Herbert Hochberg

Russell's Paradox, Russellian Relations, and the Problems of Predication and Impredicativity

Russell's paradox and the resultant distinction of logical types have been central topics of philosophical discussion for almost a century. In this essay I claim that a more fundamental distinction, that which distinguishes properties and relations as monadic, dyadic, etc., provides a basis for blocking Russell's paradox as applied to properties, not sets, without distinctions of type (or equivalent distinctions). The distinction also points to fundamental features of predication that bear on the nature of relations, the extension of relations, the problem of the analysis of relational order, and questions about symmetrical relations (identity) and purported monadic relational properties (self-identity). In particular, an early unpublished analysis of relational order that Russell proposed is examined in detail and contrasted with the standard set-theoretical analysis of Wiener and Kuratowski. I argue that the analysis of Kuratowski presupposes a basic ordering relation, and hence does not provide an analysis of order, whereas a modification of Russell's analysis provides a viable alternative analysis. The discussion of Russell's paradox and relational predication brings out a connection with Bradley's lesser-known paradox of predication, which, like the Russell paradox, is found to stem from a mistaken conception of the exemplification relation.

Though I argue that a *Principia*-style schema without type distinctions avoids the familiar paradox involving self-predication, I defend Russell's rejection of impredicative properties, and his use of a ramified theory of orders, from Ramsey's arguments attacking Russell's claim that a "vicious circle" is involved in the use of impredicative predicates. This discussion involves a consideration of Ramsey's view of quantification, which he derived from Wittgenstein. As both the ramified theory, involving Russell's axiom of reducibility, and the acceptance of impredicative predicates are found to be problematic, I conclude that complex predicate abstracts involving predicate quantifiers should not be taken to represent properties.

I. Self-predication and Paradox

Russell and others have held that the Russell paradox arises for both properties and classes, given an unrestricted comprehension rule. The paradox is sometimes mistakenly said to arise from the introduction of a certain expression into a schema. Thus, if one introduces a predicate "I" defined by:

(1) $I(f) = df \neg f(f)$

so that we have:

(2) $(f)[I(f) \equiv \neg f(f)]$,

which we instantiate to:

(3) $I(I) \equiv \neg I(I)$,

we have a contradiction. Quine has stressed the point that this is mistaken since the paradox results only when one holds that the property (or class) I exists.[1] The paradox does not arise from the mere introduction of "I" into the schema by (1). Thus, the additional assumption of "$(\exists f)(f = I)$" or "$(\exists f)(g)[f(g) \equiv \neg g(g)]$" is needed to generate (3). This, for Quine, is implicit in the instantiation from (2) to (3), since, as Quine sees it, one is committed to the existence of something if one instantiates to a sign "for it" or existentially generalizes from such a sign. The instantiation from (2) to (3) also involves letting "I" occupy subject place, even though we define the sign, in (1), only as it occurs in predicate place. This, too, fits with Quine's notions of "ontological commitment," since, for Quine, we arrive at the paradox of (3) only by means of an instantiation of (2), or an equivalent rule for substitution covering the free variable in (1), which presupposes the appropriate existential claim. Thus, Quine and others avoid the paradox by adopting a comprehension axiom warranting existential statements like "$(\exists f)(g)[f(g) \equiv \ldots g \ldots]$" if restrictions on the context "..." are satisfied. The restrictions limit instantiations to "non-problematic" predicates. Russell's theory of types amounts to such a restrictive comprehension axiom, since it excludes contexts like "$\neg f(f)$" as ill formed, hence blocking existential statements like "$(\exists f)\neg f(f)$."

The use of a definition like (1) is problematic in that while "I" is a defined predicate, it cannot be replaced in (3). This is not only due to the use of "I" in both subject and predicate place in (3), while it occurs only in predicate place in the definiendum of (1). What is also peculiar is the use of the free variable "f" in both subject and predicate place in the definiens. These peculiarities suggest that the definition in (1) is not really a definition. But we can define "I" by the use of an abstract in the form of a definite description:

(4) $I = df\,(\iota f)(g)[f(g) \equiv \neg g(g)]$.

One can now attempt to generate the Russell paradox, along lines Russell sometimes used in the exposition of it, by use of the law of excluded middle.[2] Thus,

(5) $I(I) \vee \neg I(I)$

is taken as an instance of "$p \vee \neg p$," and the paradox supposedly results since each disjunct in (5) entails its negation. But the matter is not so simple. Replacing "I" by its definition, given in (4), in "$I(I)$," we get

$$(6)\quad (\iota f)(g)[f(g) \equiv \neg g(g)]\ \{(\iota f)(g)[f(g) \equiv \neg g(g)]\},$$

with the use of braces in (6) to set off the occurrence of the description in subject place. If we retain standard features of Russell's theory of descriptions in our "type-free" notation, (6) is readily seen to entail (in fact it is equivalent to):

$$(7)\quad E!(\iota f)(g)[f(g) \equiv \neg g(g)],$$

which is the existential condition Quine takes to be necessary to generate the paradox. (6) and (7) are easily seen to be contradictions upon expansion of the descriptions. Replacing "I" by its definition in "$I(I)$" we get:

$$(8)\quad \neg(\iota f)(g)[f(g) \equiv \neg g(g)]\{(\iota f)(g)[f(g) \equiv \neg g(g)]\}.$$

By Russell's theory of descriptions we now face the familiar ambiguity of scope. If we take the scope to be secondary, then (8) may be read as the denial that there is a unique Russell property that applies to itself. Upon expansion of the description in (8), (8) is clearly not a contradiction. If we take the scope (or α scope) as primary, (8) is contradictory. But, to treat the scope as primary is not to take (8) as an instance of "$p \vee \neg p$." Thus, a well known feature of the theory of descriptions blocks the attempt to derive a paradox by the use of "$p \vee \neg p$." One may then allow the predicate "$(\iota f)(g)[f(g) \equiv \neg g(g)]$" into a type-free schema without paradox, just as one may allow "$(\iota f)(f \neq f)$" as a predicate. One can no more claim that "$E!(\iota f)(g)[f(g \equiv \neg g(g)]$" holds than one can adopt "$E!(\iota f)(f \neq f)$." However, there is a problem involved in treating "I" as an abbreviation of a definite description or lambda abstraction in a calculus where predicates are taken to represent properties, rather than classes.

The problem is easily seen in a simpler context. Let "R" and "S" be two predicates that we take to stand for properties, say being red and being square. If we limit ourselves to lower functional logic we may define a predicate "RS" by

$$(D_1)\quad RS(x) = df\, R(x)\ \&\ S(x)$$

or use an abstract

$$R(\hat{x})\ \&\ S(\hat{x})$$

so understood that

$$(D_2)\quad [R(\hat{x})\ \&\ S(\hat{x})](y) = df\, R(y)\ \&\ S(y)$$

is assumed. But, if we consider "RS" or "$R(\hat{x})\ \&\ S(\hat{x})$" in a higher functional calculus, (D_1) and (D_2) do not suffice. They do not provide for the elimination of the defined signs in all contexts where such signs occur as subject signs, as in

"$(\exists f)(f = RS)$" and "$(\exists\phi)\phi(RS)$" or in "$(\exists f)(f = [R(\hat{x})\ \&\ S(\hat{x})])$" and "$(\exists\phi)\phi[R(\hat{x})\ \&\ S(\hat{x}]$." Hence, we may use definite descriptions (or lambda abstracts) so that "RS" and "$R(\hat{x})\ \&\ S(\hat{x})$" are construed in terms of "$(\iota f)(x)[f(x) \equiv (R(x)\ \&\ S(x))]$." But, doing this presupposes that we hold that only one property has the extension had by *RS*, namely, *RS* itself. This is precisely what we should not assume if we take predicates to stand for properties, rather than classes.[3] Making such an assumption amounts to the introduction of an extensionality axiom. Even if one holds that necessary or logical equivalence guarantees identity, as Carnap once suggested, one presupposes a variant of an extensionality axiom. Aside from the familiar problems associated with such a claim, it simply amounts to a stipulation regarding "identity conditions" for complex properties. Moreover, it would not only force one to hold that there is only one tautologous and one contradictory property, but it also goes against the obvious point that complex properties with different constituents are different—$R\hat{x}\ \&\ \neg R\hat{x}$ and $G\hat{x}\ \&\ \neg G\hat{x}$, for example. To avoid making an unwarranted assumption about equivalence as a condition of identity, we should not use such definite descriptions (or lambda abstracts where such abstracts are construed "extensionally") to define expressions like "RS" or "$R(\hat{x})\ \&\ S(\hat{x})$," where the latter are taken to stand for a property: the property we would normally take as being red and square. We can then take the juxtaposition of "R" and "S" in "RS" or the use of "&" in "$R(\hat{x})\ \&\ S(\hat{x})$" to be devices for forming compound predicates, from other predicates, that represent "complex" properties. The explicit use of "&" in "$R(\hat{x})\ \&\ S(\hat{x})$," and its implicit use in "RS," is not as a truth-functional connective. Rather, "&" operates as a primitive sign that is used to form predicates from predicates. The predicates formed by its use are, then, undefined complex predicates. Hence, we require

(9) $(y)([R(\hat{x})\ \&\ S(\hat{x})](y) \equiv [R(y)\ \&\ S(y)])$

as an instance of an axiom schema governing the use of "&" *in* predicate expressions. Not having (D_1) or (D_2), (9) is not a consequence of any definitional pattern.

