Part II

Extension

4

Predication

21. The First-Order/Higher-Order Controversy

First-order quantifier logic is complete; higher-order quantifier logic is not. A few formally minded philosophers of logic—such as Quine and some of his followers—appear to believe that this is sufficient grounds for concluding that the only legitimate quantifier logic is first-order, not higher-order. However, most leading formally minded philosophers of logic over the past hundred years—Frege, Russell, Church, Carnap, Henkin, Montague, Kaplan—believe that the higher-order approach is a natural generalization of the first-order approach and therefore that quantifier logic is properly identified with higher-order quantifier logic. I depart from this majority opinion. In §10 I gave several formalistic reasons (including completeness) for preferring the first-order approach over the higher-order approach. But formalistic reasons tell us little about the issues of naturalness and generality. In this chapter I will discuss the underlying philosophical differences between the two approaches to quantifier logic.[1] My hope is that the greater naturalness and generality of the first-order approach will become evident in the course of the discussion.

Consider the following intuitively valid argument:

$$\frac{x \text{ is red and } y \text{ is not red.}}{\therefore\ \text{There is something that } x \text{ is and that } y \text{ is not.}}$$

There are two approaches to the representation of this argument—the first-order approach and the higher-order approach. On the higher-order approach the argument is represented as an instance of second-order existential generalization:

$$\frac{Rx \mathbin{\&} \neg Ry}{\therefore\ (\exists f)(fx \mathbin{\&} \neg fy)}$$

where R is a name of the color red and f is a predicate variable for which R is a substituend. On the first-order approach the argument is represented as an instance of first-order existential generalization:

$$\frac{x \,\Delta\, r \,\&\, y \not\Delta\, r}{\therefore\ (\exists z)(x \,\Delta\, z \,\&\, y \not\Delta\, z)}$$

where r is a name of the color red and Δ is a distinguished 2-place logical predicate that expresses the *predication relation*, a relation expressed by the *copula* in natural language.[2]

There are analogous examples involving relations rather than properties:

x and y are husband and wife, and u and v are not husband and wife.

∴ There is something which x and y are that u and v are not.

On the higher-order approach this intuitively valid argument is represented as an instance of second-order existential generalization:

$$\frac{H^2(x, y) \,\&\, \neg H^2(u, v)}{\therefore\ (\exists f^2)(f^2(x, y) \,\&\, \neg f^2(u, v))}$$

where the 2-place predicate H^2 is construed as a name for the relation holding between husband and wife and f^2 is a 2-place predicate variable. On the first-order approach the argument is represented as an instance of first-order existential generalization:[3]

$$\frac{\langle x, y\rangle \,\Delta\, [H^2(x, y)]_{xy} \,\&\, \langle u, v\rangle \not\Delta\, [H^2(x, y)]_{xy}}{\therefore\ (\exists z)(\langle x, y\rangle \,\Delta\, z \,\&\, \langle u, v\rangle \not\Delta\, z).}$$

Philosophically speaking, how do the higher-order and first-order approaches differ? In the next few sections I suggest an answer to this question.

22. Expressive Power

It is often thought that a higher-order language has greater expressive power than the first-order counterpart. However, for appropriate first-order languages (such as L_ω with Δ and $\approx_N$), this is not so, and indeed the situation is typically the other way around. In fact, since the variables in L_ω are free to range over all the objects falling within the range of any given higher-order variable, higher-order notation can be contextually defined in L_ω with Δ and $\approx_N$.

For illustrative purposes I will show how this can be done for a sample second-order language. More complex higher-order languages can be dealt with on analogy.[4] I begin with some preliminary definitions:

$$\text{Individual Particular } (x)\ \textit{iff}_{\text{df}}\ \Box(\forall y)(y \,\Delta\, x \equiv y = x)$$

$$\text{Proposition } (x)\ \textit{iff}_{\text{df}}\ (\exists y)(x \approx_N [x \,\Delta\, y]^y)$$

$$\text{Property } (x)\ \textit{iff}_{\text{df}}\ (\exists y)(x \approx_N [z \,\Delta\, y]_z^y)$$

$$N\text{-ary Relation } (x)\ \textit{iff}_{\text{df}}\ (\exists y)(x \approx_N [\langle z_1, \ldots, z_n\rangle \,\Delta\, y]_{z_1 \ldots z_n}^y)$$

$$\text{True } (x)\ \textit{iff}_{\text{df}}\ (\exists y)(x \approx_N [(\exists z) z \,\Delta\, y]^y \;\&\; (\exists z) z \,\Delta\, y)$$

$$\ldots\{x, y\}\ldots\ \textit{iff}_{\text{df}}\ (\exists z)((\forall w)(w \,\Delta\, z \equiv (w = x \vee w = y)) \;\&\; \ldots z \ldots)$$

where z is a new variable not occurring in . . .

$$\langle v_1 \rangle =_{\text{df}} v_1$$

$$\langle v_1, v_2 \rangle =_{\text{df}} \{\{v_1\}, \{v_1, v_2\}\}$$

$$\langle v_1, \ldots, v_{n+1} \rangle =_{\text{df}} \langle \langle v_1, \ldots, v_n \rangle, v_{n+1} \rangle.$$

