9

PM's Circumflex, Syntax and Philosophy of Types

Kevin C. Klement

From *The Palgrave Centenary Companion to Principia Mathematica*, edited by Nicholas Griffin and Bernard Linsky. New York: Palgrave Macmillan, 2013, pp. 218–246. (Author's typesetting with matching pagination.)

Introduction

While second-order logic has its share of proponents, and specialized forms of type theory play a role in contemporary computer science and linguistics, I think it is fair to say that there's relatively little contemporary interest in the sort of *full-blown* higher-order logic exemplified by the simple and ramified theories of types, at least for its own sake. One does not often see, for example, a new theory or development using it as its base system. The reasons for this are no doubt many. I think one major contributing factor, however, is a disconnect between the logicians who first advocated such an approach to logic and those who have been responsible for formulating it with modern standards of rigor. *Principia Mathematica* (*PM*), remains, it is fair to say, the best known exemplar of a type-theoretic approach to logic, but exactly what its type-theory is is far from agreed-upon. Whitehead and Russell are accused of unclarity, sloppiness or even outright confusion with regard to the syntax of their language, their system's axiomatic foundations, and even its philosophical justification. More recent formulations of simple and ramified type-theories, such as those in Alonzo Church's work, although formally unambiguous and irreproachable, are seen as idiosyncratic and needlessly restrictive ways of codifying the "iterative conception" of sets or classes, more of a curiosity than a genuine rival to more flexible rival ways of codifying the same conception, such as ZFC and related set theories. But this is not surprising. The modern rigorous formulations have been done in detachment from, if not complete ignorance of, the real—or Russellian, at least—philosophical motivation of type theory, and are often done in ways that obscure that motivation.

It is not uncommon, for example, to find Church's system of ramified types (or *r*-types), or something very similar, offered in place of an explanation of *PM*'s syntax (see, e.g., Church 1976, Linsky 1999, Urquhart 2003). Even if, contrary to what I shall argue, that system were equivalent to what *PM* was meant to be, or would become if properly reconstructed, to offer *only that* is unhelpful to would-be readers of *PM*. The syntax of that system is flatly *unrecognizable* in what one finds in *PM* itself. What is needed is an historically minded intepretation of the actual *PM*, but one that does not sacrifice contemporary standards of rigor. Ideally, this would consist first in a formulation of the syntax of *PM*, which, once the definitions, abbreviations and conventions adopted by Whitehead and Russell (Peano's dot notation, the "typical ambiguity" method of suppressing type-indices, and so on) were accounted for, predicts precisely why and how the actual numbered propositions of *PM* appear the way they do. This should be presented alongside a philosophical explanation of the motivation for the type hierarchy.

If successful, this project might show that even if Whitehead and Russell did not think of the formulation of a logical system and its syntax and semantics exactly the way contemporary logicians tend to, their approach had its own rhyme and reason. Contemporary practices have been shaped to a large extent by the demands of logical meta-theory. While I do not agree with those who argue that Russell's views of logic are *antithetical* somehow to the very project of logical meta-theory,[1] it was not his own focus. One cannot get very far in a metatheoretic proof without a full recursive definition of a well-formed expression of the object language, for example. If one's aim is rather to use a given object language to state and demonstrate mathematical theorems, it perhaps suffices to make the notation clear enough that the mathematical content of those proofs is not obscured, leaving enough flexibility for refinements to the syntax to be made on the fly. Indeed, I suspect that Whitehead and Russell were deliberately less than fully explicit about the details of their system in hopes that the core of their mathematical proofs could be maintained even through substitution of a different precise understanding of the type system (cf. *PM* p. vii). Nevertheless, for the most part, it is still possible to determine what they had in mind.

This project, pursued in its entirety, is a large one. My aim here is a relatively modest one of getting clear about the syntax *PM* uses for expressions (variables and perhaps other terms) of higher-order: so-called "propositional functions". My particular emphasis is on the use, or lack thereof, of the circumflex notation for function-abstraction. I argue that in many ways the notation used in *PM* is a kind of intermediate between the approach to the syntax of a theory of types found in Frege's theory of functions of different levels, involving "incomplete" expressions with different structured kinds of incompleteness, and later, more familiar, devices for function abstraction, such as the λ-abstracts of the typed λ-calculus, where one can form complete terms for functions of any type. The discussion is one small part of trying to get a better handle on the philosophical justification, or perhaps inevitability, Russell thought there was for type-theory, a discussion which has been helped immensely in recent years by the availability of the surviving pre-*PM* manuscripts. I shall take for granted the conclusion reached not only by myself (Klement 2004, Klement 2010), but also by others (e.g., Landini 1998 and Stevens 2005), that these manuscripts show that even in 1910, Russell did not understand the type-hierarchy of *PM* as a hierarchy of entities of different logical kinds, *whether* those entities are to be understood as sets or classes at various stages of the iterative hierarchy *or* as abstract attributes-in-intension, "propositional functions" understood as mind- and language-independent real things. Indeed, my aim is largely to attempt to explain how this reading is compatible with taking the syntax of *PM* to include a very limited role for *terms* apparently standing "for" propositional functions, formed with the circumflex.

In the appendices, I briefly sketch my reconstruction of the syntax and semantics of the 1910 first edition of *PM* (—I here bracket the question as to whether and to what extent the second edition is different—) though there are aspects to my reconstruction that require more commentary and justification than what's given here, particularly with regard to those aspects that go beyond the discussion of the circumflex and notation for types.

It is worth beginning our discussion with brief recaps of the contrasting approaches.

Frege's approach

On the Fregean approach, variables of different types ("levels" in Frege's vocabulary) are literally of different shapes. A term for an object or an individual is a complete syntactic unit, whether simple or complex. An expression for a first-level function can be regarded as what is obtained from a complex term for an object by removing (one or more occurrences of) a simpler part which is itself a name for an object. In this way function expressions are gappy, incomplete, or as Frege says, "unsaturated".

A second-level function expression is obtained by removing a first-level function expression from a complex term, and thus, while gappy, is not gappy in precisely the same way that a first-level function expression is. Only another incomplete expression can complete it. Compare the first-level function expression "$F(\ )$" and the second-level function expression "$(\exists x) \ldots x \ldots$".[2] The latter's argument expression, unlike the former's, must itself have a place to receive the "x". The two expressions mutually saturate. This difference is reflected even when *variables* are used in Frege's notation. For a variable for second-level concepts, Frege writes not (e.g.) "M", but "$M_\beta \ldots \beta \ldots$" (Frege 1964, §25). Thus the canonical notation for a type-2 function taking as argument a variable type-1 function is:

(1) $$M_\beta\, f(\beta)$$

If "$f(\ )$" here is instantiated to some complexly defined instance, such as "$(F(\ ) \vee G(\ ))$", the mutually saturating mark, "β" is placed inside both gaps to create:

(2) $$M_\beta(F(\beta) \vee G(\beta))$$

If "$M_\beta \ldots \beta \ldots$" is itself instantiated to a complexly defined instance, such as "$(x)(\ldots x \ldots \supset \ldots x \ldots)$", the mutual saturation takes place over the whole:

(3) $$(x)((F(x) \vee G(x)) \supset (F(x) \vee G(x)))$$

The expression (3) is of the form "$M_\beta\, f(\beta)$", but multiply so, since it can also be seen as the result of giving $f(\ )$ and $M_\beta \ldots \beta \ldots$ the values $((F(\ ) \vee G(\ )) \supset (F(\ ) \vee G(\ )))$ and $(x) \ldots x \ldots$, or the values $G(\ )$ and $(x)((F(x) \vee \ldots x \ldots) \supset (F(x) \vee \ldots x \ldots))$, respectively, instead, and so on.