The same situation arises in a type-free schema in the case of contexts like

(10) $R(R)$

and

(11) $\neg R(R)$.

We can form the predicates

(12) $\hat{\phi}(\hat{\phi})$
(13) $\neg\hat{\phi}(\hat{\phi})$

by abstraction. But, there is a problem about how to treat such predicate abstracts in sentential contexts. So long as we have taken predicates to stand for properties, we cannot treat such abstracts as abbreviations of definite descriptions, such as "$(\iota f)(g)[f(g) \equiv \neg g(g)]$." Just as we may not use "$(\iota f)(x)[f(x) \equiv (R(x) \& S(x))]$" in a calculus where the predicates are taken to stand for properties, so we may not use "$(\iota f)(g)[f(g) \equiv \neg g(g)]$," and, hence, we may use neither "*I*," defined as in (4), nor (13), construed as an abbreviation for "$(\iota f)(g)[f(g) \equiv \neg g(g)]$."

(13) must then be construed as a primitive complex predicate that is not eliminable by definition. Hence, as in the case of "$R(\hat{x}) \& S(\hat{x})$," a question arises about the treatment of such abstracts in sentential contexts. Again, let us consider a simpler case. From the contexts "$(\exists x)R(x)$" and "$(\exists f)f(a)$," where "*a*" is a proper name, we may construct the abstracts "$(\exists x)\hat{f}(x)$" and "$(\exists f)f(\hat{x})$" and read them, respectively, as "applying to something" and "having some property." If we then consider the appropriate sentences used to attribute such properties to *R* and *a*, respectively, to be

$$[(\exists x)\hat{f}(x)](R)$$

and

$$[(\exists f)f(\hat{x})](a)$$

we must assume, as instances of axiom schemata for such abstracts,

$$[(\exists x)\hat{f}(x)](R) \equiv (\exists x)R(x)$$

and

$$[(\exists f)f(\hat{x})](a) \equiv (\exists f)f(a),$$

in order to have the predicate abstracts carry the "sense" intended. Therefore, it is simpler, following Russell and Whitehead in PM, to express the attribution of $(\exists x)\hat{f}(x)$ to *R* by replacing the occurrence of "*f*" by "*R*" to yield "$(\exists x)R(x)$," and employ a similar pattern in the case of "*a*" and "$(\exists f)f(\hat{x})$." This way of construing such predicate abstracts can be taken to be a reason for denying that such abstracts stand for properties if their use is limited to predicate place. Where such abstracts function as subject terms, as in "$\phi[(\exists)\hat{f}(x)]$," no such replacement is possible, just as in the earlier case of "*RS*" and "$R(\hat{x}) \& S(\hat{x})$." In the present cases we do not construct the predicate abstract by combining predicate signs, as in the case of "$R(\hat{x}) \& S(\hat{x})$," but by "abstracting from" a constant predicate (or "open sentence" if one prefers). Interestingly enough, just as one may deny that such abstracts stand for properties, if they are confined to use in predicate place, since, for example, "$[\exists(x)\hat{f}(x)](R)$" is merely an alternative rendition of "$(\exists x)R(x)$," one may also object to taking such abstracts to stand for properties if they are not so confined. For, in the latter circumstances, such abstracts must be construed as undefined but complex predicates, since they are not eliminable, as subject terms, by axiom schemata or rules of replacement. Ruling out such properties on such grounds would, of course, rule out the Russell property.[4] But ruling out the Russell property by denying that complex, yet undefined, predicates stand for properties is not a resolution I wish to pursue here. Rather, there is a more fundamental point about the abstract "$\neg\hat{\phi}(\hat{\phi})$" and the "Russell property" that will block the paradox.

Given the abstract "$\neg\hat{\phi}(\hat{\phi})$" as a sign for the "Russell property," one may then, purportedly, form

(14) $\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi})]$,

as the relevant self-predication.[5] By our understood replacement rule we would take the subject abstract to replace each token of "$\hat{\phi}$" in the predicate term and obtain

(15) $\neg[\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi})]]$,

which contradicts (14). One might object to the legitimacy of "$\hat{\phi}(\hat{\phi})$" and "$\neg\hat{\phi}(\hat{\phi})$" as predicate expressions, since their use as predicates leads to an unending application of the replacement rule. Thus, by that rule, (15) will generate another statement, equivalent to (14), that will generate yet another statement, and so on. One will never arrive at a statement where the abstract in predicate place is removed by application of the replacement rule. This is reminiscent of the peculiarity of "I," defined as in (1), as it occurs in (3) and, hence, points to another way of dismissing the Russell paradox. But there is still a more fundamental point to be made about $\neg\hat{\phi}(\hat{\phi})$. This we can see by considering the argument purportedly establishing the paradox. Supposedly, given the derivation of (15) from (14), one concludes that (14) is contradictory and hence that $\neg\hat{\phi}(\hat{\phi})$ cannot apply to itself. And, if we assume that it does not, i.e., assume (15), then by the replacement rule we obtain

(16) $\neg[\neg[\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi})]]]$,

which is the other strand of the familiar paradox. Thus, assuming our replacement rule and that either (14) or (15) holds, we seem to arrive at the familiar contradiction.

The appearance is deceiving and illustrates a crucial point about the "Russell property." Neither (15) nor (16) is obtained by a legitimate application of the replacement rule. The purported Russell property $\neg\hat{\phi}(\hat{\phi})$ is really a two-term relation, and the abstract "$\neg\hat{\phi}(\hat{\phi})$" a two-term relational predicate. Yet to arrive at (15) and (16) we replaced a monadic predicate variable abstract "$\hat{\phi}$" by a two-term relational predicate. Once again, the point is easily seen in a simpler case.

Consider the sentence "*Ra*." We form an abstract by abstraction on the sign "*a*" and obtain "$R\hat{x}$", which we may take to be another sign for the property *being an R*, i.e., *R* itself. If we abstract from the predicate sign and obtain "$\hat{\phi}a$," we get a sign for *being a property of a*. Suppose we abstract from both signs and obtain "$\hat{\phi}\hat{x}$." What do we have a sign for? One obvious answer is *the relation* exemplification that obtains between properties and objects. Thus, we can take

(17) $\hat{\phi}\hat{x}(R, a)$

to be a way of stating that a has R and, by appropriate use of the replacement rule, take (17) to yield "Ra." The crucial point is that "$\hat{\phi}\hat{x}$" is a relation sign and that we must apply the replacement rule to two sign tokens to obtain "Ra" from (17). Notice also that one replaces the token of "$\hat{\phi}$" in (17) by a monadic predicate, "R." The same situation is present in the case of "$\hat{\phi}(\hat{\phi})$" and "$\neg\hat{\phi}(\hat{\phi})$." They are relation signs, just as "$\hat{\phi}\hat{x}$" is. The Russell property is really a relation, not a monadic property. $\hat{\phi}(\hat{\phi})$ may be taken as the relation of self-exemplification for monadic properties and $\neg\hat{\phi}(\hat{\phi})$, the Russell "property," as the relation of non-self-exemplification for monadic properties. Thus, (14) is not well formed, irrespective of a theory of types, since "$\neg\hat{\phi}(\hat{\phi})$" is a relational predicate that is used as a monadic predicate in (14). In place of (14) we should seek to use

(18) $\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi}), \neg\hat{\phi}(\hat{\phi})]$

to state that the Russell relation applies to itself, i.e., that it stands in the relation to itself. One would then seek to derive the contradiction by replacing "$\neg\hat{\phi}(\hat{\phi})$" for each occurrence of "$\hat{\phi}$" in the predicate expression "$\neg\hat{\phi}(\hat{\phi})$" in (18). But, whereas "$\neg\hat{\phi}(\hat{\phi})$" is a two-term relation sign, "$\hat{\phi}$" is a monadic predicate abstract and must be replaced by a monadic predicate sign. Recall (17) and the replacement of "$\hat{\phi}$" by "R." Thus, such a replacement in (18) is illegitimate.[6] The derivation of the paradox is also blocked if we seek to use

(19) $(g)[\neg\hat{\phi}(\hat{\phi})g \equiv \neg g(g)]$

in place of the replacement rule, for (19) must be modified to

(20) $(g)[\neg\hat{\phi}(\hat{\phi})(g, g) \equiv \neg g(g)]$

to be well formed, given that "$\neg\hat{\phi}(\hat{\phi})$" is a relational predicate. But, then, we can, *at best*, only instantiate to

(21) $\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi}), \neg\hat{\phi}(\hat{\phi})] \equiv \neg[\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi})]]$

and not to

(22) $\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi})] \equiv \neg[\neg\hat{\phi}(\hat{\phi})[\neg\hat{\phi}(\hat{\phi})]]$.