The definition of individual particular has interesting historical roots as far back as works by Peter Abelard and Leibniz and, more recently, in Lesniewski's 'Ontology' and Quine's 'New Foundations' and *Mathematical Logic*. The definition says that x is an individual particular if and only if it is necessary that x is predicable of itself and itself only. What is particular about particulars is that necessarily they are predicable of themselves and themselves only. The definition of truth says that x is true if and only if there is a property y such that x is necessarily equivalent to the proposition that y has an instance and y does in fact have an instance. So, for example, let y be the property of being something z such that x is true (i.e., $[x \text{ is true}]_z^x$). Then, x is a proposition if and only if x is necessarily equivalent to the proposition that y has an instance. And, in turn, x is true if and only if y in fact has an instance. Incidentally, although this definition of truth is logically adequate, it is not the official definition that I will offer in §45. It is given here for illustrative purposes. Finally, if our background theory should be T1 instead of T2′, then $=$ should replace $\approx_N$ in the above definitions.

Now in order to contextually define any given second-order sentence C in the first-order language L_ω with Δ and $\approx_N$, simply convert C into the sentence C' of L_ω by means of the following conversion rules:

(1) Second-order atomic formulas (where p_i is a sentential variable and f_i^n is a predicate variable):
$p_i \Rightarrow \text{True}\,(p_i)$.

$f_i^n(t_1, \ldots, t_n) \Rightarrow \langle t_1, \ldots, t_n\rangle\, \Delta f_i^n$.

(2) Restricted quantifiers (where v_j is a new variable not occurring in $A(a_i)$):

$(\forall a_i)A(a_i) \Rightarrow (\forall v_j)(v_j$ is an individual particular $\supset A(v_j))$.

$(\forall p_i)A(p_i) \Rightarrow (\forall v_j)(v_j$ is a proposition $\supset A(v_j))$.

$(\forall f_i^1)A(f_i^1) \Rightarrow (\forall v_j)(v_j$ is a property $\supset A(v_j))$.

$(\forall f_i^n)A(f_i^n) \Rightarrow (\forall v_j)(v_j$ is an n-ary relation $\supset A(v_j))$.

To apply these conversion rules to a given higher-order sentence C, begin with the innermost formula in C and apply rules (1) and (2) in that order; then, working outward in C, repeat this process until no higher-order notation remains. The result is the sentence C' of L_ω with Δ and $\approx_N$. C is then contextually defined as follows: C *iff*$_{df}$ C'.

Since L_ω with Δ and $\approx_N$ has a single sort of variable that ranges over everything, it is actually more expressive (and in this sense, more general) than the typical higher-order language. Yet it is possible, though quite uncommon, for a higher-order language to have just one sort of variable. Therefore, greater expressive power cannot be used as a fail-safe criterion for distinguishing the first-order approach to logic from the competing higher-order approach. For such a criterion we must look to the subject/predicate distinction.

23. The Subject/Predicate Distinction

The first-order approach adopts the traditional linguistic distinction between subject and predicate, between noun and verb; the higher-order approach does not. That is, on the first-order approach an absolute distinction is made between linguistic subjects and linguistic predicates such that a linguistic subject (noun) cannot except in cases of equivocation be used as a linguistic predicate (verb) and conversely. The higher-order approach does not impose such a restriction.[5]

The distinction between linguistic subject and linguistic predicate is evident in the surface syntax of natural language. To see the distinction there, notice that English predicates, e.g., the verbs 'repeats' and 'cycles', can never (without equivocation) occur as subjects: e.g., 'repeats = cycles' is just not a sentence. Likewise, subjects, like the noun phrases '1/3' and '.333...', can never (without equivocation) occur as predicates: e.g., '1/3 .333...' is not a sentence either. By contrast, subjects and predicates, when combined with each other in the proper order, do form sentences: e.g., '.333... repeats' and '1/3 cycles' are sentences. Thus, at least as far as the surface syntax of English is concerned, there does seem to be a sharp distinction between linguistic subjects and linguistic predicates. And this is the distinction that is built into the syntax of first-order languages. In the syntax of higher-order languages, however, this distinction is glossed over.

Although in natural language predicates cannot be used as subjects, it is possible to transform predicates into legitimate subjects by means of certain abstraction operations. (The resulting linguistic subjects are complex abstract noun phrases.) So by nominalizing the verb 'repeats', we may transform it into the gerund 'repeating', and by nominalizing the verb 'cycles', we may transform it into the gerund 'cycling'. Since these nominalized expressions are legitimate linguistic subjects, they can be combined with verbs (e.g., '=' and 'is') to form sentences. Hence, e.g., 'repeating = cycling', '.333... is repeating' and '1/3 is cycling' are sentences. Nominalizations are naturally represented in first-order language by means of the bracket notation. These three sentences may thus be represented by '$[Rx]_x = [Cx]_x$', '$.333 \ldots \Delta [Rx]_x$', and '$1/3 \, \Delta \, [Cx]_x$', respectively. By contrast, in a higher-order language, where the distinction between a predicate and its nominalization is glossed over, the above three English sentences would typically be represented by '$R = C$', '$R(.333\ldots)$', and '$C(1/3)$', respectively.