This approach suggests a certain kind of philosophy for the levels hierarchy itself. There is not necessarily one privileged decomposition of an expression into function and argument. A function expression is not necessarily one unified "piece" of notation; there is no such thing as a "function term", and hence, no such thing as placing a term of the wrong sort where a term of another sort ought to go. In a sense it is not even possible to violate the grammatical restrictions between types. One just *cannot* place "$(\exists x) \ldots x \ldots$" into its own argument spot: it simply won't fit. A Fregean function is not a "thing" in the same sense as an object. While Frege does speak informally of incomplete expressions as referring to functions and concepts, as if there were such "things," only non-object things, out there to be referred to, he himself admits that this way of talking is misleading and inexact (Frege 1892). On one hand, type indices are not really necessary for Fregean notation: the different structures of the variables speak for themselves. On the other, however, since each type of variable has a different "shape", it is very difficult to describe the syntax of an infinitary hierarchy, even in a schematic way, or state replacement rules for arbitrary types uniformly. It is no coincidence then that this method is typically used only for systems restricted to second or third-order systems, not a full-blown infinite system of types.[3]

The λ-calculus approach

The approach taken with the λ-calculus is to make use of single letters for variables of any type. In the typed λ-calculus (which is our focus here), the variables are given type indices to provide restrictions on how they may combine. Thus, using o for the type of individuals, (o) for the type of (propositional) functions of individuals, and $((o))$ for functions of such functions, etc.,[4] the analogue of Frege's (1) could be written simply:

(4) $$M^{((o))}(f^{(o)})$$

Complex function expressions are represented using λ-abstracts, which, rather than containing "gaps", contain rather variables bound by a λ-operator. The type of the abstract is determined by the bound variable(s); in particular, a term of the form $\ulcorner \lambda x^{\tau}.\, \alpha \urcorner$ has type (τ). The analogue of Frege's (2) appears:

(5) $$M^{((o))}(\lambda y^{o}.(F^{(o)}(y^{o}) \vee G^{(o)}(y^{o})))$$

λ-abstracts may occur not only in argument-position, but in function-position, so giving $M^{((o))}$ here the value $\lambda g^{(o)}.\, (x^{o})(g^{(o)}(x^{o}) \supset g^{(o)}(x^{o}))$ directly yields:

(6) $$\lambda g^{(o)}.\, (x^{o})(g^{(o)}(x^{o}) \supset g^{(o)}(x^{o}))(\lambda y^{o}.(F^{(o)}(y^{o}) \vee G^{(o)}(y^{o})))$$

Here, the two values we have given "$M^{((o))}$" and "$f^{(o)}$" from (4) are still clearly recognizable, and make up discrete and unified pieces of the symbolism. This is *the* way of writing *that* type-2 function taking *that* type-1 function as argument. If I wished instead to write the value of $\lambda g^{(o)}.(x^o)((F^{(o)}(x^o) \vee g^{(o)}(x^o)) \supset (F^{(o)}(x^o) \vee g^{(o)}(x^o)))$ taking $G^{(o)}$ as argument, I would write:

(7) $\lambda g^{(o)}.(x^o)((F^{(o)}(x^o) \vee g^{(o)}(x^o)) \supset (F^{(o)}(x^o) \vee g^{(o)}(x^o)))(G^{(o)})$

To be sure, the λ-calculus contains expansion and contraction rules making (6) and (7) logically equivalent. In particular, by applying the λ-reduction[5] rule to (6), the argument to $\lambda g^{(o)}.(x^o)(g^{(o)}(x^o) \supset g^{(o)}(x^o))$ takes the place of its bound variable, and we get:

(8)
$(x^o)(\lambda y^o.(F^{(o)}(y^o) \vee G^{(o)}(y^o))(x^o) \supset \lambda y^o.(F^{(o)}(y^o) \vee G^{(o)}(y^o))(x^o))$

Applying it again to the (now) internal λ-abstracts yields:

(9) $(x^o)((F^{(o)}(x^o) \vee G^{(o)}(x^o)) \supset (F^{(o)}(x^o) \vee G^{(o)}(x^o)))$

And then, (7) can be obtained from (9) by one step of λ-expansion. Obviously, (9) is more recognizably an analogue of Frege's (3) than is either (6) or (7). Nevertheless, despite their interderivability, the difference in syntactic form suggests that there may be some semantic difference between them.

I doubt very much that Alonzo Church, who invented the notation of the λ-calculus, would claim that it is necessarily wedded to any semantic or philosophical conception of the nature of logical types.[6] Indeed, as the resulting system can easily be made to be fully equivalent and intertranslatable with Frege's, it remains perfectly open to an adherent of Frege's philosophical understanding of the levels divisions to adopt Church's notation, claiming that its use of type indices captures the unsaturated nature of functions, only in a different way. She could then fully exploit the greater ease of stating the syntactic formation and inferential replacement rules in a uniform way.[7] Nonetheless, I personally cannot help but feel that there is something inappropriate or misleading about using this sort of notation while pushing for a Fregean understanding of the motivation for the type hierarchy. In particular, because (6) contains a discrete term for $\lambda g^{(o)}.(x^o)(g^{(o)}(x^o) \supset g^{(o)}(x^o))$, one gets the impression that it is *about* this function in a way that (9) is not. And insofar as this function is something some propositions can be "about", the function must be a *thing* of some sort. It has a name after all, and while it is a violation of the type system to use this name in its own argument spot, this is not because it is somehow physically impossible to write:

$$\lambda g^{(o)}.\,(x^o)(g^{(o)}(x^o) \supset g^{(o)}(x^o))(\lambda g^{(o)}.\,(x^o)(g^{(o)}(x^o) \supset g^{(o)}(x^o)))$$

I just have. The explanation for why the same things cannot be said about *this* thing, as can be said about others, would have to be sought elsewhere.