But, even (21) is illegitimate, since we instantiate a relation sign for a monadic predicate variable, "g," to obtain the right-hand side of the biconditional. Thus, (21) is not only not contradictory, it is not even well formed.

Taking the Russell property as a relation, as we should, the Russell paradox does not arise for properties and relations in a type-free schema that distinguishes between monadic and relational properties and predicates in the familiar way. The paradox does arise for classes, without the familiar restrictions, and for a schema that would permit definite descriptions of properties, based on extensional conditions, as in (4), and where (7) is assumed. But the use of such descriptions for properties is problematic in its own right. Thus, for properties and relations, one may recognize self-predication and the Russell relation without paradox, if one employs a viable schema for the representation of properties. Such a schema need not employ either a version of type theory or a corresponding restricted comprehension rule. The distinction between monadic and relational properties suffices as a restriction that prohibits the paradox. Thus, we may conclude that a fundamental feature of exemplification or predication suffices to block the paradox. It is worth recalling that Russell once took the "fact" that particulars were of one kind while attributes were of logically different kinds, monadic, dyadic, etc., to be the ultimate distinguishing feature between particulars and attributes (properties).

II. Properties, Relations, and Identity

Besides the difference between classes and properties that prevents the use of a sign like "$(\iota f)(g)(f(g) \equiv \neg g(g))$" for a property, another reason the paradox arises for classes, but not for properties, is that there is no distinction for classes like that between monadic, dyadic, etc., attributes. Taken "in extension," as one says, a dyadic relation is a class, though a class of ordered pairs or, following the Wiener-Kuratowski procedure, a class of two-membered classes (whose members, in turn, are classes). Thus, one can speak of the class of classes that are not members of themselves. In the case of attributes one must speak of the dyadic relation of non-self-exemplification that monadic attributes stand in (or do not) to themselves. It is clear that a monadic Russell property does not exist since "$(\exists f)(g)(f(g) \equiv \neg g(g))$" is paradoxical. Moreover, if one claims that there is a monadic property had by all properties that stand in the Russell relation to themselves,

$$(23) \quad (\exists f)(g)(f(g) \equiv \neg \hat{\phi}\hat{\phi}(g, g)),$$

a paradox results as well, since, by the replacement rule, we arrive at "$(\exists f)(g)(f(g) \equiv \neg g(g))$" from (23). This means that the abstract "$(g)(\hat{\phi}(g) \equiv \neg g(g))$," taken as an abstract standing for a monadic property that applies to any monadic property that is a Russell property, stands for an empty property. The class of monadic Russell properties is empty. No paradox results, however, from recognizing $(g)((\hat{\phi})(g) \equiv \neg g(g))$ as an attribute (see note 5), just as no paradox results from recognizing $\neg \hat{\phi}\hat{\phi}$ as a dyadic relation.[7] Moreover, as no paradox results from "$(\exists R_2)(R_2 = \neg \hat{\phi}\hat{\phi})$" or "$(\exists R_2)(g)(R_2(g, g) \equiv \neg \hat{\phi}\hat{\phi}(g, g))$" or "$(\exists R_2)(g)(R_2(g, g) \equiv \neg g(g))$," it is clear that there are several ways in which one can say that the Russell relation exists without paradoxical consequences.

The relation $\hat{\phi}(\hat{\phi})$ poses a problem ignored in the preceding discussion. Suppose one considers "$\hat{\phi}(\hat{\Psi})$" to represent monadic exemplification for properties. Thus

(24) $\hat{\phi}(\hat{\Psi})(G, G)$

would express the self-exemplification of G, and (24) would be equivalent to "$G(G)$."

To say that G exemplifies G is to say that G stands in *the relation of exemplification* to G (itself). But if one also recognizes $\hat{\phi}(\hat{\phi})$ we have a second situation obtaining: that G stands in the relation of self-exemplification to G. The situations would be different since one involves $\hat{\phi}(\hat{\Psi})$ while the other involves $\hat{\phi}(\hat{\phi})$. Yet, both "$\hat{\phi}(\hat{\Psi})(G, G)$" and "$\hat{\phi}(\hat{\phi})(G, G)$" are equivalent to "$G(G)$," by the understood replacement rule. This points to the peculiarity of recognizing a relation of self-exemplification, as well as a relation of exemplification. It is problematic to take "$\hat{\phi}(\hat{\phi})$" to stand for a relation, since what we have when we use "$\hat{\phi}(\hat{\phi})$" in "$\hat{\phi}(\hat{\phi})(G, G)$" is simply a case of the use of "$\hat{\phi}(\hat{\Psi})$." There is no more a relation of self-exemplification in addition to that of exemplification than there is a relation of self-identity in addition to the identity relation. What can lead one to recognize such an additional relation is the confusion whereby one takes such a relation to be a monadic property, which is easily done. Thus, one may think of *self-identity* as a monadic property that an entity exemplifies, and not as a reflexive relation. Being a monadic property, it is then different from a reflexive relation. But this is mistaken. If we take "$\hat{x} = \hat{y}$" as a predicate for the identity relation, then "$\hat{x} = \hat{x}$" only appears to be a monadic predicate, since it involves two tokens of the same type. Consider "$L\hat{x}\hat{y}$" to represent the relation of being to the left of. It is clearly absurd to take "$L\hat{x}\hat{x}$" to represent the further relation, or property, of *being to the left of itself*. There are really two questions involved. First, is there an additional relation or monadic property in the case of *self-identity* and *being to the left of itself*? Second, if there is, is it a monadic property or a relation? It is clear that we do not have a monadic property. For when one asserts that an object a is not to the left of itself and is to the left of an object b, it must be that one denies that the object a stands to itself as it stands to the object b. That is, what is denied is that *the pair* (a, a) stands in the very same relation that obtains of *the pair* (a, b). But, even if it is taken to be a relation, $L\hat{x}\hat{x}$ would have to be taken to hold of the pair (a, a) only when the relation $L\hat{x}\hat{y}$ holds of the pair. To recognize $L\hat{x}\hat{x}$ in addition to $L\hat{x}\hat{y}$ is not only pointless by Occam's razor, but introduces the difficulty of forcing one to recognize the distinct situations involving the different relations, while having both "$L\hat{x}\hat{y}(a, a)$" and "$L\hat{x}\hat{x}(a, a)$" being elliptical for "Laa." But, then, we no more have "$L\hat{x}\hat{x}$" standing for a relation than for a monadic property. What goes for $L\hat{x}\hat{x}$ goes for self-identity and $\hat{\phi}(\hat{\phi})$. Thus, neither "$\hat{\phi}(\hat{\phi})$" nor "$\neg\hat{\phi}(\hat{\phi})$" stands for a relation. The purported Russell relation $\neg\hat{\phi}(\hat{\phi})$ may thus be avoided, and with it the purported paradox disappears in yet another way.

III. Relations and Order

Not acknowledging $L\hat{x}\hat{x}$, $\hat{x} = \hat{x}$, and $\hat{\phi}(\hat{\phi})$ as relations, for the preceding reasons, shows something further about relational predication. It has been common to construe properties in extension as classes. Relations, then, are taken as classes of ordered pairs or, following the Wiener-Kuratowski procedure for the construal of such pairs, as classes of classes. This means that in taking relations as classes, one either recognizes a further kind of entity – an ordered pair – or treats relations like $L\hat{x}\hat{y}$, which obtain among particulars, as higher-order classes of classes of classes, as opposed to monadic properties of objects, which become first-order classes of particulars. By so treating relations one obliterates a logical distinction between monadic and relational properties by, in effect, treating relations as monadic properties of ordered pairs. Yet, by so doing one appeals to a different logical distinction: that between the objects, particulars as opposed to pairs, or that between the order of classes relative to the particulars, since relations like $L\hat{x}\hat{y}$ become classes of classes of classes of particulars.

In the case of the use of the Wiener-Kuratowski procedure there is an interesting consequence. The pair (a, a) becomes the class $\{\{a\}\}$, a one-membered class of a unit class of a particular. This, in a way, correlates with the notion that signs like "$\hat{x} = \hat{x}$," "$L\hat{x}\hat{x}$," and "$\hat{\phi}(\hat{\phi})$" are signs for monadic properties. Yet, there is also an obvious correlate of the point that "$\hat{x} = \hat{x}$," etc., are not signs for monadic properties, for the appropriate class for a monadic property of objects a, b, etc., would be $\{a, b, \ldots, \}$, whereas the appropriate class correlated to the relational predicate "$\hat{x} = \hat{x}$" would contain $\{\{a\}\}$, $\{\{b\}\}$, etc. The difference between a and $\{\{a\}\}$ as members of the extension of predicates reflects the difference between a monadic property and a reflexive relation. Of course, if one thinks of "open sentences" being satisfied, one can speak of "$x = x$" and "Gx" both being satisfied by, say, a, and hence of classes of objects of the same kind as the extensions of both "predicates." But this already overlooks the relational form of "$x = x$" by treating it along the lines of "Gx & Hx." In both cases there are two tokens of an individual variable, but no relational predicate occurs in "Gx & Hx."

