:

$$Common(P) \leftrightarrow \exists x\ \exists y\ (x \neq y \land P(x) \land P(y))$$

We've just noted that the property of being a cube is common in any world in which there are many cubes; similarly, the property of being large is common in any world in which there are many large things. What about the property of being common? Is it common?

Well, the property of being a cube is common, as we noted above, since there are many cubes, and so is the (different) property of being large. So we have:

$$Cube \neq Large \land Common(Cube) \land Common(Large)$$

Hence, by **∃ Intro**, we have:

$$\exists P\ \exists Q\ (P \neq Q \land Common(P) \land Common(Q))$$

But this means that the property of being common is itself common—several things have it. That is:

Common(Common)

This is a sentence of second-order logic that says that the property of being common has the property of being common. That is, it says that this property has itself as one of its properties!

Problems with second-order logic

It may seem to be no more than an oddity that there is a property that has itself as one of its instances—a self-exemplifying property. But worse is yet to come. For we might note that although the property of being common is self-exemplifying, not very many properties are like this. For example, the property of being a cube is not itself a cube; the property of being large is not large, etc. Facts such as these seem to be expressible in a second-order language as follows:

$\neg$Cube(Cube) $\neg$Large(Large) $\neg$Tet(Tet)   …

One might wish to mark this difference between *common*, on the one hand, and *cube*, *large*, and *tetrahedron*, on the other, by saying that *cube* and its ilk are **ordinary** properties, while *common* is extraordinary. *Ordinariness*, that is to say, is a property of most, but not all, of the properties we've considered so far. Since it is a property of at least some properties, we should be able, in a second-order language, to express the facts we just noted as follows:

Ordinary(Cube)
Ordinary(Large)
Ordinary(Tet)
$\neg$Ordinary(Common)

In fact, we can use our second-order language to define *ordinariness* as follows:

$\forall$P (Ordinary(P) $\leftrightarrow$ $\neg$P(P))

That is, a property is ordinary just in case it does not have itself as one of its properties.

Unfortunately, we now face a most difficult question: What about the property of being ordinary? Is it ordinary? The answer must be either *yes* or *no*, but both answers seem to get us into trouble. For if the answer is *yes*, then the property is self-exemplifying, which makes it extraordinary; and if the answer is *no*, then the property is not self-exemplifying, which makes it ordinary after all. We have arrived at a contradiction.

The contradiction emerges immediately from our definition of *ordinariness* above when one realizes that its universal quantifier $\forall$P ranges over all properties. So we can apply **$\forall$ Elim** to the definition and obtain:

Ordinary(Ordinary) $\leftrightarrow$ $\neg$Ordinary (Ordinary)

Second-order logic, then, runs the risk of falling into contradiction.

Russell's Paradox

Alert readers with a little knowledge of set theory will no doubt have noticed the similarity between this result and what Russell's Paradox shows about naïve set theory (see *LPL*, pp. 405-6, 432-3). For those who don't know about Russell's Paradox, here is a brief presentation of it. We'll start with some background information about sets.

A set is just a collection or a group of objects. These objects are the *members* of the set. The set of all teacups, for example, has all the teacups as its members. The set of all saucers is another set. Given any two sets, we can combine them in various ways to form new sets. For example, the *union* of two sets is the set which has as its members anything which is a member of either of them.

(The union of the two sets mentioned above is the set whose members are all the teacups and all the saucers—that is, anything that is either a teacup or a saucer.) The *intersection* of any two sets is the set which has as its members anything which is a member of both of them. (The intersection of the two sets mentioned above is the set which has as its members anything that is both a teacup and a saucer. In this case, the intersection is *empty*—that is, it has no members, since no teacup is a saucer.) One set can also be *included in* (or be a *subset* of) another. The set of porcelain teacups is a subset of the set of teacups, since every porcelain teacup is also a teacup.

Now comes a point that is crucial both to set theory and to Russell's Paradox: we can also treat sets as objects that are themselves members of other sets. For example, we can form a set, *T*, that has exactly two members—(1) the set of teacups and (2) the set of saucers. Notice that the set of teacups is a *member* of *T*, but it is not *included in T*. (That is because the members of *T* are sets, not teacups, and the set of teacups is not a teacup.) Set membership is not the same thing as set inclusion.

When a set does have sets as members, it does not typically have *itself* as a member. But it would seem that in some cases, a set might have itself as a member. For example, since sets are abstract objects, the set of all abstract objects would be a member of itself, since it is an abstract object.

This is where Russell's Paradox begins. Just as we defined an *ordinary* property as one that is not self-exemplifying, Russell defined an *ordinary* set as one that is not a member of itself. Then Russell asked us to consider *R*, the set of all ordinary sets. That is, *R* is defined as follows (where 'α' is a variable ranging over sets and '$x \varepsilon \alpha$' means 'x is a member of α'):

$$\forall\alpha\ (\alpha\ \varepsilon\ R \leftrightarrow \neg\ \alpha\ \varepsilon\ \alpha)$$

That is, *R* has as its members all and only the sets that are not members of themselves. We then apply **∀ Elim** to this definition and obtain:

$$R\ \varepsilon\ R \leftrightarrow \neg\ R\ \varepsilon\ R$$

R, the set of all ordinary sets, is a member of itself iff it is not a member of itself. So naïve set theory leads to a contradiction.

Prospects for second-order logic

The similarity between second-order logic and set theory has led some people to say that second-order logic is just "set theory in sheep's clothing" (Quine, *Philosophy of Logic*, 1970). This claim, however, is controversial (for criticism of it, see Boolos, "On Second-Order Logic," *J. Phil.* 72 (15), 1975). And just as set theory (albeit not naïve set theory) can be developed in a way that avoids Russell's Paradox, so can second-order logic be developed in a contradiction-free way.

For example, one can distinguish between first-level properties (properties of individuals) and second-level properties (properties of first-level properties), and insist that no quantifier can range over properties of both levels. Thus, Cube and Tet are first-level properties, and Common is a property of such first-level properties as Cube and Tet, But although Common has many instances (just as Cube has many instances), that does not mean that Common has itself as a property. What makes Cube common is that it has many *individual* instances, whereas what makes Common common is that it has many *first-level properties* as instances. So the property that Common has in virtue of having many instances is not the same as the property that Cube has in virtue of having many instances. Hence, Common does not, after all, have itself as a property. Rather, the first-level property Common_1 has a second-level property Common_2.

This means that our definition of *ordinariness*:

$$\forall P\ (\text{Ordinary}(P) \leftrightarrow \neg P(P))$$

is defective, in that the universal quantifier ∀P purports to range over both first and second level properties.

However, even if second-order logic can be given a consistent formulation, it has other shortcomings. It has been shown, for example, that there can be no set of inference rules for second-order logic that is both sound and complete. This contrasts sharply with FOL, for which we have system $\mathcal{F}$, among others, that are both sound and complete.