PM's approach

4.1 The development of Russell's views

Let us return to Russell's views, working our way towards the treatment in *PM*. To best understand them, I think one must appreciate their development. In earlier work, Frege considered and rejected the Fregean doctrine of functions understood as "incomplete" or "gappy things" (e.g., Russell 1931, §482). In manuscripts from 1903–1905, published only posthumously (in Urquhart 1994), he also anticipated much of the notation and methods of the λ-calculus (irrelevant notational details notwithstanding), though eventually abandoned the approach. I have discussed his historical confrontation with both these views elsewhere (Klement 2003, Klement 2005), so here I offer only a very crude summary.

Russell's principal objection to Frege's view was its inability to explain the difference between, e.g., the functions corresponding to the open sentences "$x \leq y$" and "$x \leq x$"—if both are thought of obtained from *removing* the relata of the relation, it is natural to "think of" both as $(\ ) \leq (\ )$. Russell's conclusion in 1903 was that

> . . . in general it is impossible to define or isolate the constant element in a propositional function, since what remains, when a certain term, wherever it occurs, is left out of a proposition, is in general no discoverable kind of entity. Thus the term in question must be not simply omitted, but replaced by a *variable*. (Russell 1931, §107)

Through most[8] of the period prior to 1905, Russell seems to have conceived of variables realistically, and of propositional functions as proposition-like complexes containing variables.

This view is incompatible with thinking of propositional functions as constituents of their values.[9] The function that "x is wise" represents contains a variable, whereas the proposition "Socrates is wise" represents does not. However, early Russell did believe that there was a proposition containing and about the function itself equivalent with "Socrates is wise". In a notation used in 1903, this equivalence would have been written (e.g., Urquhart 1994, pp. 50ff.):

(10) $\acute{x}(x \text{ is wise})|\text{Socrates} \equiv \text{Socrates is wise}$

And in 1904, instead as (e.g., Urquhart 1994, pp. 128ff.):

(11) $(\hat{x} \text{ is wise})\frac{\text{Socrates}}{\hat{x}} \equiv \text{Socrates is wise}$

It was during this period that Russell came closest to anticipating the λ-calculus approach. It is easy to regard the difference between the two sides of (10) and (11) as roughly the same as the difference that would exist in the λ-calculus between "$\lambda x^o.\,\text{Wise}^{(o)}(x^o)(\text{Socrates}^o)$" and "$\text{Wise}^{(o)}(\text{Socrates}^o)$". This is of course a difference of precisely the same sort as what exists between (6), (7) and (9), for which there is no difference in Frege's syntax.

Russell's reasons for rejecting this approach ultimately are complicated, but one core difficulty is that he regarded there to be a way of occurring in a proposition—as "logical subject" or as an "individual"—that he regarded it as "self-contradictory" (Russell 1931, §§47–49, cf. Frege 1980, p. 134) to deny the capacity of any entity to occur within. If this is right, then Russell had no explanation for why it is that a propositional function whose arguments are individuals cannot take itself as argument, hence no solution to the propositional functions version of Russell's paradox. Eventually, this seems to have led him to reject a realist ontology of propositional functions as extra-linguistic entities altogether. And this in turn led him to think that terms such as "$\hat{x}$ is human", which appear to name such entities, must either be rejected altogether, or reinterpreted as "incomplete symbols", i.e., expressions that make meaningful contributions to the complete propositions in which they appear without having their own semantic values, or "things" that are their meanings. As he wrote in a pre-*PM* manuscript, "A function must be an incomplete symbol. This seems to follow from the fact that $\phi!(\phi!\hat{z})$ is nonsense" (Russell 1906, p. 498). From late 1905 to early 1907, he took the tack of eschewing such terms altogether. Rather than quantifying over propositional functions, he then quantified over propositions and made use of a four-place relation, written "$p/a; x!q$", meaning that q is just like p except containing x wherever p contains a. Then any claim one would wish to make about "$\hat{x}$ is human" could be rephrased instead making use of claims involving this relation where the first two relata are, e.g., the proposition *Socrates is human*, and Socrates, respectively. However, other paradoxes regarding propositions as logical subjects remained unsolved, and he then came to the conclusion that propositions too must not be taken as language-independent entities. This led him, in *PM*, to return partly, but not wholly, to making use of function abstracts, though, on my interpretation, he now regarded their use much differently.

4.2 PM's propositional function nominalism

On the interpretation I favor, the higher-order variables and quantifiers of *PM* are to be understood substitutionally, so that the truth-conditions of formulæ containing them are defined recursively in terms the formulæ obtained by replacing the variables with their substitution-instances, which cannot, on pain of circularity, contain bound variables of the same order (whence *PM*'s ramification). In taking this general line of interpretation, I agree with other commentators such as Gregory Landini (1998) and Graham Stevens (2005). In what follows, I shall take the correctness of this general line of interpretation for granted.

One of the virtues of the reading is that it takes Russell at his word when, post-*PM*, he wrote such things as the following. (I shall assign letters as tags to quotations in order to refer back to them later.)

> (A) . . . a propositional function in itself is nothing: it is merely a schema. (Russell 1956c, p. 234)
>
> (B) In the language of the second order, variables denote symbols, not what is symbolized . . . (Russell 1940, p. 192)
>
> (C) A propositional function is nothing but an expression. It does not, by itself, represent anything. (Russell 1958, p. 53; cf. pp. 62, 92)

A substitutional interpretation of higher-order quantifiers can explain how it is that higher-order quantifiers—"apparent variables" for propositional functions, as Russell might say—can be understood, without there being extra-linguistic "things" for them to range over. However, this leaves one more puzzle. Whitehead and Russell continue to use a notation $\ulcorner \phi\hat{x} \urcorner$, which they tell us, "means the function itself, as opposed to an ambiguous value of the function"(*PM* p. 127; cf. pp. 15, 40). This appears to be a term, but a term *for what*? How should the circumflex be understood?

Landini (1998, pp. 265f.) has taken the heroic course of arguing that the circumflex is not a term-forming operator of the official language of *PM*. On this reading, the main use of circumflex constructions is within the informal discussion of the system in the metalanguage, to speak of the system's open sentences, as opposed to an arbitrary sentence obtained by assigning values to the variables. On Landini's reading, the only terms of the object-language are variables. Ignoring for the moment the distinct use of the circumflex in expressions of the form $\ulcorner \hat{x}\, \phi x \urcorner$, used in the contextual definition of classes, one almost never finds the circumflex used in a numbered proposition of *PM*. Whitehead and Russell even remark on this:

> (D) In fact we have found it convenient and possible—except in the explanatory portions—to keep the explicit use of symbols of the type "$\phi\hat{x}$," either as constants [*e.g.* $\hat{x} = a$] or as real variables, almost entirely out of this work. (*PM* p. 19)

What gives pause, of course, are the words "convenient" and "almost", which seem to suggest that there are exceptions, or perhaps, that while the syntax does allow such expressions, it is more convenient to use distinct formulations when available. In practice, the entire apparatus of propositional functions mainly serves as a stepping stone to introduce the contextually defined notation for classes and relations-in-extensions in *PM*, and once these are in place, variables and other terms for propositional functions disappear from the remainder of *PM*.