Russell was long preoccupied by the problems posed by the analysis of relational facts. One problem he was concerned with was that of the logical form of propositions. That problem would be resolved by recognizing different exemplification relations for monadic, dyadic, etc., facts. Thus, one difference between the facts expressed by "Ga" and "Lab" would be that the former involved the two-term exemplification connection expressed by "$\hat{\phi}\hat{x}$," while the latter would involve a three-term connection expressed by "$\hat{R}\hat{x}\hat{y}$." The exemplification connection would be the form of the fact. But such a form would not suffice to distinguish the facts expressed by "Lab" and "Lba." Russell's most detailed attempt to deal with the problems of relational predication occurs in the recently published *Theory of Knowledge*.[8] It has been suggested that his solution of the difficulty was similar to the Wiener-Kuratowski procedure in that he appeals to higher-order relations.[9] But this is not accurate. What Russell did was to suggest coordinating a relation, say $L\hat{x}\hat{y}$, to two relations, say L_1 and L_2, which were relations that obtained between objects, like a, and "complexes," as nonlinguistic entities. Thus, the fact or proposition[10] that-*Lab* was the fact or proposition to which a stood in the relation L_1 and to which b stood in the relation L_2. The complex that-*Lab* was thus denoted by a definite description:

$$(R_1)\ (\iota p)[(aL_1p)\ \&\ (bL_2p)].$$

Such an analysis is problematic, since it distinguishes *aLb* from *bLa* by holding that the first stands in one relation to a, while the second stands in another relation to a. This means that one assumes that one may distinguish two entities by means of relational properties. Ironically, Russell had, following Moore, earlier argued that one could not do this.

This problem is avoided by modifying Russell's analysis and doing the following. Let us take "$(L\hat{x}\hat{y}, a, b)$" to indicate a "complex" that obtains when one of the two particulars is to the left of the other, without specifying any "order." Let L_1 and L_2 be relations between a particular and the complex indicated by "$(L\hat{x}\hat{y}, a, b)$." Then

$$(R_2)\ [aL_1(L\hat{x}\hat{y}, a, b)]\ \&\ [bL_2(L\hat{x}\hat{y}, a, b)]$$

can be taken to state that a is the first element, and b the second, of an instance of $L\hat{x}\hat{y}$. Hence, "*Lab*" and (R_2) express the same proposition or situation. But (R_2), unlike "*Lab*," makes no use of order in several senses. First, conjunction is commutative.[11] Second, the sign "$(L\hat{x}\hat{y}, a, b)$," like a set sign, makes no use of the order of the terms. Thus, $(L\hat{x}\hat{y}, a, b) = (L\hat{x}\hat{y}, b, a) = (a, b, L\hat{x}\hat{y}) =$ etc. Third, given the difference in the kinds of terms of L_1 and L_2—individuals and situations or propositions—one can take "$aL_1(L\hat{x}\hat{y}, a, b)$" and "$(L\hat{x}\hat{y}, a, b)L_1a$" to state the same thing: that the particular a is the first term of an instance of a left-of relation (with b).

Russell's use of a description rather than a sign like "$(L\hat{x}\hat{y}, a, b)$" avoids the redundancy of one of the conjuncts of (R_2). (Actually, given the redundancy in (R_2) one may take "*Lab*" to be analyzed in terms of one conjunct.) But his description is peculiar. For Russell replaces a *sentence*, such as "*Lab*," by a *description*, since we understand that the sentence "*Lab*" stands for a complex that the description picks out. We, at best, have "*Lab*" as an abbreviation for the description, but it is not an abbreviation for a sentence, unless

$$(R_3)\ E!(\iota p)[(aL_1p)\ \&\ (bL_2p)]$$

is the appropriate sentence, a sentence Russell uses to assert that the complex exists. Russell's use of a description instead of an unordered complex sign like "$(L\hat{x}\hat{y}, a, b)$," may be partly motivated by an apparent circularity in the use of such a complex. He takes the relation $L\hat{x}\hat{y}$ to be determined by L_1 and L_2. And, if we think of specifying what "*Lab*" asserts in terms of (R_1) or (R_3), as opposed to (R_2), we see an apparent problem with (R_2) that is not present in (R_3) or (R_1).[12] The problem is only apparent. For "L_1" and "L_2" are used in different ways in (R_2), on the one hand, and (R_1) and (R_3) on the other. Russell's L_1 and L_2 determine $L\hat{x}\hat{y}$ as well as the order of the terms *a* and *b* in the expressed proposition or situation. By using (R_2), one takes L_1 and L_2 simply to supply the order. Thus, whereas Russell would require two completely different relations, playing the roles analogous to L_1 and L_2, in the case of another two-term relation, say *below*, "L_1" and "L_2" may be used for such a relation, as those terms are used in (R_2). As referred to in (R_2), the relations L_1 and L_2 simply determine the order of terms, and not the relation involved. Thus, there is no circularity in the use of "L_1" and "L_2" in (R_2). Rather the pattern recognizing ($L\hat{x}\hat{y}$, *a*, *b*) separates the two features involved: the content supplied by one relation rather than another *and* the ordering of the terms standing in the relation. On the pattern, L_1 and L_2 are the basis for the analysis of order in propositions (or situations or facts). Moreover, one does not appeal to Russell's vague talk of L_1 and L_2 "determining" $L\hat{x}\hat{y}$. For what this amounts to is simply the running together of the two quite different aspects of a relational fact: the content supplied by the relation and the order of the terms.

The Wiener-Kuratowski procedure appears to offer an alternative analysis of order without appealing to ordering relations like L_1 and L_2. But this is misleading. If one were to employ such a procedure in the analysis of facts or propositions, one would have to introduce higher-order classes or some correlate of such classes as constituents of facts or propositions. Either alternative is problematic. For not only are additional entities introduced, but the ordering relations are not avoided. It is easy to see why they are not. Suppose one takes $L\hat{x}\hat{y}$ to be a property of the class $\{\{a\}, \{a, b\}\}$, and thus takes the fact that $-Lab$ to be the exemplification of $L\hat{x}\hat{y}$ by that class. It is clear that two relations, say *is a member of a unit class in* and *is not a member of a unit class in*, are implicitly employed in the analysis of relational facts by the use of classes like $\{\{a\}, \{a, b\}\}$. The former relation replaces L_1 as an ordering relation, since we understand that *a* is the first element, as in the standard definition

$$\langle a, b\rangle = \text{df } \{\{a\}, \{a, b\}\}.$$

The appeal to order is not eliminated by the use of the Wiener-Kuratowski procedure. Rather, one uses a property (really a relation) like *being a member of a unit set* as an ordering property instead of something like *being the first member of a pair*. In short, what the procedure shows is that a property like *being a member of a unit set* can be used to order a pair of elements. When we have an ordered pair, $\langle a, b\rangle$, we have two elements, and one is the first of the pair and the other is the second. We express that by the linear ordering of the signs in "$\langle a, b\rangle$." With

"{ {a}, {a, b} }" we take *is the member of the unit set* to perform the role of *is the first.*[13]

We may conclude that Russell has proposed a way of analyzing relational facts that appeals to ordering relations that are not really avoided by procedures of the Wiener-Kuratowski type. On the pattern suggested here, derived from Russell's, one need not introduce classes as constituents of facts nor appeal to ordered pairs as basic entities. However, on such a pattern, the analysis of a fact or situation, such as that – Lab, involves the particulars, a and b; the relation $L\hat{x}\hat{y}$;[14] the ordering relations L_1 and L_2; the unordered compound ($L\hat{x}\hat{y}$, a, b); and compounds like $aL_1(L\hat{x}\hat{y}, a, b)$.[15]

IV. Bradley's Paradox, Russell's Paradox, and Exemplification

While Russell's paradox has preoccupied philosophers and logicians since the turn of the century, Bradley's paradox of predication has received relatively little attention. Yet it is far more threatening, since it involves the claim that predication is incoherent. There are many ways of construing the problem posed by Bradley. One way that is germane to the preceding discussion of Russell's paradox and to Russell's concern with relational predication is the following. We take an atomic sentence "Ga" to state that a particular has or exemplifies a property. The existence of the indicated fact – the particular exemplifying the property – is the truth condition for the sentence. The fact is taken to consist of the particular, a, and the property, G, in the exemplification relation, $\hat{\phi}\hat{x}$. But, supposedly, it cannot be so taken. For there must be a further constituent: a three-term exemplification connection that obtains of G, a, and $\hat{\phi}\hat{x}$. Thus, the fact must be construed to consist of a, G, and $\hat{\phi}\hat{x}$ in the three-term exemplification relation. But, then, there must be a further constituent connecting these four constituents, and so on. The supposed problem can be taken to be that to acknowledge $\hat{\phi}\hat{x}$ as a constituent of the fact that a is G is to prohibit allowing one to specify the factual truth condition for the sentence "Ga," since it is not the fact consisting of a and G in the relation $\hat{\phi}\hat{x}$. For that purported fact turns out to be a fact containing the three-term connection that, in turn, must be construed as a fact containing a four-term relation and so on. Frege, it should be noted, took Bradley's problem to be a serious problem prior to Bradley's statement of it. He suggested one type of response to it, a response that Russell adopts at some places. This response holds that properties (or at least relations) do not require connecting relations to combine with particulars. This is a, if not *the*, fundamental difference between particulars and properties. It is a logical feature of a property or relation that it provides the connecting link in a fact (or proposition). In this vein one may say that the various logical kinds of properties – monadic, dyadic relations, etc. – provide the logical form of the facts in which they attributively enter. There has been a second type of response to Bradley's problem. This is to hold that there is an exemplification connection (or many such) but that such connections need not, in turn, be related to what they connect. This singular feature of an exemplification relation is marked by classifying it as a *tie*, a *nexus*, or a *logical* relation.