What function does the subject/predicate distinction have? First, in speech the distinction shows up as follows. A subject expression is the kind of expression that functions to identify a thing about which something is to be said. A predicate expression, by contrast, functions to say something about things so identified. As Strawson might put it, subjects fix the subject matter, and predicates (verbs) do the saying. Secondly, the subject/predicate distinction plays a role in syntax. For example, in the syntax for first-order extensional language there are three primitive syntactic categories—subject, predicate, operator—and one defined syntactic category—sentence (open or closed). The definition of sentence is roughly this: subjects combine with predicates to form sentences, and operators combine with sentences to form sentences. Hence, a very natural syntax.[6] Thirdly, the subject/predicate distinction plays a role in the construction of a natural, economical semantics that tallies with the intuitive concept of meaning. Let me explain.

Consider the kind of semantics that I call *Russellian semantics*. In this semantics, unlike a Fregean semantics, there is just one fundamental kind of meaning, and the familiar semantic relations of naming and expressing are defined in terms of it, together with the syntactic notions of subject and predicate. Naming is just the restriction of the meaning relation to syntactically simple linguistic subjects:

x names y *iff*$_{\mathrm{df}}$ x is a syntactically simple linguistic subject and x means y.

And expressing is the restriction of the meaning relation to linguistic predicates and syntactically complex expressions:

x expresses y *iff*$_{\mathrm{df}}$ x is a linguistic predicate or a syntactically complex expression and x means y.

In a first-order language, since no linguistic predicate or formula is a linguistic subject, linguistic predicates and formulas do not, according to a Russellian semantics, name at all. This result tallies with the intuitive notion of naming. For according to the intuitive notion, predicates and sentences do not name. (What do 'runs', 'equals', 'is', 'x runs', 'Everything equals something', etc. name? Intuitively, they name nothing at all.) By contrast, in higher-order languages all predicates and sentences are also linguistic subjects. Thus, by suppressing the distinction between linguistic predicates (and sentences) and linguistic subjects, the higher-order approach yields the counterintuitive consequence that all linguistic predicates (and sentences) name something.

A related difficulty arises in connection with Frege's question of how a true sentence '$a = b$' can differ in meaning from '$a = a$'. Frege's two-kinds-of-meaning semantics is expressly designed to answer this question. In §38, however, I show that for an idealized representation of natural language Russellian semantics is every bit as adequate as Fregean semantics. The argument makes use of the fact that strings such as '$F = G$' and '$F = F$' are ill-formed in a first-order language (since linguistic predicates are not counted as linguistic subjects). But such strings are well-formed in higher-order language (since linguistic predicates are there counted as linguistic subjects). Thus, in a higher-order setting, unlike a first-order one, we need special assurances that strings such as '$F = G$' and '$F = F$' do not constitute problematic new instances of Frege's puzzle. (This is Church's worry about Russellian semantics; see §38.) In this way, our simple and natural Russellian semantics becomes problematic when we move to a higher-order setting from a first-order one. This then is one more way in which the traditional subject/predicate distinction, as it is incorporated in first-order language, plays a role in linguistic theory.

The distinction between linguistic subjects and linguistic predicates is, of course, reminiscent of Frege's distinction between object-names and function-names. There are important differences, however. One of these differences is ontological in character. According to Frege's theory, object-names name things called *objects*, and function-names name things called *functions*. Objects are what Frege calls complete (or saturated); functions are what he calls incomplete (or unsaturated). (He further distinguishes ordinary functions from functions whose values are truth values. The latter he calls *concepts*. I will suppress this distinction in the present remarks.) However, in the framework of the first-order theory of PRPs there is a far more natural ontological distinction that does much the same job as Frege's function/object distinction. What I have in mind is the distinction between things that are ontological predicates and things that are not. Something is an *ontological predicate* if and only if in principle it could be expressed by a linguistic predicate. Now let us agree that an *object* is anything that could be named by a linguistic subject. While it is true that any ontological predicate is ontologically distinctive (for it is either a property or relation), it is also true that each ontological predicate is an object. Indeed, any property or relation can simply be assigned as the value of a first-order variable.[7] Herein lies the difference between ontological predicates and Frege's functions, for on Frege's theory no function can ever be an object. And so Frege must say that the concept horse is not a concept!

How did Frege arrive at this bizarre distinction? My suspicion is that the distinction had its origin in none other than Frege's proclivity to treat the logical syntax of natural language as higher-order and, specifically, in his proclivity to treat all constants in natural language as names, including even those constants that were traditionally identified as linguistic predicates. Let me explain.

Frege was well aware of natural language phenomena such as the following: for all linguistic subjects *b* and all linguistic predicates *F* if $\ulcorner \ldots b \ldots \urcorner$ has a truth value (or makes sense), then barring equivocation $\ulcorner \ldots F \ldots \urcorner$ does not have a truth value (or sense). For example, 'Cycling is a property' has a truth value (makes sense), but 'Cycles is a property' does not. (See p. 50, Frege, 'On Concept and Object'.) When Frege sought to explain such linguistic phenomena, he arrived at an ontologically based semantical explanation. 'Cycling is a property' has a truth value (sense) because 'cycling' and 'is a property' name (express) things that by their nature combine together to yield something else; that thing is the nominatum (sense) of 'Cycling is a property'. By contrast, 'Cycles is a property' does not have a truth value (sense) because 'cycles' and 'is a property' name (express) things that cannot by their nature combine together to yield something. Hence, there is nothing with which to identify the nominatum (sense) of 'Cycles is a property'.