On my own interpretation, circumflexes of this sort do have a *very narrow* role to play in the syntax of *PM*, though not nearly as much as λ-abstracts have in the λ-calculus. To fully understand their use, I think one needs to understand the way Russell speaks about "propositional functions." At times, he means only open sentences. At times, however, he seems to want to refer to the "would-be" things that open sentences would stand for *if only there were such things*. Both uses are evident in (A). It is a propositional function in the first sense that is a "mere schema" or "an incomplete symbol". It is a propositional function in the latter sense that is a "nothing", or "not a definite object", as in the following important quotation from *PM* itself:

> (E) . . . a function is essentially an ambiguity, and that, if it is to occur in a definite proposition, it must occur in such a way that the ambiguity has disappeared, and a wholly unambiguous statement has resulted. A few illustrations will make this clear. Thus "$(x).\, \phi x$," which we have already considered, is a function of $\phi\hat{x}$; as soon as $\phi\hat{x}$ is assigned, we have a definite proposition, wholly free from ambiguity. But it is obvious that we cannot substitute for the function something which is not a function: "$(x).\, \phi x$" means "ϕx in all cases," and depends for its significance on the

> fact that there are "cases" of ϕx, *i.e.* upon the ambiguity which is characteristic of a function. ... Take e.g. "x is a man," and consider "$\phi\hat{x}$ is a man." Here there is nothing to eliminate the ambiguity which constitutes $\phi\hat{x}$; there is thus nothing definite which is said to be a man. A function, in fact, is not a definite object, which could be or not be a man; it is a mere ambiguity awaiting determination, and in order that it may occur significantly it must receive the necessary determination, which obviously it does not receive if it is merely substituted for something determinate in a proposition.* [*Footnote*: * Note that statements concerning the significance of a phrase containing "$\phi\hat{z}$" concern the *symbol* "$\phi\hat{z}$," and therefore do not fall under the rule that the elimination of the functional ambiguity is necessary to significance. Significance is a property of signs. Cf. pp. 40, 41.] (*PM* pp. 47–48)

I believe Russell makes it fairly clear that he does not regard an open sentence—whether the variable is circumflected or not!—as being independently meaningful. The way in which such an expression contributes to the meaning of the complete propositions in which it appears depends on the way in which the "ambiguity" is removed.

As is clear from the footnote, it is not necessary to "eliminate the ambiguity" when the circumflex construction is used to mention the open sentence rather than use it. In the footnote, he refers back to pages 40–41, where it is claimed that "$\phi\hat{x}$ is a function," is an unambiguous statement, and that "the value for $\phi\hat{z}$ with the argument $\phi\hat{z}$ is true," is a meaningful, albeit false, proposition. I take it that *these* uses of the circumflex are to be taken as mentions rather than uses. However, they are not found in the technical portions of *PM*, only in the informal discussion. Indeed, he there makes it clear that "$\phi(\phi\hat{z})$" is not to be interpreted this way.

What Russell has in mind by eliminating the ambiguity, when it *is* necessary to do so, is exemplified by such things as binding the variable in the open sentence with a quantifier. The ways in which this can be done are quite limited, as he makes clear in a later work:

> (F) We do not need to ask, or attempt to answer, the question: "What *is* a propositional function?" A propositional function standing all alone may be taken to be a mere schema, a mere shell, an empty receptacle for meaning, not something already significant. ...
>
> There are, in the last analysis, only two things that can be done with a propositional function: one is to assert that it is true in *all* cases, the other to assert that it is true in at least one case, or in *some* cases ... All the other uses of propositional functions can be reduced to these two. (Russell 1919, pp. 157–58; cf. Russell 1956c, p. 230, Moore forthcoming)

The claim that there are only two things is, of course, an exaggeration. Elsewhere he is a bit more generous, including also asserting a particular value of the propositional function (e.g., Russell 1950, p. 377, Russell 1958, p. 62). But the more important part of what Russell is saying here is that he thinks that all uses of propositional functions are derived somehow from these. For example, the following represent different assertions which use the propositional function "*x* is wise", and while none of them simply are universal or existential quantification applied with this open sentence, it is easy to see the dependence:

(12) $(x)(x \text{ is wise} \supset a \text{ is wise})$

(13) $(\exists x)(x \text{ is wise} \,.\, x \neq a)$

(14) $((\exists x)\, x \text{ is wise}) \,.\, (\sim (x)\, x \text{ is wise})$

These are all cases in which the "ambiguity" has been removed, and not coincidentally, there is no need to use a circumflex in any of them. The propositions expressed by (12)–(14) are not about some entity whose name is "$\hat{x}$ is human"; and indeed, Russell's explicit view is that the above, if true, are not made true by any simple fact or complex, but by numerous complexes, whose constituents include Wisdom, the entity *a*, and various *values* of the variable (*PM* p. 46). To be sure, when explaining quantification informally, he sometimes paraphrases, e.g., "$(\exists x)\, x$ is wise" by something like "the propositional function '*x* is wise' is sometimes true" (e.g., Russell 1919, p. 159) as if it was *about* the open sentence or some other "thing", but I think this is loose talk phrased for the benefit of an audience that, in general, would have had almost no familiarity with modern quantificational logic. The paraphrase gives the truth-conditions for the sentence, but is not quite synonymous with it. Indeed, Russell explicitly says that we "more correctly" speak of "functions of functions" than "statements about functions" (Russell 1919, p. 186). The examples (12)–(14) represent values of second-order functions when "*x* is wise" is the argument, but are not *about* that function.

It is clear I think that Russell believes that the circumflex notation is not needed at all when the ambiguity involved in the *use* of a propositional function is removed, as with (12)–(14) above, and that, in 1910 at least, Russell would not regard there as being any need for distinct forms in which that notation reappears such as, e.g.:

(15) $(x)((\hat{z} \text{ is wise})\frac{x}{\hat{z}} \supset (\hat{z} \text{ is wise})\frac{a}{\hat{z}})$

(16) $(\exists x)((\hat{z} \text{ is wise})\frac{x}{\hat{z}} \,.\, x \neq a)$

(17) $((\exists x)((\hat{z} \text{ is wise})\frac{x}{\hat{z}})) \,.\, (\sim (x)((\hat{z} \text{ is wise})\frac{x}{\hat{z}}))$

Notation such as this was regarded by him as appropriate only when he thought there was such a "thing" denoted by "$\hat{x}$ is wise". This is roughly to say that Russell would not have gone along with the proliferation of different inter-convertible or equivalent notations found in the λ-calculus. If propositional functions always must appear in such a way that they "receive the necessary determination" to "eliminate the ambiguity which constitutes" them, and doing so always takes a form in which a circumflex construction is not needed, it is hard to see why the circumflected expressions are needed at all in the notation of *PM*.