The Fregean response suffers from giving each property a further, yet common, role in a fact or proposition. The alternative response faces the charge of stipulating that exemplification relations are unique in order to avoid Bradley's problem. It is thus an ad hoc solution. What I wish to argue is that the construal of exemplification as a logical form provides us with a solution that is not ad hoc.

Consider a list of primitive monadic, dyadic, etc., predicates representing properties and relations. Exemplification could not be represented by one of the basic dyadic predicates. For we must have

$$(E_1)\ Ex(G, a) \equiv Ga$$

and, in general,

$$(E_2)\ (f)(x)[Ex(f, x) \equiv fx]$$

as logical truths, with "*Ex*" representing exemplification. But there is no justification for (E_1) and (E_2) being such truths if "*Ex*" is a primitive predicate. Moreover, it is clear that one appeals to such a relation, exemplification, by the use of sentential patterns like "*fx*" and "*Ga*." In effect, sentence structure is taken to represent such a relation. Yet, there is a trap one must avoid in making such a claim. The sentence "*Ga*" represents the purported fact that *a* is *G*. How, then, does the structure of the sentence represent exemplification? The point of the question is that as "*a*" represents an object and "*G*" a property, the sentence pattern represents the fact that *a* is *G*, and not the structure of the fact. The abstraction device provides a way of representing the structure, with "$\hat{\phi}\hat{x}$" as a sign for the exemplification relation and "$\hat{\phi}\hat{x}(G, a)$" as an alternative rendition of "*Ga*." In a way, "$\hat{\phi}\hat{x}$" is not a constituent sign of "*Ga*" as "*G*" and "*a*" are, but in that we form "$\hat{\phi}\hat{x}$" by abstraction and in that "$\hat{\phi}\hat{x}(G, a)$" is an alternative rendition of "*Ga*," one can see what is meant by the claim that the sentence structure of "*Ga*" represents the exemplification relation. For it is clear that a sign like "$\hat{\phi}\hat{x}$" presupposes sentential patterns, since without such patterns one could not form "$\hat{\phi}\hat{x}$" by abstraction. Hence, without such patterns one could not have "$\hat{\phi}\hat{x}$" on a list of signs standing for relations. Thus, we already recognize and represent exemplification by the use of "*Ga*." Removing the content terms "*G*" and "*a*" from "*Ga*" we are left, by abstraction, with "$\hat{\phi}\hat{x}$," which represents the form of monadic exemplification between a particular and a property. The exemplification relation $\hat{\phi}\hat{x}$ may then justifiably be called the form of monadic atomic facts containing particulars and properties.

Bradley's paradox may then be taken to amount to the claim that to recognize the form of a fact, as a constituent of a fact, requires acknowledging that there is a further form with respect to which the first form is a constituent among constituents. Or, to put it another way, the claim is that one cannot specify the form of a fact, since whatever one takes to be the form will merely be a constituent requiring a further form. But, in view of the contrast between the form $\hat{\phi}\hat{x}$, and the content constituents, G and a, of the fact that–Ga, Bradley's problem loses its force. For clearly, there is nothing ad hoc about the difference between a form and the things, particulars and properties, that are "in" it. It is worth recalling that monadic exemplification can be represented by a sign like "$\hat{\phi}\hat{x}$" only if we already represent it by the sentential juxtaposition of the subject and predicate signs. And, if we introduce "$\hat{\phi}\hat{x}$" and, subsequently, from

$$\hat{\phi}\hat{x}(G, a)$$

form something like

$$\hat{R}(\hat{\phi}, \hat{x})$$

and subsequent abstracts, then the members of the series of sentences

$$\hat{\phi}\hat{x}(G, a)$$
$$\hat{R}(\hat{\phi}, \hat{x})[\hat{\phi}\hat{x}, G, a]$$
$$\ldots\ldots\ldots\ldots\ldots$$

all "reduce" to "Ga," by the understood replacement rule. Ironically, Bradley's purported paradox would be bothersome only if such a reductive chain were not present. One way of avoiding such a reductive chain would be to introduce "Ex" as a primitive predicate representing exemplification; but then we do not get (E_1) and (E_2) as formal truths.

There is a further irony. Even if one holds that the fact that-Ga is to be analyzed as the fact that–$\hat{\phi}\hat{x}(G, a)$, and so on, one cannot conclude that we have not specified the form of the fact that–Ga unless one also holds that we can only specify the form of a fact that is not further analyzable. Such a position involves a kind of assumption characteristic of the atomism of Russell and Wittgenstein. And it is only on such an assumption that we could make our version of the Bradley problem a paradox.

We can see another point. On the Fregean-style resolution of Bradley's problem, a property (concept) supplies the form to a proposition. This is revealed by the use of the sign "$G\hat{x}$" What then happens when a property exemplifies another? It would appear that both properties "carry" conflicting forms into the proposition or fact. Thus one can be led to deny that properties can be subjects in prepositions or facts–as Frege and Russell, at times, were led to do in their respective ways and as Wittgenstein may also have done in the *Tractatus*. Separating the form from the property or concept, as Russell did at other places, avoids this pointless problem, for the property $G\hat{x}$ can enter into the form $\hat{\phi}\hat{x}$ or the form $\hat{\Psi}(\hat{\phi})$. Or, if we have a general form for monadic predication without regard to types, say $\hat{\phi}\hat{\alpha}$, where α can be either a property or a particular, then $G\hat{x}$ can combine in such a form as either attribute or term. All the "$\hat{x}$" in "$G\hat{x}$" reveals is that the attribute referred to is a monadic attribute of particulars and not that it must be "predicated" in a fact or proposition.

Bradley's purported paradox can be seen to be a reflection of the fact that, given a procedure for producing abstracts, we can carry on the series of abstracts

$\hat{\phi}\hat{x}$
$\hat{R}\,(\hat{\phi}, \hat{x})$
$\hat{R}_3\,(\hat{R}, \hat{\phi}, \hat{x})$
....

and the series of sentences

$\hat{\phi}\hat{x}(G, a)$
$\hat{R}(\hat{\phi}, \hat{x})[\hat{\phi}\hat{x}, G, a]$
$\hat{R}_3\,(\hat{R}, \hat{\phi}, \hat{x})[\hat{R}(\hat{\phi}, \hat{x}), \hat{\phi}\hat{x}, G, a]$
....

indefinitely. But this is not paradoxical, unless one stipulates that the successive members of the series reveal the correct logical form of the previous members. What is then paradoxical is that one can never state or show the form of a fact correlated to an atomic sentence. Without such a stipulation, the series of sentences is no more paradoxical than the fact that we can generate, with an appropriate *truth* predicate, the series

Ga
T' Ga'
T' T' Ga''
....

What one may conclude is that it is pointless to introduce "$\hat{\phi}\hat{x}$," by abstraction, as a predicate expression. For, as we noted earlier, doing so presupposes the recognition of exemplification by the very use of sentential structure. Thus, the "paradoxes" of Russell and Bradley stem from a common implicit and, in a way, unrecognized assumption: that exemplification is, logically, a relation and, hence, representable by a predicate. This overlooks the fundamental difference between what is a logical form and the "objects" that may *stand in* a form. This, in turn, invites one to hold that the objects and the form stand in a form and so on. Recognizing the difference, one avoids both paradoxes while acknowledging the form $\hat{\phi}\hat{x}$. But this involves rejecting $\hat{\phi}\hat{x}$ as a relation among relations. Thus, Russell's paradox and Bradley's paradox disappear, with respect to the attribution of properties, without appealing to types or to a special relation that does not require to be related in turn.

Wittgenstein rejected Russell's paradox in the *Tractatus* by an appeal to logical form in a Fregean manner.[16] Taking the appropriate predicate form for a monadic predicate to be "ϕx" he rejected the pattern "$\phi(\phi)$," since "ϕx" could only be a subject term for a predicate whose form was "$\Psi(f\hat{x})$." Hence, as its form is "ϕx" and not "$\Psi(f\hat{x})$," no predicate of the first form can be its own argument. But Wittgenstein's claim either over-relies on the use of "ϕx" in place of "ϕ" or builds the type distinction into the predicate form. His solution either repeats Russell's in different words or simply and arbitrarily forbids the substitution of "ϕx" for "x" in "ϕx." Wittgenstein's appeal to logical form to reject Russell's paradox is thus quite different from the use of logical form in this essay. Yet, the rejection of Bradley's paradox I have advocated is in keeping with Wittgenstein's insistence that form be shown, not represented.