In contrast to Frege's ontologically based higher-order semantical explanation, the first-order explanation of the above natural language phenomena is *syntactic*. Complex expressions have truth value (make sense) if and only if they are syntactically well-formed formulas. To be a syntactically well-formed formula a complex expression must be built up according to the syntactic formation rules. However, the syntactic formation rules prohibit using linguistic predicates as linguistic subjects. This simple syntactic line of explanation was unavailable to Frege, for on his theory both linguistic subjects and predicates are names. Thus, Frege could explain the failure of the substitutability of linguistic predicates for linguistic subjects only by positing a bizarre ontological distinction between the kind of things named by linguistic predicates and the kind of things named by linguistic subjects, i.e., by positing the distinction between functions and objects.

24. The Property/Function Distinction

Frege's bizarre ontological distinction between functions and objects has not had much impact historically. Nevertheless, the Fregean doctrine that predicates name functions has had a persistent influence on subsequent higher-order formulations of logic. Here, of course, the theory that functions cannot be objects is suppressed. The practice of treating predicates as naming functions has been taken up by Russell (in *Principia Mathematica*), Church, Henkin, Montague, Kaplan, and David Lewis, to name a few. I will now make some criticisms of this practice.

Consider the following intuitively valid argument:

x is red and red differs from blue.
∴ There is something that x is and it differs from blue.

The standard higher-order representation of this argument is:

$$\frac{R(x)\ \&\ R \neq B}{\therefore\ (\exists f)(f(x)\ \&\ f \neq B)}$$

where the predicates R and B are construed as names of the properties red and blue, respectively, and f is a 1-place predicate variable. Now if in accordance with the common higher-order practice (n-ary) predicates are also construed as naming (n-ary) functions, then the properties red and blue must be identified with 1-ary functions.[8] Indeed, all properties (i.e., all 1-ary intensional entities) must on this higher-order approach be identified with 1-ary functions. And similarly, n-ary relations (i.e., n-ary intensional entities, for $n \geq 2$) must be identified with n-ary functions.

But how unnatural such identifications are. Joy, the shape of my hand, the aroma of coffee—these are not functions. When I feel joy, see the shape of my hand, or smell the aroma of coffee, it is not a function that I feel, see, or smell. (For more on sensing and feeling, see §49.) Indeed, from the intuitive point of view n-ary functions are just a special kind of $n + 1$-ary relations, namely, those $n + 1$-ary relations that are univocal. Thus, the higher-order practice results in an identification of properties with 2-ary relations, 2-ary relations with 3-ary relations, 3-ary relations with 4-ary relations, etc. This outcome is entirely unintuitive.

On the first-order approach, this unintuitive outcome is easy to avoid. The above argument, for example, is straightforwardly represented as

$$\frac{x \,\Delta\, r \,\&\, r \neq b}{\therefore\ (\exists w)(x \,\Delta\, w \,\&\, w \neq b)}$$

where r and b are singular terms denoting the properties red and blue, respectively. On this approach properties are just what they should be—1-ary intensional entities. Likewise, n-ary relations-in-intension are just what they should be—n-ary intensional entities. And propositions are just what they should be—0-ary intensional entities.

The higher-order practice of identifying properties with functions has often led to another difficulty. To dramatize this difficulty consider the following propositions, where x is some particular:

$$[Fx]^x$$

$$[x \,\Delta\, [Fy]_y]^x$$

$$[x \,\Delta\, [u \,\Delta\, [Fy]_y]_u]^x$$

$$[[Fy]_y \,\Delta\, [x \,\Delta\, v]^x_v]^x$$

$$[\langle x, [Fy]_y\rangle \,\Delta\, [u \,\Delta\, v]_{uv}]^x$$

$$[\langle [Fy]_y, x\rangle \,\Delta\, [u \,\Delta\, v]_{vu}]^x \quad \ldots.$$

Although on conception 1 these propositions are identical, on conception 2, which concerns intentional matters, these propositions are all distinct. Now consider any higher-order functional approach to intensional logic that does not avail itself of a primitive Δ-predicate. (If a theory does avail itself of a Δ-predicate, one can hardly see the point of making the theory higher-order; recall §22.) On such a higher-order functional approach, the above propositions would be represented:

$$F_{o_1\iota}(x_\iota)$$

$$(\lambda y_\iota)(F_{o_1\iota}(y_\iota))(x_\iota)$$

$$(\lambda u_\iota)((\lambda y_\iota)(F_{o_1}(y_\iota))(u_\iota))(x_\iota)$$

$$(\lambda f_{o_1\iota})(f_{o_1\iota}x_\iota)(\lambda y_\iota)(F_{o_1\iota}(y_\iota))$$

$$(\lambda f_{o_1\iota})(\lambda u_\iota)(f_{o_1\iota}(u_\iota))((\lambda y_\iota)(F_{o_1\iota}(y_\iota)), x_\iota)$$

$$(\lambda u_\iota)(\lambda f_{o_1\iota})(f_{o_1\iota}(u_\iota))(x_\iota, (\lambda y_\iota)(F_{o_1\iota}(y_\iota))) \quad \ldots.$$

However, given the usual laws for λ, if the above propositions are represented in this way, they would all have to be identical. Therefore, intensional distinctions relevant to the logic for intentional matters are lost on the above kind of higher-order functional approach. Where does this approach go wrong?