4.3 The role of the circumflex

The answer to this riddle is that there is sometimes need for stating general definitions, and asserting general truths, employing propositional functions, in which it is not specified precisely how that ambiguity is removed. Indeed, this is explained by Whitehead and Russell themselves in the context immediately preceding (D):

> (G) In the definition of "$\sim\{(x).\,\phi x\}$" only the function considered, namely $\phi\hat{z}$, is a real variable; thus so far as concerns the rule in question, x need not appear on the left. But when a real variable is a function, it is necessary to indicate how the argument is to be supplied, and therefore there are objections to omitting an apparent variable where (as in the case before us) this is the argument to the function which is the real variable. This appears more plainly if, instead of a general function, $\phi\hat{x}$, we take some particular function, say "$\hat{x} = a$," and consider the definition of
>
> $$\sim\{(x).\, x = a\} \;.\!=.\; (\exists x).\sim(x = a) \quad \text{Df.}$$
>
> But if we had adopted a notation in which the ambiguous value, "$x = a$," containing the apparent variable x, did not occur in the *definiendum*, we should have had to construct a notation employing the function itself, namely "$\hat{x} = a$." This does not involve an apparent variable, but would be clumsy in practice. (*PM* p. 19)

To understand what is being said here, it is perhaps best to take examples. If one wanted, for example, to introduce a definition or abbreviation for the type-2 function applied to "x is wise" in (12), one might write:

(18) $$(Qx)\,\phi x\ .{=}.\ (x)(\phi x \supset \phi a)\ \ \text{Df.}$$

Here the defined sign is used as a variable-binding operator and goes along with a variable "x" which can occupy the argument position to the letter "ϕ", used to make the definition general. This is possible here, because the definiens fills the argument positions of "ϕ" in similar fashion. This is not always possible, especially if the definition is of a more complicated sort. Here we may consider the contextual definition of class-abstracts given in *PM* *20.01 (involving the other use of the circumflex).

(*20.01) $$f(\hat{z}\,\psi z)\ .{=}{:}\ (\exists\phi){:}\,\phi!x\ .{\equiv_x}.\ \psi x : f(\phi!\hat{z})\ \ \text{Df.}$$

In the definiendum here, the argument-place to "ψ" is filled by the variable z, which is effectively bound in that context. However, if this definition were unpacked, and the definiendum replaced with the definiens, the precise way in which the argument place of the variable letter "$\phi!$" is filled depends on the context $f(\ )$ in which the class abstract appears. If one were applying this definition in the context, "$a\ \epsilon\ \hat{z}\,\psi z$", then, given additionally the definition of ϵ (*20.02), one would end up with "$(\exists\phi){:}\,\phi!x\ .{\equiv_x}.\ \psi x\ :\ \phi!a$", but if one were to apply it rather to "$(y).\,y\ \epsilon\ \hat{z}\,\psi z$", the result would be "$(y)\,(\exists\phi){:}\,\phi!x\ .{\equiv_x}.\ \psi x\ :\ \phi!y$", and of course much more complicated forms are possible too. The point is that when a particular context $f(\ )$ is supplied, the ambiguity characteristic of the function should be removed, but since the definition is so general that it is to be applicable *however* this is done, it is stated in a way that requires the circumflex term, so as not to presuppose some particular way of removing that ambiguity.

By itself, the use of the circumflex as in *20.01 is consistent with Landini's position that the circumflex is not needed in the object language of *PM* at all. This requires taking a certain sort of stance on the use of the letter f there. On the surface this appears to be a higher-type object-language variable, but it is also not fixed as to its order. Similar uses of an order-unspecified variable are found, for example, in the Axiom of Reducibility, which, in the monadic case, appears:

(*12.1) $$(\exists f){:}\,\phi x\ .{\equiv_x}.\ f!x$$

The purpose of the axiom is to assert that "[a]ny function of one argument ... is equivalent to a predicative function of the same argument or arguments." The shriek "!" on the variable f in *12.1 indicates that it is restricted to predicative order, but to serve the work for which it is intended, ϕ would seem to be able to take values of any order consistent with its having x as an argument. Of such order-unspecified variables as ϕ in *12.1 and f in *20.01, Russell writes the following:

> (H) We require, however, a means of symbolizing a function whose order is not assigned. We shall use "ϕx" or "$f(\chi!\hat{z})$" etc. to express a function (ϕ or f) whose order, relative to its argument, is not given. Such a function cannot be made into an apparent [i.e., bound] variable, unless we suppose its order previously fixed. As the only purpose of the notation is to avoid fixing the order, such a function will not be used as an apparent variable; the only functions which will be so used will be predicative functions ... (*PM* p. 165)

Something very similar occurs in the summary of *PM* Russell sent to Carnap,[10] where after listing the Axiom of Reducibility, Russell states, "[a] predicative function can be an apparent variable; a general function cannot", contrasting the two kinds of variables the axiom contains. The suggestion seems to be that the language contains certain variables which can only be used unbound, to express that a certain something holds of *any* value of that variable, where those values are not restricted to an order. The necessity of the order hierarchy prevents that sort of variable being used with quantifiers, or to express something about *all* values (cf. Russell 1956b, pp. 67ff.). This is a fairly difficult position to wrap one's head around, and indeed, even Wittgenstein chided Russell about this in a letter in 1913 (Wittgenstein 1979, p. 122). Landini interprets these "general variables" as best understood not as object-language variables at all, but as schematic letters, so that, for any open sentence one might substitute for ϕx, we have a distinct instance of *12.1, where *12.1 should not be understood as a single axiom, but as an axiom schema. While it is probably not the case that Russell himself had a clear understanding of the difference between object-language unbindable variables and schematic letters of the metalanguage, Landini's suggestion is a charitable one, consistent with the actual uses of these variables in *PM*. In fact, the suggestion makes better sense of some of Russell's actual practice than his "official" explanation. (We shall see an example later.) Indeed, most likely, I think Russell would have welcomed the re-description of his use of this variables had the more contemporary vocabularly been available to him.