V. Wittgenstein, Ramsey, and the Ramified Theory of Types

So far, the discussion of Russell's paradox has focused on the role of the exemplification relation and the consequent connections of the paradox with Bradley's attack on the exemplification relation and with issues raised by a consideration of relations. By so doing, I have avoided a problem concerning versions of Russell's paradox that arise in complex contexts, and which involves further issues raised by so-called impredicative properties. Thus, consider

(25) $I(f) = df\,(\exists g)[(g = f)\ \&\ \neg f(g)]$.

To claim that "I," or the abstract "$(\exists g)[g = \hat{f})\ \&\ \neg\hat{f}(g)]$" that it abbreviates, is a relational abstract, as "$\neg\hat{\phi}\hat{\phi}$" is, poses a number of problems, though I believe one can block the ensuing paradox along such lines. Aside from that, (25) involves a context that is *impredicative* in the sense that a quantifier is used to form a sign for a property and that property is in the range of the quantifier. Thus, we are concerned with impredicative contexts relevant to Russell's ramified theory of types. I will argue (1) that Russell was correct, as opposed to Ramsey and Carnap, to rule out such impredicative properties; (2) that the ramified theory of types is problematic; and hence, (3) that contexts like "$(f)fa$" do not yield abstracts, "$(f)f\hat{x}$," which represent properties whether or not one takes the quantifier to be general or restricted by considerations of ramification. The argument for (1) will rule out contexts like (25).

Russell's ramified theory of types was intended to avoid functions like $(f)f\hat{x}$ and $(\exists f)f\hat{x}$ as well as $\neg\hat{\phi}(\hat{\phi})$ and an unrestricted truth property. Ramsey argued that the ramified theory was an unnecessary complication.[17] His argument was twofold. First, he argued that Russell's claim that a "vicious circle" was involved in the recognition of impredicative functions was unfounded. Russell had held that impredicative functions "presupposed" a totality, the functions comprising the domain of the quantifier, and hence could not belong to such a totality. Thus, they could not be elements of the domain of the quantifier. Ramsey argued that impredicative functions only apparently posed a problem due to our inability to itemize an infinite list of functions. He might have meant that it would be mistaken to think we could not determine whether a function, $(f)f\hat{x}$, applied to it. For it is obvious that we do not determine such matters by going through an infinite list. We know that $(f)f\hat{x}$ does not apply to anything, since it is a contradictory function, just as we know that $(\exists f)f\hat{x}$ applies to everything, since it is an analytic function. Ramsey clearly argued that there was nothing wrong with including an item in a totality it "presupposed." Thus, for example, in the case of the conjunction "p & q" we have an item that is logically equivalent to a "totality"

(26) $p \mathbin{\&} q \mathbin{\&} (p \mathbin{\&} q)$,

of which it is an "element." Thinking of the quantifiers in Wittgenstein's fashion, Ramsey took "$(f)f\hat{x}$" and "$(\exists f)f\hat{x}$" to represent an infinite conjunctive and disjunctive function, respectively. And, as in the case of (26) and "p & q," he saw nothing wrong with one argument of such functions being the function itself. Second, Ramsey noted no contradiction was forthcoming from permitting *impredicative* functions, as opposed to *self-predicative* functions like $\neg\hat{\phi}(\hat{\phi})$, which violated a simple type restriction.[18] Ramsey's rejection of ramified type theory has prevailed, and there has been little interest in the ramified theory of types. But, even aside from its dependence on Wittgenstein's problematic view of quantification, Ramsey's argument is not cogent. Russell's worry about a "vicious circle" is well founded.

Russell's worry is that the quantifier has to be specified as governing a domain of "objects." The issue is, then, whether impredicative functions "involving" a quantifier may belong to the domain. It is as if Russell thinks of "introducing" the quantifier with respect to a domain. Hence, nothing in that domain can be specified by use of the quantifier. To speak of "involving" and "introducing" a quantifier is vague, and unfortunately it must be. For it is not clear just what quantifiers are construed to be, as nonlinguistic items, if they are construed to be *anything* at all, and how, as nonlinguistic items, they are construed to be related to impredicative properties (as the correlates of impredicative predicates). Linguistically, the matter is clearer. A quantifier (sign) and its bound variable are constituents of the sign for the impredicative property. But leaving aside problems about the nature of quantification, we can see the force of Russell's worry in another way. Whatever we take the quantifiers to be, the quantification signs are understood in terms of their connection with the rules of universal instantiation and existential generalization. These rules are the correlates of truth tables for the connectives, for the rules codify the use and, in that sense, the "meaning" of the quantifiers (signs). Suppose we take "$(f)f\hat{x}$" and "$(\exists f)f\hat{x}$" to stand for properties—the property of having every property and the property of having a property. Then,

(27) $(\exists f)f\hat{x}(a)$

states that a has the property of having some property. The existential quantification from (27), over the predicate, would be

(28) $(\exists g)g(a)$.

But (28) is not a generalization of (27): it *is* (27), by the understood replacement rule. Thus, there is no sense to the notion of an existential generalization from "$(\exists f)f\hat{x}$" in (27). (A related point can be made about the purported universal instantiation involving "$(f)f\hat{x}$" and "$(g)g(a)$.") This points up Russell's worry. His concern can be taken to be that for (28) to be true a must have some property. Yet, if we allow "$(\exists f)f\hat{x}$" to stand for a property, satisfying that property could not be the instance that warrants the existential generalization, since to say that a has $(\exists f)f\hat{x}$ is to state the existential generalization. But, then, we do not use the quantifier in an appropriate way in going from (27) to (28). This puts quite specifically the point that functions that *presuppose* the use of a quantifier to specify them may not belong to the domain of quantification. Ramsey's argument overlooks a fundamental asymmetry between the case of the quantifiers and the case of conjunction. Moreover, if, as Ramsey suggests, "$(\exists f)f\hat{x}$" represents an infinite disjunctive function, one constituent of which is the infinite disjunctive function itself, then, like the familiar case of the label on the bottle, which contains a picture of a bottle with the label, we have an embedded infinite regress with respect to specifying the function.

There are three issues involved. First, there is a question about whether we can specify the meaning and use of the quantifier "$(\exists f)$" if we include $(\exists f)f\hat{x}$ in the domain of properties over which the quantifier ranges. I have argued that we cannot. Second, there is a question as to whether including $(\exists f)f\hat{x}$ in that domain is like the case of "p & q" and "p & q & (p & q)," where we have a constituent included in a totality to which it is equivalent. The cases are not the same for the simple reason that the specification of the meaning and use of "&" is provided by the truth table for that sign and not, as in the case of the quantifier, by an inference rule connecting the sign to a specified domain of "entities." Third, there is a question as to whether construing the quantifiers in terms of infinite disjunctive and conjunctive functions enables one to avoid the problem raised by the first question and to hold that such infinite functions can unproblematically be specified, while containing "themselves" as "p & q & (p & q)" unproblematically contains "p & q." This third question is complex. For it involves, as a first step, the construal of the quantifiers in Wittgensteinian fashion in terms of the connectives. This is, at least, problematic. It also involves the additional issue regarding the claim that there is no problem in specifying the infinite disjunctive (and conjunctive) function even if we construe the quantifiers in terms of the connectives. My concern here is with this second step and not with the general problem raised by the Wittgensteinian interpretation of quantification. The problem is not about the specification of the meaning and use of the quantifiers, since that is supposedly resolved by the Wittgensteinian move. Rather, the problem is about the specification of the infinite functions, given that such functions are elements of themselves.

In a sense Ramsey has a point. Given an infinite domain of functions or properties $F_1, F_2, \ldots, F_n, \ldots$ and an infinite conjunctive function, ϕ_1, compounded from them, we may assume that we have an infinite conjunctive function, ϕ_2, compounded from the original F_i and ϕ_1. ϕ_2 is logically equivalent (hence identical) to ϕ_1. Ramsey makes use of this point, but he does so in an illegitimate way. For he includes ϕ_1, identified with ϕ_2, among the original F_i. One can see how he may be thinking. Since "*p* & *q*" is the logical product of "*p*," "*q*," and "*p* & *q*" and hence *is* "*p* & *q* (*p* & *q*)," a conjunction can contain itself. So, if "(*f*)*fa*" is a conjunction it can contain itself. Moreover, when we consider conjunction in terms of a truth table for "*p* & *q*," it is understood that "*p*" and "*q*" may be replaced by any propositional signs, including conjunctions like "*p* & *q*" and "*p* & *q* & (*p* & *q*)." Hence, in a way, we have a domain of propositions over which "&" ranges, and that domain includes conjunctive compounds. One may then think of specifying conjunction in terms of applying to a domain that includes conjunctions, and hence applying to a totality that includes itself. But there is a significant difference in the case of Ramsey's infinite functions. *Given* an infinite conjunction *C*, which contains a conjunct *K*, we may identify *C* with *C* & *K*. The problem is with the specification of *C* if we take it to include itself as a conjunct. Ramsey cannot specify such a function in general. For he is faced with an infinite embedded and self-referential series. By contrast, we can specify both the truth functional connective expressed by "&," by the truth table, as well as the field of propositions to which it applies. Identifying "*p* & *q*" with "*p* & *q* & (*p* & *q*)" does not preclude specifying *the* conjunction. Allowing a function to be one of the original F_i over which "(*f*)" ranges and the logical product of functions compounded from the F_i does preclude specifying the function in some cases.