Without attempting a detailed analysis, I think that I can in a rough way indicate the source of the problem. Consider the first two propositions $[Fx]^x$ and $[x \Delta [Fy]_y]^x$. Given the algebraic methods developed in chapter 2, we have the following:

$$[Fx]^x = \text{Pred}_0([Fy]_y, x)$$
$$[x \Delta [Fy]_y]^x = \text{Pred}_0(\text{Pred}_0([u \Delta v]_{uv}, [Fy]_y), x).$$

Thus, whereas the proposition $[Fx]^x$ is obtained by applying the predication operation to the property $[Fy]_y$ and x, the proposition $[x \Delta [Fy]_y]^x$ involves not only the predication operation but also the predication relation (the Δ-relation). The error in the above sort of higher-order functional approach is something like this. It in effect collapses the predication operation and the predication relation into the single Fregean operation of application of function to argument.

In view of the difficulties facing the functional approach to higher-order logic, why do higher-order theorists persist in treating predicates as names of functions rather than as names of properties? Beyond mere tradition and preoccupations with mathematics rather than natural logic, the major impetus for this practice is that it makes possible a relatively simple kind of semantics for higher-order language. A property-theoretic semantics, which would be more natural than a function-theoretic semantics, has to my knowledge never been accomplished for higher-order language. My conjecture is that the simplest way to construct one is, ironically, to translate the higher-order language into a first-order language, perhaps along the lines of §22, and then to do the property-theoretic semantics for the first-order language, perhaps along the lines of §§13–14.

25. The Origin of Incompleteness in Logic

We now return to the issue of incompleteness in logic, the issue with which this chapter began. Gödel showed that first-order number theory is incomplete. Since first-order number theory can be modeled within first-order set theory, first-order set theory is incomplete as well. A thesis of the next chapter is that first-order set theory can in turn be modeled within the first-order logic for the predication relation.[9] It follows that this logic is incomplete. Thus, in view of the results of chapter 2 we obtain the following fuller picture of the stages of completeness and incompleteness in first-order theories.[10]

COMPLETENESS AND INCOMPLETENESS IN FIRST-ORDER THEORIES

(1)	first-order quantifier logic with identity and the numerals	Complete
(2)	first-order quantifier logic with identity and extensional abstraction	Complete
(3)	first-order quantifier logic with identity and intensional abstraction	Complete
(4)	first-order quantifier logic with identity, the numerals, addition, and multiplication	Incomplete
(5)	first-order quantifier logic with identity and set membership and with or without extensional abstraction	Incomplete
(6)	first-order quantifier logic with identity and predication and with or without intensional abstraction	Incomplete

What is the origin of the incompleteness in logic? In view of (1), (2), and (3) in the above picture, the ontology of abstract entities clearly is not responsible. So in view of (4) in the above picture, one might be inclined to the view that the standard number-theoretic operations are responsible. However, if this answer is not elaborated, it is unconvincing. For on the face of it, operations from number theory do not even belong to logic *per se*, i.e., to the science of valid thinking. Similarly, in view of (5) in the above picture, one might be inclined to identify the relation of set membership as the source. However, as with operations from number theory, the relation of set membership does not on the face of it belong to the domain of logic *per se*.[11]

A thesis of chapter 6 is that all the usual operations from number theory are definable in L_ω in terms of the predication relation. And a thesis of chapter 5 is that, insofar as set theory has any utility in mathematics or empirical science, an ϵ-relation having all the properties attributed to ϵ in axiomatic set theory is definable in terms of the predication relation. Therefore, if these theses are correct and if the predication relation indeed falls within the domain of logic *per se*, then the incompleteness in logic can in this sense be traced to defined number-theoretic operations or to a defined ϵ-relation. However, since the logical character of these defined notions derives from their definability in terms of the predication relation, this relation, if it indeed belongs to the domain of logic *per se*, must be identified as the ultimate source of the incompleteness in logic. (See (6) in the picture opposite.)

Does the predication relation belong to the domain of logic *per se*? That is, is the theory for the predication relation truly part of the science of valid thinking? The answer to this question is obvious: if any theory at all ever qualifies as part of logic, the theory for the predication relation does; the predication relation is the very paradigm of a purely logical relation. This point, which has been neglected by virtually all twentieth century philosophers of logic,[12] cannot be stressed enough. The copula is a logical constant *par excellence*, and the theory for the copula is part of logic.

Therefore, my conclusion is this. It is not the infinite abstract ontology of logic, i.e., the infinite ontology of properties, relations, and propositions, that is responsible for the incompleteness in first-order logic. Rather, the ultimate source of the incompleteness is a fundamental logical relation on that abstract ontology, the predication relation. The logic for properties, relations, and propositions, the logic for L_ω is provably complete as long as no predicate is singled out as a distinguished logical predicate expressing the predication relation. However, as soon as a predicate is singled out in this way, the resulting logic is rendered incomplete.