If we regard the $f(\ )$ in the definition *20.01 as a schematic letter, it raises the possibility that the circumflex notation—used there because it is unclear how the ambiguity is removed—is only necessary because the definition is stated schematically. It seems in fact that most of the ways the $f(\ )$ can be assigned result in the disappearance of the circumflex, at least when all defined expressions are fully resolved. The definition *20.01 is often resolved in a context in which the apparent class-term to be eliminated appears on the right side of the membership sign ϵ, itself defined as follows:

(*20.02) $$x \epsilon (\phi ! \hat{z}) \ .{=}.\ \phi ! x \quad \text{Df.}$$

This is an odd definition in that the definiendum is syntactically more complex than the definiens, and only the definiendum contains the circumflex notation. I surmise that this definition is written in the way it is precisely to make it easier to apply the contextual definition *20.01 to contexts in which the circumflex notation otherwise would not appear. As we have seen, when both definitions are unpacked, "$a \ \epsilon \ \hat{z}\, \psi z$" becomes "$(\exists \phi) : \phi ! x \ .{\equiv}_x .\ \psi x : \phi ! a$", where no circumflex is necessary at all. Similarly, interpreting the contextual definition with narrow scope, "$(y)(y \ \epsilon \ \hat{z}\, \psi z)$", would become "$(y)\,(\exists \phi) : \phi ! x \ .{\equiv}_x .\ \psi x : \phi ! y$", and "$(\exists y)(y \ \epsilon \ \hat{z}\, \psi z)$" turns into "$(\exists y)\,(\exists \phi) : \phi ! x \ .{\equiv}_x .\ \psi x : \phi ! y$". Given that Russell believes that the only things that can be "done with" a propositional function are to assert its truth for all or some values, or assert some particular values, or something constructible somehow from these, one might think that any possible context the schematic variable "$f(\ )$" could represent would be one in which no circumflex is necessary. To my knowledge, *20.01 and *20.02, their immediate corollaries, their analogues for resolving the definition for relations-in-extension into statements about dyadic functions (*21.01 and *21.02), and their immediate corollaries, are the only numbered propositions in which what appear to be circumflex terms for propositional functions appear at all explicitly in *PM*. If the circumflex only appears there because of their schematic nature, perhaps the circumflex notation is not needed in the object-language at all.

Unfortunately, this is not quite right, since in addition to such schematic letters as "$f(\ )$" from *20.01, *PM* does have object-language variables for higher-type functions, and unlike Frege, Russell does not require these variables to themselves work as variable-binding operators. Thus

*20.01 must be applicable when $\ulcorner f(\hat{z}\,\psi z)\urcorner$ takes such a form as "$\chi!(\hat{z}\,\psi z)$", where "χ" is a variable letter of the object language, and indeed such contexts are necessary to interpret statements about classes of classes. Unpacking the contextual definition *20.01 in the following:

$$(\exists\chi)\,\chi!(\hat{z}\,\psi z) \tag{19}$$

yields:

$$(\exists\chi)((\exists\phi)\colon \phi!x\ .{\equiv_x}.\ \psi x : \chi!(\phi!\hat{z})) \tag{20}$$

I take (20) to be an object-language formula of *PM*, completely free of schematic letters, but here the circumflex is still necessary. I do not, however, take a statement such as (20) to be a serious counterexample to Russell's claims about "the only" things that can be done with propositional functions, nor in any way to require us to posit some "thing" that the term "$\phi!\hat{z}$" stands for to understand its semantics. This is because, on the reading given here, the higher-type variables are given a substitutional semantics, so that the truth of (20) depends recursively on the truth of such formulæ as "$(\exists\phi)\colon \phi!x\ .{\equiv_x}.\ \psi x : \phi!a$", "$(\exists\phi)\colon \phi!x\ .{\equiv_x}.\ \psi x : (y)\,\phi!y$", and "$(\exists\phi)\colon \phi!x\ .{\equiv_x}.\ \psi x : (\exists y)\,\phi!y$". It is true just in case one (or more) of those are true. Being taken as an argument to a higher-type function is not some wholly "additional" thing that can be done with the functions that are the values of $\phi!\hat{z}$ here, but just a way of making generalizations about those things we have already acknowledged can be done.

In sum, while I do think the circumflex notation does have a role to play in the object language of *PM*, this is only in situations in which the circumflected term appears as an argument to a higher-type propositional function *variable*. In such circumstances, moreover, the higher-type function variable must be part of some quantified statement whose truth or falsity depends upon statements in which that particular circumflex term disappears. In particular, the circumflected variable is replaced by either a particular value or a variable bound by a universal or existential quantifier. In no case are we required to assign some one "thing" as "the" semantic value of the circumflex abstract. For this to hold good of the example (20) it must be that none of the formulæ its truth depends upon are ones in which "$\phi!\hat{x}$" occurs in argument position to *another* higher-type variable. The formulæ it depends upon, I take it, are those that result from assigning "χ" a value that yields a closed sentence. Because "χ" is predicative, however, none of these instances can contain quantifiers binding variables of the right type to take "$\phi!\hat{x}$" as argument (and instances using unbound variables of that type are not closed sentences).

Moreover, I agree with Landini that contrary to the traditional reading of ramified type theory, in the language of *PM*, aside from the "general variables" discussed earlier (here interpreted schematically), all higher-type variables are predicative. We have already seen in quotation (H) above that Russell claims that only predicative variables would be used in *PM* as apparent (bound) variables (and cf. Russell 1956b, p. 87). Just before this passage, it is claimed that "non-predicative functions always result from such as are predicative by means of generalization" and that "therefore, it is possible without loss of generality, to use no apparent variables except such as are predicative" (*PM* p. 165). In the introduction, he writes that "we need not introduce as variables any functions except predicative functions" (*PM* p. 54), a remark that confirms that at some level Russell did not regard the order-unspecified "general variables" as real variables of the object language. With the Axiom of Reducibility assumed, quantification just over predicative variables is, for extensional contexts, nearly as good as quantification unrestricted to order, and for mathematical contexts, nothing else is needed. Still, it is perhaps arguable that these remarks are not *conclusive*: to claim that no bound non-predicative function variables will be used is not necessarily to claim that they are somehow not even part of the language. Nevertheless, as Landini has explained (Landini 1998, chap. 10), the explanation of the type system, and even conventions regarding certain contextual definitions in *PM*, are simply inadequate if non-predicative variables are included as well. This by itself seems to me to be an indication that *likely* the language wasn't intended to include them. While these considerations may not be conclusive, and explorations of a reconstructed system allowing such variables is welcome, it simplifies matters greatly to restrict the language to predicative higher-type variables.

4.4 Is the circumflex a complex term-forming operator?

I have argued that the circumflex is only used when the function abstract appears in argument position to a higher-type *variable*. Landini's position seems to be that the circumflex is only used when one function variable appears in subject position to another (higher-type) function. He claims in footnote that "[t]his use is quite minor and is easily omitted" (1998, p. 265). The thought, I take it, is that so long as type indices were specifically listed, the portion "$\chi!(\phi!\hat{z})$" of (20) could be simplified to something such as "$\chi^{((o))}(\phi^{(o)})$". This goes hand in hand with his claim that the only terms of *PM* are variables, and that "comprehension" for propositional functions is effected not through having complex circumflex terms represent valid substituends of function variables, but rather through the Axiom of Reducibility itself. This, as we have seen, is to be understood schematically, and posits a value among those quantified over by predicate function variables of a certain type one coextensive with any open formula, whether or not that open formula is predicative. Landini points out that problems exist if the circumflex is understood as a term-forming operator applicable to complex formulæ. On the present interpretation, one significant problem is that if circumflection is allowed on formulæ containing quantifiers, then even if all variables are of predicative types, some circumflex terms would be non-predicative, but yet if circumflex terms are only used in argument position to higher-type variables, without higher-type variables with non-predicative arguments, there could be no complete formulæ for them to appear within.