If one allows for functions like $(f)f\hat{x}$ and $(\exists f)f\hat{x}$, one should resort to something like ramification to avoid impredicative properties and preserve the asymmetry of the instantiation and generalization rules. But the ramified theory has an insoluble problem. Russell and Whitehead introduced the axiom of reducibility to overcome problems connected with their definition of identity as

$$x = y. =: (\phi) :\phi!x .\supset .\phi!y \; Df$$[19]

and with the need in mathematics for statements "which will usually be equivalent to what we have in mind when we (inaccurately) speak of 'all properties of x.' "[20] They were concerned with the status of the axiom as a truth of logic and with whether of not it could be deduced from other logical truths.[21] But there is a more basic problem that Wittgenstein noted in a letter to Russell:

> Your axiom of reducibility is
>
> $$:(\exists f):\phi x \equiv {}_x f!x;$$
>
> now is this not all nonsense as this proposition has only then a meaning if we can turn the ϕ into an *apparent* variable. . . . The axiom as you have put it is only a schema and the real Pp ought to be
>
> $$:(\phi):(\exists f):\phi(x) \equiv {}_x f!x,$$
>
> and where would be the use of that?[22]

Wittgenstein's point is that the axiom cannot be stated without an unrestricted quantifier that violates the restrictions of the ramified theory of types. What one can state are indefinitely many axioms (or meta-axioms or statements in a background language) for various orders of functions. It is as if one were to state in a metalanguage or background schema that there is an axiom for every order or function of the system. But this background statement involves a quantified expression not governed by "the axiom" itself. Any statement of the axiom violates the point of the ramified theory of types. Yet, without such an axiom, or something equivalent to it, the problems Russell and Whitehead noted about the ramified theory of types remain.[23]

Wittgenstein's criticism overlooks a distinction that lies behind Russell's way of stating the axiom. The distinction was based on the supposed difference between "all" and "any." As Russell put it in 1908:

> If ϕx is a propositional function, we will denote by "(x).ϕx" the proposition "ϕx is always true.". . . Then the distinction between the assertion of all values and the assertion of any is the distinction between (1) asserting (x).ϕx and (2) asserting ϕx where x is undetermined. The latter differs from the former in that it cannot be treated as one determinate proposition. . . . In the case of such variables as propositions or properties, "any value" is legitimate, though "all values" is not. Thus we may say: "p is true or false, where p is any proposition," though we can not say "all propositions are true or false." The reason is that, in the former, we merely affirm an undetermined one of the propositions of the former "p is true or false," whereas in the latter we affirm (if anything) a new proposition, different from all the proposition of the form "p is true or false." Thus we may admit "any value" of a variable in cases where "all values" would lead to reflexive fallacies.[24]

and, specifically about the axiom of reducibility, as formulated by Wittgenstein above, Russell writes:

> This is the axiom of reducibility. It states that, given any function ϕx, there is a predicative function f!x such that f!x is always equivalent to ϕx. Note that, since a proposition beginning with "(∃f)" is, by definition, the negation of one beginning with "(f)," the above axiom involves the possibility of considering "all predicative functions of x." If ϕx is *any* function of x, we can not make propositions beginning with "(ϕ)" or "(∃ϕ)," since we can not consider "all functions," but only "any function."[25]

The real problem, however, is whether Russell's distinction between "all" and "any" (and its connection with his notions of "undetermined value," "ambiguous denotation," "ambiguous statement" and "statement about an ambiguity") makes his statement of the axiom viable. What is of course specious about Russell's claim, aside from questions about his account of "denotation," is, first, the claim that we can assert an indeterminate proposition by the use of a free (real) variable, and second, the use of a free variable with the "power" of a universally quantified (apparent) variable while holding that no determinate proposition is asserted. His overlooking of Russell's discussion of "any" and "all" notwithstanding, Wittgenstein's point is well taken. Moreover, even on his own terms, Russell's view is in trouble. First, he must admit that some primitive proposition of *Principia* cannot be symbolized but must be expressed in words. And he goes on to suggest the introduction of a new symbolic device to carry the sense of the words. Thus the symbol "[ϕy]" is introduced in the primitive proposition

$$:[\phi y].\supset.(x).\phi x$$

to symbolize "ϕy is true however y may be chosen."[26] Second, Russell's discussion of the use of a free variable in the axiom of reducibility is inconsistent with the primitive proposition of *Principia* amounting to the rule of universal generalization.

> *9.13 In any assertion containing a real variable, this real variable may be turned into an apparent variable of which all possible values are asserted to satisfy the function in question.[27]

Of course one can point out that it does not apply in such a case because the resulting universal generalization is an "illegitimate" statement. But that merely points up the specious use of a free variable to state the axiom of reducibility. Given the cogency of Wittgenstein's criticism, one may conclude that with or without ramification, abstracts like "$(f)f\hat{x}$" and "$(\exists f)f\hat{x}$," involving quantifiers, should not be taken to stand for properties. Hence, we may conclude that there are no "complex properties" represented by such quantified abstracts.

We may consider paradoxes of the Russell type to be of two kinds. One kind, the "pure" paradoxes of predication, involves only negated elementary subject-predicate contexts, such as "$\neg\phi(\phi)$," for monadic properties, "$\neg$R(R,R)" for dyadic relations, and so on. These are all blocked by the same considerations leading to the construal of "$\neg\hat{\phi}(\hat{\phi})$" as, at best, a relational abstract and by the basic distinction between monadic, dyadic, etc., predicates. In this vein "$\neg\hat{R}(\hat{R},\hat{R})$" will be a three-term relational abstract, and so on. The other kind of paradox, making use of complex quantified contexts, like "$(\exists g)((g = f) \,\&\, \neg f(g))$," involves the use not only of a term in an "unstratified" context but also of a quantifier ranging over such a term. Thus, even if one could not block such versions of the paradox by construing abstracts like "$(\exists g)((g = \hat{f}) \,\&\, \neg\hat{f}(g))$" as relational abstracts (and I believe that one can block them in this way), they can be blocked by not acknowledging "impredicative" properties. Hence, assuming that one allows ramified predicate abstracts to stand for properties, which I have argued we should not do, the paradoxes arising from quantified contexts can be blocked by a variant of Russell's ramified theory of "orders," which does not make use of a type distinction as that is normally construed. Consider a familiar way of presenting the distinction between simple and ramified type theory – a way of presentation that is, though familiar, *not* an accurate account of the theory of Russell and Whitehead. One distinguishes *types* of predicates and properties – properties of O-level objects, properties of properties of O-level objects, and so on. These are properties of the first type, the second type, and so on. Within each type of property one then distinguishes orders of properties.[28] In type 1, for example, the properties of the first order are taken to be, say $f_1^{1,1}$, $f_2^{1,1}$, $f_3^{1,1}$, . . . which constitute the domain for a first type, first-order quantifier "$(f_1^{1,1})$." Then, an abstract like "$(f_1^{1,1})f_1^{1,1}(\hat{x})$" would be a predicate of order 2, type 1. Such an abstract does not represent a property in the domain of properties over which "$(f_1^{1,1})$" ranges. Modify the familiar presentation so that we do not distinguish types, but simply consider a domain of properties $G^1{}_1$, $G^1{}_2$, $G^1{}_3$, of order 1, but where we allow contexts like "$G^1{}_1(G^1{}_1)$." Yet, the quantifier "$(f^1{}_1)$," ranging over the properties (predicates) of order 1, will be taken to form an abstract "$(f^1{}_1)f^1{}_1(\hat{x})$" of order 2, and hence that abstract (predicate) will not represent a property in the domain of "$(f^1{}_1)$." Such a ramified theory of orders does not make the type distinction, but the separation of properties into orders will block the paradoxes making use of quantified contexts, as in "$(\exists g)((g = f) \,\&\, \neg f(g))$." However, such a theory, by recognizing complex properties represented by ramified abstracts, faces the problems that led to (and are involved in) the axiom of reducibility. But, whatever one might come to hold about the representation of properties by abstracts containing predicate quantifiers, the main concern of my discussion has been with what I have called the "pure" paradoxes of predication. These, I have argued, should be looked at and resolved in terms of the distinction that Russell took to be the essential demarcation between properties and particulars – the division of properties into monadic, dyadic, etc. – and the early insights Russell had regarding the need to acknowledge logical forms.[29]

Notes

1. W. V. O. Quine, *Set Theory and Its Logic* (Cambridge, MA: Harvard University Press, 1963), p. 39.

2. B. A. W. Russell, *Introduction to Mathematical Philosophy* (London, Allen & Unwin, 1953), p. 136.

3. For a detailed consideration of this point see H. Hochberg, "Properties, Abstracts, and the Axiom of Infinity," *Journal of Philosophical Logic*, 6 (1977), pp. 193-207. Russell often speaks of "propositional functions" in connection with predicate abstracts. He uses such a notion in many ways. Here I will take predicates and predicate abstracts to represent properties or attributes. In short, "propositional functions" will be construed as attributes, which fits with one of Russell's uses of "propositional function."