What is the source of the incompleteness in higher-order theories? In higher-order settings, unlike the first-order setting, philosophically relevant stages of completeness and incompleteness evidently cannot be isolated,[13] for higher-order quantification theory is incomplete from the very start. Given the hypothesis that the predication relation is the source of the incompleteness in logic, we can explain the inability to separate philosophically relevant stages of completeness and incompleteness in higher-order quantification theory. This theory is incomplete from the start because the notation for the predication relation is built into the syntactic structure of higher-order languages[14] and, thus, the semantic import of this notation is never permitted to vary from one standard model to another. However, if higher-order quantification theory is treated as a derived theory constructed within the first-order logic for the predication relation (as in §22), then the source of the incompleteness in higher-order quantification theory—namely, the predication relation—becomes transparent.

26. The Logical, Semantical, and Intentional Paradoxes

The source of incompleteness in first-order logic, I have argued, is traceable to the predication relation. It should be no surprise, then, that I also hold that the predication relation lies at the heart of the familiar paradoxes that have plagued logicians over the years, e.g., the paradoxes of Russell, Cantor, Burali-Forti and the paradoxes of Epimenides, Berry, Grelling, and Richard. Specifically, I hold that, when properly analysed, each of these paradoxes involves some kind of self-refuting predication.

How do the paradoxes arise? The algebraic semantic technique provides a new perspective on this question. Consider the standard model structure $\mathcal{M}$ for L_ω with Δ:

$$\langle \mathcal{D}, \mathcal{P}, \mathcal{K}, \mathcal{G}, \mathrm{Id}, \bar{\Delta}, \mathrm{Conj}, \mathrm{Neg}, \mathrm{Exist}, \mathrm{Exp}, \mathrm{Inv}, \mathrm{Conv}, \mathrm{Ref}, \mathrm{Pred}_0, \mathrm{Pred}_1, \mathrm{Pred}_2, \ldots \rangle$$

Here $\bar{\Delta}$ is the relation-in-intension in $\mathcal{D}_2$ that is expressed by the predicate Δ on its standard interpretation. Now, what would one think is the extension of the predication relation $\bar{\Delta}$? Intuitively, one would think that a pair x, y is in the extension of the predication relation $\bar{\Delta}$ if and only if x is in the extension of y. That is, one would think that $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D} : x \in \mathscr{G}(y)\}$ and, more generally, that $(\forall H \in \mathscr{H})H(\bar{\Delta}) = \{xy \in \mathscr{D} : x \in H(y)\}$. But this is impossible, as the following model-theoretic analogue of Russell's paradox shows. Suppose that $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D} : x \in \mathscr{G}(y)\}$.

(1)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \in \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	Premise
(2)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \notin \mathscr{G}(\mathrm{Ref}(\bar{\Delta}))$	By (1) & Neg-rule[15]
(3)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \notin \mathscr{G}(\bar{\Delta})$	By (2) & Ref-rule
(4)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \notin \{xy \in \mathscr{D} : x \in \mathscr{G}(y)\}$	By (3) & hypothesis
(5)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \notin \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	By (4) & set theory
(1′)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \notin \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	Premise
(2′)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \in \mathscr{G}(\mathrm{Ref}(\bar{\Delta}))$	By (1′) & Neg-rule
(3′)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \in \mathscr{G}(\bar{\Delta})$	By (2′) & Ref-rule
(4′)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \in \{xy \in \mathscr{D} : x \in \mathscr{G}(y)\}$	By (3′) & hypothesis
(5′)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \in \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	By (4′) & set theory

Thus, given the law of the exluded middle, the hypothesis that $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D} : x \in \mathscr{G}(y)\}$ leads to a contradiction.

Another way to see the difficulty is this. Given the algebraic semantics for L_ω, the following holds for all formulas A:

$$\langle \mathscr{A}(v_1), \ldots, \mathscr{A}(v_j)\rangle \in \mathscr{G}(\mathrm{D}_{\mathscr{J}\mathscr{A}\mathscr{M}}([A]_{v_1 \ldots v_j})) \textit{ iff } \mathrm{T}_{\mathscr{J}\mathscr{A}\mathscr{M}}(A).$$

Therefore, if $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D} : x \in \mathscr{G}(y)\}$, then $v \,\Delta\, [A]_v \equiv A$ would have to be true for all formulas A. But this is just the principle of predication from which the property-theoretic analogue of Russell's paradox follows immediately:

$$[v \not\Delta v]_v \,\Delta\, [v \not\Delta v]_v \equiv [v \not\Delta v]_v \not\Delta [v \not\Delta v]_v.$$

What is going on? The language L_ω is *semantically complete* in the sense that, for every formula A, there is a singular term (namely, the normalized intensional abstract $[A]_\alpha$) that denotes the meaning of A. That is, all *expressible* properties, relations, and propositions are *denotable*. To my mind any language that provides an ideal treatment of modal and intentional matters ought to be semantically complete in this sense. Now consider a semantically complete language (e.g., L_ω) whose sentential and quantificational logic is classical. If such a language has a predicate (e.g., Δ) that expresses the predication relation, then necessarily the extension of the predication relation is different from what one would naively take it to be. If classical logic is sound, then, paradoxes in a semantically complete language originate in a mistake concerning the extension of the predication relation.