My own view is the intermediate one that circumflection is allowed when applied to a variable occurring within a quantifier-free open expression, but that these expressions only occur within a complete well-formed expression in argument position to a higher-type variable. It should be recalled that the strategy of the austere quantification theory of *9 of *PM* is, strictly speaking, only to allow quantifiers at the beginning of a formula, exploiting the possibility of writing any expression in which quantifiers appear subordinate to a truth-functional connective in an equivalent form with the quantifiers in what these days we would call prenex normal form. The definitional conventions of *9.01–*9.08 of *PM* allow us to interpret quantifiers when directly subordinate to a negation sign or disjunction sign, but they provide no interpretation of something of the form "$\chi!((\psi)\,\psi!(\hat{x}))$". But I do think Russell would allow for circumflection to be used as forming the expression for a complex argument to a higher-type function variable when no quantifiers are used internally, e.g., something such as:

(21) $$(\chi)(\chi!(\psi!\hat{x} \vee \phi!\hat{x}) \supset \chi!(\psi!\hat{x} \vee \phi!\hat{x}))$$

I take it Landini would not acknowledge even such uses. The evidence for Russell's acceptance of them is, admittedly, not entirely conclusive, but Russell does in informal discussion speak of, e.g., "$\phi\hat{x} \vee \psi\hat{x}$" (*9.61) and "$\sim f\hat{x} \vee \chi(\hat{x}, \hat{y})$" (*10.221). Moreover, Russell's acknowledgement in 1919 that *PM* had erred in not stating replacement rules for the variables (Russell 1919, p. 151n) suggests to some extent that Reducibility alone is not entirely responsible for functional "comprehension" in *PM*. If such complex function abstracts or function "terms" (using that phrase loosely) are allowed only before quantifiers are added, there need be no worries about non-predicative abstracts.

However, this difference between myself and Landini is partly indicative of a wider difference involving whether or not *PM* was committed to the presence in the language of non-logical constants and how this relates to the semantics of individual variables and predicative second-order variables. While we agree upon reading most higher-order quantification in *PM* substitutionally, Landini advocates a realist interpretation of the first-order variables, and even predicative lowest-type function variables, to serve as the basis for the recursive theory of truth in which the truth or falsity of higher-order variables is to be defined (Landini 1998, p. 238). On my own view, all forms of quantification in *PM* are to be understood substitutionally. Even though no non-logical vocabulary is actually *employed* within the strictly logical and mathematical content of the book, the semantics BR has in mind presupposes a base language—indeed, a logically ideal language—in which every individual has a proper name, and every simple universal (quality or relation) is represented by a simple predicate. This comes out, for example, in *Introduction to Mathematical Philosophy* where Russell claims that non-logical words are needed "for giving values to the variables" (Russell 1919, p. 201) in mathematics and pure logic even though the non-logical words are not actually used in these fields.[11]

Although I think that *ultimately*, the truth of falsity of any statement of the language of *PM* that includes the circumflex depends recursively on those that do not, this dependence proceeds in stages, and to make sense of the semantics of a "purely logical" proposition such as:

(22) $$(\phi)\,(\exists\chi)\,\chi!(\phi!\hat{x})$$

one must understand it as true just in case *all* such propositions as:

(23) $$(\exists\chi)\,\chi!(\hat{x} \text{ is green} \vee \hat{x} \text{ is blue})$$

and similar instances, are true. To be sure, (23) in turn ultimately depends for its truth on circumflex-free formulæ such as "(a is green $\vee$ a is blue) $\supset$ (b is green $\vee$ b is blue)", and so on, but one still must acknowledge the well-formedness of (23) in order to understand the semantics of (22). Landini, who would not understand the semantics of the quantifier (ϕ) at the start of (22) in quite this way, needs make no such acknowledgement. A full explanation for my deviance from Landini on this point must be left for another occasion.

4.5 Comparison with the other approaches

So to return to the discussion of the ways in which Russell's approach falls in between Frege's approach and the approach embracing explicit function abstracts, as in the λ-calculus, the following conclusions seem apt. Russell's approach is similar to the latter in that it does make use of function abstracts, and rejects the notion of a function as something "incomplete" or "gappy". Russell's "$\chi!(\phi!\hat{x})$" segments cleanly into function expression and argument expression; the "$\hat{x}$" goes with the ϕ, unlike Frege's $M_\beta f(\beta)$, where both β's are to be considered as belonging with the M, not with the f. However, unlike in the λ-calculus, the circumflex "terms" are allowed *only* in argument position to a higher-type variable. For this reason, Russell's notation does not have the multiplicity of equivalent forms as exemplified by (6), (7) and (9) in the λ-calculus. It is a well known result in typed λ-calculi that every well-formed expression is equivalent to one in so-called "normal form", i.e., one in which λ-abstracts appear only in argument position (Hindley and Seldin 1986, pp. 323ff.) or not at all; with the examples just given, this is (9). In some ways, it is as if Russell's notation only included the λ-calculi's normal forms; when a higher-type variable in function position (not just argument position) is given a particular value, the necessary conversion to reach this reduced form is applied at the same time. A further subtle difference is that Russell's notation does not use the circumflex notation when a specific higher-type function takes a propositional function as argument, as with universal or existential quantification; one writes "$(x)\,\phi x$", not "$(x)\,\phi\hat{x}$", whereas in Church's notation, "$(x^o)\,\phi^{(o)}(x^o)$" abbreviates "$\Pi^{((o))}(\lambda x^o.\,\phi^{(o)}(x^o))$".

This suggests a somewhat different attitude in how it is natural to understand the syntax of circumflex "terms" for Russell; these terms are not to be understood as terms for something, and different types of circumflex terms are not to be understood as referring to distinct logical categories of entities. As Russell put it in quotation (E), a propositional function is "not a definite object," but "a mere ambiguity awaiting determination," and indeed, the circumflex is only necessary when it is not known how that ambiguity is removed, because different ways of removing it—represented with a schematic letter or variable—are allowed in the context. The different types making up the type theory correspond to how different ways of removing ambiguities pair up with ambiguities to be removed. The quantifier "(x)" can be applied to the open sentence "x is human" because it makes the resulting sentence's truth depend recursively on the substitutions of actual designating terms for the "x" in the sentence, terms which do stand for a thing that may or may not be human. The quantifier "(ϕ)" cannot be used to remove its ambiguity, since the substitution instances for its variable are themselves things whose ambiguity needs to be removed by the context into which they are placed; if they are placed in a context which a independent semantically meaningful term is necessary, nothing meaningful results.