4. For other recent attempts to rule out the Russell paradox without recourse to type theory or the standard restrictions on a comprehension axiom see H. Castaneda, "Ontology and Grammar," *Theoria*, 42 (1976), pp. 44-93; N. Cochiarella, "Whither Russell's Paradox of Predication," in *Logic and Ontology*, ed. M. Munitz (New York: New York University Press, 1973), pp. 133-58; R. Grossmann, "Complex Properties," *Nous*, 6 (1972), pp. 153-64.

5. The abstract "$(g)(\hat{\phi}(g) \equiv \neg g(g))$" would be an abstract for a property that applied to *any* Russell property; i.e., it would apply to any property that applied to all monadic properties that did not exemplify themselves. Interestingly, the self-predication involving such an abstract leads to contradiction with the replacement rule, but the denial of such a predication does not (as in the case of (8) with secondary scope).

6. Hence, (18) is not really well formed. One cannot, then, say that the Russell relation does not (or does) apply to itself. As on the theory of types, the crucial statements are ill formed.

7. One must keep in mind that when an abstract like "$(g)(\hat{\phi}(g) \equiv \neg g(g))$" occurs in predicate place, as in "$[(g)(\hat{\phi}(g) \equiv \neg g(g))](R)$" one obtains "$(g)(R(g) \equiv \neg g(g))$" by replacing the occurrence of "$\hat{\phi}$" by "R," but when such an abstract occurs in subject place, no replacement of "$\hat{\phi}$" is permitted.

8. B. A. W. Russell, *Theory of Knowledge*, vol. 7 of *The Collected Papers of Bertrand Russell*, ed. E. Eames et al. (London: Blackwell, 1984).

9. D. Lackey, "Russell's 1913 Map of the Mind," in *The Foundations of Analytic Philosophy*, ed. P. French, et al. (Minneapolis: University of Minnesota Press, 1981), pp. 125-42.

10. Russell speaks of a "complex." What he has in mind is a fact, the existence of which provides a ground of truth for a sentence or proposition. Since the problem of order arises in the case of propositions (as nonlinguistic entities), facts, and "possibilities" or "situations," I will speak indiscriminately, in this discussion, of facts, complexes, propositions, and situations.

11. It would be foolish to object that the conjunction (R_2) makes use of order, since we cannot interchange "*a*" and "*b*" in the different conjuncts. For it is clearly not order that is made use of, as the conjuncts can be commuted (recall, also, that only one is needed). Rather, what differentiates them is the occurrence of "*a*" in one and "*b*" in the other as in "*Fa* & *Gb*" and "*Fb* & *Ga*" These latter conjunctions do not differ in the order of the terms "*a*" and "*b*." They differ in that "*a*" goes with "*F*" in one and with "*G*" in the other, and similarly for "*b*." This is not a question of the order of the terms "*a*" and "*b*" in "*Fa* & *Gb*" and "*Fb* & *Ga*."

12. Another reason for the use of a description would be an attempt to deal with the problems posed by the nonexistence of the complex in the case of a false sentence. Thus, Russell attempts to avoid the problems connected with "possible" and "negative" facts by the use of descriptions, just as he avoids a corresponding problem about nonexistent objects.

13. It is as if both the set sign "$\{\{a\}, \{a, b\}\}$" and the sign "$\langle a, b\rangle$" are used to implicitly express the fact that there is a pair of elements *and* that *a* is the first element of the pair.

14. The rejection of unordered complexes or ordering relations forces one, I believe, to hold that relational order is not analyzable and, hence, that facts like *Lab* and *Lba* cannot be distinguished by means of differing constituent entities. The difference between those facts cannot then be accounted for. (For an argument that relational facts are unanalyzable and that introducing basic ordered pairs amounts to accepting different *unanalyzed ordered complexes*, see my "Logical Form, Existence and Relational Properties," in *The Foundations of Analytic Philosophy*, pp. 215-37.) But I see no objections to acknowledging L_1, L_2, and $(L\hat{x}\hat{y}, a, b)$ other than tedious, vague, and oversimplified declarations about what is or is not given in experience. Russell, by the way, took himself to be *acquainted*

with ordering relations and logical forms. Gustav Bergmann has recently proposed an analysis of relational order based on the Wiener-Kuratowski procedure. In effect, Bergmann introduces unordered pairs, say (a, b), as complex entities, and then forms further unordered pairs, say $(a, (a, b))$, etc. While there are several problems with Bergmann's analysis (see his "Notes on Ontology" and my "Intentionality, Structure, and Bergmann's Ontology," both in *Nous*, 15 [1981], pp. 131–54 and 155–64, respectively) the fundamental defect relevant to the present discussion is his attempt to do without ordering relations, like L_1 and L_2. His analysis fails for the same reasons that the Wiener-Kuratowski procedure does not give us an analysis of order in facts, propositions, or other ordered complexes. If Bergmann were to acknowledge ordering relations, then a modification of his analysis would be similar to the modification of Russell's pattern. In place of $(L\hat{x}\hat{y}, a, b)$ we would have the unordered complex (a, b) with $aL_1(a, b)$ replacing $aL_1(L\hat{x}\hat{y}, a, b)$. Thus, the fact that-*Lab* could be analyzed in terms of $aL_1(a, b)$ & $bL_2(a, b)$ & L (a, b); (a, b) being *the term* of the predicate "*L*." This involves accepting two kinds of complexes: those like (a, b) which do not contain properties, and those like $L(a, b)$, which do.

15. Relational order is analyzed, on the modification of Russell's view, in that facts like that-Lab and that-Lba are not taken to consist of the same constituents in a "different arrangement." They contain different constituents. Relational order may be taken to be unanalyzed in that (1) basic ordering relations, L_1 and L_2, are acknowledged, (2) complexes like $(L\hat{x}\hat{y}, a, b)$ are acknowledged, and (3) the difference in kind between particulars, like a, and complexes, like $(L\hat{x}\hat{y}, a, b)$, as terms of the ordering relations L_1 and L_2, is crucial to the analysis. (As William Demopoulos put it, a kind of "type" distinction is employed.) The situation is similar to the case of predication. In one sense, a philosopher like Russell offers an analysis of predication by recognizing a fundamental relation, or tie, or logical form of exemplification as involved in a fact. In another sense, as the ontological ground of predication is taken to be a basic relation or form, it is unanalyzed.

16. See 3.333 in L. Wittgenstein, *Tractatus Logico-Philosophicus*, trans. D. F. Pears and B. F. McGuiness (London: Routledge & Kegan Paul, 1961), p. 31.

17. F. P. Ramsey, "The Foundations of Mathematics" and "Mathematical Logic" in *The Foundations of Mathematics*, ed. R. B. Braithwaite (London: Routledge, 1931), pp. 1–61 and 138–55, esp. 38–54, 77–79.

18. Ramsey, "Foundations of Mathematics," p. 41.

19. A. N. Whitehead and B. A. W. Russell, *Principia Mathematica*, 2nd ed. (Cambridge: Cambridge University Press, 1950), vol. 1, p. 57.

20. Ibid., p. 166.

21. In the second edition Whitehead and Russell suggested that the axiom of reducibility was equivalent to the assumption that "any combination or disjunction of predicates is equivalent to a single predicate" (pp. 58–59). But they noted that the "combination or disjunction is supposed to be given intentionally" (p. 59). Thus, while they were thinking somewhat in the vein of Ramsey and Wittgenstein, about the quantifiers, they were concerned about the specification of Ramsey's infinite functions.

22. Printed in L. Wittgenstein, *Notebooks, 1914–1916*, trans. G. E. M. Anscombe (New York: Harper & Row, 1969), p. 122.

23. There is an additional problem concerning whether or not the axiom of reducibility is really a metalinguistic statement about *predicates*, not functions, that specify classes. This complication I ignore.

24. B. A. W. Russell, "Mathematical Logic as Based on the Theory of Types," reprinted in *Logic and Knowledge*, ed. R. C. Marsh (New York: Macmillan, 1956), pp. 66–67. For a link with Russell's earlier account of denoting see B. A. W. Russell, *The Principles of Mathematics* (London: Allen & Unwin, 1956), p. 94.

25. "Mathematical Logic as Based on the Theory of Types," p. 87.

26. Whitehead and Russell, *Principia Mathematica*, p. 132.

27. *Ibid.*

28. For presentations along such lines see I. M. Copi, *Symbolic Logic*, 4th ed. (New York: Macmillan, 1973), p. 302; R. Carnap, *Logical Syntax of Language* (London: Routledge & Kegan Paul, 1959), p. 86; W. and M. Kneale, *The Development of Logic* (Oxford: Oxford University Press, 1962), p. 659.

29. *Principia Mathematica*, p. xix.