If classical logic is not to be tampered with,[16] then a resolution of the paradoxes in semantically complete languages must involve modifications in what one naively takes to be the extension of the predication relation. So it is quite pleasing to see that this is precisely what happens when the standard resolutions of the paradoxes in naive first-order set theory are adapted to first-order intensional logic with predication. Until we find an ideal resolution of the paradoxes of predication, we may therefore follow this maxim: to obtain a workable resolution of these paradoxes, determine the best resolution of the paradoxes in first-order set theory and then adapt it to the setting of intensional logic with predication.

For illustrative purposes I will now sketch how such adaptation works in the case of the two most familiar resolutions of the first-order set-theoretical paradoxes, namely, Zermelo's resolution and von Neumann's resolution.[17] In connection with the von Neumann-style resolution I will say that an object is *safe* if and only if it has properties, i.e.,

$$S(v_i) \ \textit{iff}_{\text{df}} \ (\exists v_j)\, v_i \,\Delta\, v_j.$$

As a notational convention, let the letters $a, b, c, \ldots$ be introduced as special restricted variables that range over safe things. Accordingly, $(\forall a_i)A(a_i)$ is short for $(\forall v_j)(S(v_j) \supset A(v_j))$, and $(\exists a_i)A(a_i)$ is short for $(\exists v_j)(S(v_j) \,\&\, A(v_j))$ where v_j is a new distinct variable. Now, as I have said, the following is the naive principle of predication that is responsible for the paradoxes:

(Naive Principle of Predication)
For any formula A,
$\vdash \langle v_1, \ldots, v_j \rangle \,\Delta\, [A]_{v_1 \ldots v_j} \equiv A.$

According to the Zermelo-style and the von Neumann-style resolutions of the logical paradoxes, the naive principle of predication is modified as follows:

(Zermelo-Style Principle of Predication)
For any formula A having the form $(v_1, \ldots, v_j \,\Delta\, u \,\&\, B)$,
$\vdash \langle v_1, \ldots, v_j \rangle \,\Delta\, [A]_{v_1 \ldots v_j} \equiv A$.

(von Neumann-Style Principle of Predication)
For any formula A where, for all h, $1 \leq h \leq j$, a_h is free for v_h in A and conversely,
$\vdash \langle a_1, \ldots, a_j \rangle \,\Delta\, [A(v_1, \ldots, v_j)]_{v_1 \ldots v_j} \equiv A(a_1, \ldots, a_j)$.[18]

The L_ω counterparts of the remaining Zermelo-Fraenkel (ZF) and von Neumann-Gödel-Bernays (GB) axioms—minus extensionality—are formulated on analogy.[19] By adding the ZF-style axioms or the GB-style axioms to T1 or T2′, we obtain the rudiments of four logics for L_ω with Δ.

Now what about the logical paradoxes? Evidently, the closest we can come to, e.g., Russell's paradox in the two ZF-style logics for L_ω with Δ is

$$[x \,\Delta\, u \,\&\, x \not\Delta\, x]^u_x \,\Delta\, [x \,\Delta\, u \,\&\, x \not\Delta\, x]^u_x$$
$$\equiv ([x \,\Delta\, u \,\&\, x \not\Delta\, x]^u_x \,\Delta\, u$$
$$\& \, [x \,\Delta\, u \,\&\, x \not\Delta\, x]^u_x \not\Delta\, [x \,\Delta\, u \,\&\, x \not\Delta\, x]^u_x)$$

from which it follows merely that

$$(\forall u)([x \,\Delta\, u \,\&\, x \not\Delta\, x]^u_x \not\Delta\, u).$$

And the closest we can come to Russell's paradox in the two GB-style logics for L_ω with Δ is

$$S([x \not\Delta\, x]_x) \supset ([x \not\Delta\, x]_x \,\Delta\, [x \not\Delta\, x]_x \equiv [x \not\Delta\, x]_x \not\Delta\, [x \not\Delta\, x]_x)$$

from which it follows merely that

$$\neg S([x \not\Delta\, x]_x).$$

Thus, we may tentatively conclude that the above logics for L_ω with Δ are free of contradiction.[20]

But what about the semantical and intentional paradoxes? To set the stage for the discussion of these paradoxes, note the following surprising fact. In each of the above logics for L_ω with Δ it is possible to define a *truth predicate T for propositions* such that the following condition of adequacy is provable for all formulas A:

$$T[A] \equiv A.$$[21]

In view of Tarski's theorem on the undefinability within a given language of a truth predicate for the sentences of that language, the definability within the logic for propositions of a truth predicate for propositions might appear paradoxical. But it is not. To get a semantical paradox something more is required, specifically, a special interpretation of L_ω. Suppose that L_ω is interpreted in such a way that one of its primitive predicates (let it be M^2) expresses the *meaning relation* for L_ω. In that case a truth predicate *Tr* for the sentences in L_ω could be defined in L_ω as follows:

$$Tr(x) \ \mathit{iff}_{\mathrm{df}} \ T((\iota y)M^2(x, y))$$

i.e.,

x is a true sentence $\mathit{iff}_{\mathrm{df}}$ what x expresses is true.

Given this definition and the above condition of adequacy for T, the following condition of adequacy for *Tr* would hold for all sentences A in L_ω:

$$Tr \ulcorner A \urcorner \equiv A.$$[22]

And this does contradict Tarski's theorem. Therefore, if L_ω can be interpreted in such a way that one of its predicates expresses the meaning relation for L_ω, the ZF-style and GB-style principles of predication must be modified further.

In a similar vein, although the above logics for L_ω with Δ are as they stand free of intentional paradoxes, intentional paradoxes can easily be manufactured by suitably interpreting L_ω and by adjoining certain empirically conceivable auxiliary premises. For example, let L_ω be interpreted so that one of its predicates expresses an intentional relation, e.g., belief. And suppose that there is someone who believes that he is sometimes mistaken but (with the possible exception of some of his beliefs that are entailed by this one together with his true beliefs) all his other beliefs are true.[23] From this supposition it is possible to derive the following logical falsehood in the above ZF and GB-style logics for L_ω with Δ:

$$xB[(\exists y)(xBy \ \& \ \neg Ty)]^x \equiv \neg xB[(\exists y)(xBy \ \& \ \neg \mathrm{Ty})]^x.$$

Hence, an intentional paradox.

Despite the ease with which semantical and intentional paradoxes seem to be generated, such paradoxes may be neatly resolved simply by further adjusting the extension of the predication relation. Take any formula A. Consider any quantified occurrence of a variable v_i in A. Suppose that this occurrence of v_i is bound by an occurrence in A of a quantifier $(\exists v_i)$ or $(\forall v_i)$ that itself is not a constituent of an occurrence in A of (the expanded form of) our definition of the truth predicate T.[24] Such occurrences of variables in A will be called *ungrounded*. Let A_u be the formula that results from restricting the range of ungrounded occurrences in A to things that have u as a property, and let such formulas A_u be called *grounded*. Now consider the following modified principles of predication:

(ZF-Style Predicative Principle of Predication)
If A_u has the form $(v_1, \ldots, v_j \,\Delta\, u \,\&\, B)$, then
$\vdash \langle v_1, \ldots, v_j \rangle \,\Delta\, [A_u]_{v_1 \ldots v_j} \equiv A_u$.

(ZF-Style Impredicative Principle of Predication)
If w is distinct from $v_1, \ldots, v_j$ and is not free in A and if A has the form $(v_1, \ldots, v_j \,\Delta\, u \,\&\, C)$, then
$\vdash (\exists w)(\langle v_1, \ldots, v_j \rangle \,\Delta\, w \equiv_{v_1 \ldots v_j} A)$.[25]

(GB-Style Predicative Principle of Predication)
If for all h, $1 \leq h \leq j$, a_h is free for v_h in A_u and conversely, then
$\vdash \langle a_1, \ldots, a_j \rangle \,\Delta\, [A_u(v_1, \ldots, v_j)]_{v_1 \ldots v_j} \equiv A(a_1, \ldots, a_j)$.[26]

(GB-Style Impredicative Principle of Predication)
If w does not occur in A, then
$\vdash (\exists w)(\langle a_1, \ldots, a_j \rangle \,\Delta\, w \equiv_{a_1 \ldots a_j} A)$.

Let the L_ω counterparts of the remaining ZF and GB axioms (minus extensionality) be formulated on analogy.[27] By adding the modified ZF-style axioms or the modified GB-style axioms to T1 or T2′, we obtain four logics for L_ω with Δ. Evidently, none of the familiar semantical or intentional paradoxes can be generated in these modified ZF-style and GB-style logics even when a univocal meaning predicate and various intentional predicates are singled out.[28] And at the same time, we still can define the univocal truth predicate T such that the following modified condition of adequacy is provable for all grounded formulas A_u:

$T[A_u] \equiv A_u$.

What is the intuitive idea behind this resolution of the semantical and intentional paradoxes? It is that in all contexts of speech and thought there is an implicit limitation *u* on the things that are taken to be relevant for consideration. That is, in all contexts of speech and thought an implicit universe of discourse *u* is invoked, where *u* is something less than the totality of all things. In a given context the identity of *u* is determined pragmatically by features of the context. The semantical and intentional paradoxes result from a failure to notice and keep track of subtle contextual shifts affecting the implicit universe of discourse.[29]

The idea that the semantical and intentional paradoxes can be resolved by making explicit contextually invoked limitations on the universe of discourse ought to sound familiar. For the ramified theory of types embodies a special case of this very idea. Indeed, the modified ZF-style and GB-style logics for L_ω with Δ may be viewed as natural generalizations of ramified type theory.[30] However, these logics for L_ω with Δ are generalizations that eliminate most of the artificiality and rigidity for which ramified type theory is notorious. The bearing this fact has on the first-order/higher-order controversy, with which this chapter has been concerned, is that ramified type theory is typically formulated as a higher-order logic. So once again the naturalness and generality of first-order logic comes through.

Is higher-order logic best viewed as a natural generalization of first-order logic, or is it best viewed as an artificially restricted theory derived within first-order logic with predication? The answer, I hope, is evident.

The proposed resolution of the semantical paradoxes depends essentially on the fact that intensional entities are the primary semantical correlates of formulas. No analogous resolution is possible if instead extensional entities—namely, sets—are identified as the primary semantical correlates.[31] This problem in set-theoretical semantics is just the beginning of the troubles for a formal philosophy based on set theory. In the next chapter we shall find many more.