These different sorts of ambiguities and ambiguity-removing potentials are shown to some extent by the differing structural complexities of the variables of different types in Russell's notation. While an individual variable "x" is syntactically simple, Russell makes it clear that the full variable is not simply "f" or "χ", but "$f!\hat{x}$" for a predicative variable the next type up, or "$\chi!(\hat{\phi}!\hat{x})$" for two types up (e.g., *PM* pp. 51, 65). Again, this makes Russell's notation somewhat more like Frege's, where, as noted earlier, explicit type-indices are not required since the differing structure of the variables may speak for themselves, and unlike the λ-calculus, where type indices are crucial. However, differences between Frege's notation and Russell's make the latter not fully unambiguous without type indices. This can be seen, for example, with such an expression as the following:

$$\theta!(\chi!(\hat{x}, \psi!(\hat{y}, \hat{z}))) \tag{24}$$

Even ignoring the practice of typical ambiguity that in some contexts might allow "x", "y" and "z" to be other than variables for individuals, and insisting that they be taken as such, and taking "ψ" as a variable for a two-place function whose arguments are both individuals, this notation does not make clear what types "θ" and "χ" must be taken as having. On one reading, "χ" would be read as a variable for a function having as its arguments one individual, and one function of two individuals, and then "θ" could be understood as a function taking a function of individuals as argument. Then the whole expression's primary form equivalent in the λ-calculus might be:

$$\theta^{((o))}(\lambda x^o . \chi^{(o,(o,o))}(x^o, \lambda y^o, z^o . \psi^{(o,o)}(y^o, z^o))) \tag{25}$$

If "χ" is read instead as a variable for functions taking one individual and one function of individuals as argument, "θ" is instead taken as taking for its argument a relation between individuals. It might, for example, be read as the analogue of this:

(26) $$\theta^{((o,o))}(\lambda x^{o}, y^{o}.\, \chi^{(o,(o))}(x^{o}, \lambda z^{o}.\, \psi^{(o,o)}(y^{o}, z^{o})))$$

If something more like (26) is meant, the addition of type indices to Russell's (24) can be used to rule out reading it as akin to (25):

(27) $$\theta^{((o,o))}!(\chi^{(o,(o))}!(\hat{x}^{o}, \psi^{(o,o)}!(\hat{y}^{o}, \hat{z}^{o})))$$

Nevertheless, *even this* does not suffice to eliminate all ambiguities. It instead might be read as the analogue of:

(28) $$\theta^{((o,o))}(\lambda x^{o}, z^{o}.\, \chi^{(o,(o))}(x^{o}, \lambda y^{o}.\, \psi^{(o,o)}(y^{o}, z^{o})))$$

Given the paucity of contexts in which Russell felt this kind of notation actually needed to be used in *PM*, and that he never used multiple circumflexes explicitly in the same formula, it is likely that he ignored such complications.

Nevertheless, those wishing to state *PM* as a formal system up to modern standards of rigor cannot ignore such problems. My reconstruction of the syntax of *PM* therefore, uses both type indices and preceding λ's as variable-binding operators, but such expressions can occur as parts of well-formed formulæ *only* when in argument position to a higher-type variable. Conventions could be adopted for suppressing type indices and for substituting the circumflex notation in circumstances in which no ambiguity threatens, thereby preserving the explicitly written formulæ of *PM* as matching the syntactic rules given here.

Appendices

A. Syntax of PM

In this appendix, I sketch briefly and without further lengthy justification a reconstruction of the syntax of *PM* as I interpret it. At times I deviate slightly from the letter of *PM* for the sake of exactness, but I acknowledge when this is the case. In the present context, I list only the syntax of the more austere quantification apparatus in *9 of *PM*.

As alluded to in section 4.4, I believe *PM* presupposes a base language, including a list of proper names, one for each individual, along with simple predicates; these go together to form atomic propositions. This vocabulary of course is never used in the purely logical and mathematical content of the book, and hence is not listed in any detail. When examples are given, they are simply written out in English in sentences such as "this is red" and "Socrates is human" (*PM* pp. 39, 50). To systematize things somewhat, however, I stipulate that non-logical constants should take the form of single boldface letters. Predicates are written at the end of the sentence, which I do only to discourage the common misconception that Russell in any way thinks of the subject/predicate relationship as a form of the function/argument relationship, which I have argued against elsewhere (Klement 2004). This notation is also inspired by Russell's own occasional depiction of the basic form of subject/predicate sentences as "*xP*", which he does even in works in which the notation "'$f(x)$" is used elsewhere for propositional functions (e.g., Russell 1973, pp. 295, 305). I also assume here that there are denumerably many such constants, though I think Russell's actual intention is that there should be exactly as many constants as there are things, period. I call the letters usable as predicates "universal constants", since Russell understands them as standing for universals. At the time of the first edition of *PM*, Russell still regarded a word for the universal which one given predicate stands for as also capable of occurring in subject position to another predicate or even to itself; that is a view Russell had endorsed in *The Principles of Mathematica*, and did not abandon until having been influenced by Wittgenstein.[12] Thus, in my notation an atomic formula has such a form as:

$$\langle \mathbf{a}, \mathbf{b} \rangle \mathbf{r}$$

meaning, "*a* bears *r* to *b*", or even:

$$\langle \mathbf{r}, \mathbf{b} \rangle \mathbf{r}$$

i.e., "*r* bears itself to *b*". Hence, we can offer the following definitions:

> *A particular constant* is a bold-faced letter $\mathbf{a}, \ldots, \mathbf{e}$, written optionally with one or more primes (to ensure an infinite supply), e.g., $\mathbf{a}'$, $\mathbf{b}''$, etc.
> *A universal constant* is a bold-faced letter $\mathbf{f}, \ldots, \mathbf{z}$, with a superscript n where $n \geq 1$, written optionally with one or more primes.
> *A constant* is either a particular constant, or a universal constant.

On my interpretation, all simple constants, both those for particulars and those for universals, represent individuals, which goes along with Russell's definition of an individual as something "destitute of complexity" (1956b, p. 76) or which "exists on its own account" (*PM* p. 162).

Variables, however, can have other types, and for reasons explained above, type-symbols are used in my reconstruction.

A type symbol is defined recursively as follows: (i) o is a type symbol; (ii) if $\tau_1, \ldots, \tau_n$ are type symbols, then so is $(\tau_1, \ldots, \tau_n)$. Nothing else is a type-symbol.

The order of a type symbol τ is the greatest integer reached while parsing τ, adding one for each left parenthesis, subtracting one for each right parenthesis.

A variable letter is a (non-bold, italicized) letter $a, \ldots, o, q, r, t, \ldots, z, A, G, H, K, \ldots, Z, \phi, \psi, \chi$, or θ,[13] written optionally with one or more primes $'$ (to insure an infinite supply), and a superscripted type-symbol. (E.g., a'^{o}, $f^{(o)}$, $\theta''^{(o,\,(o))}$, etc.)

However, it is made clear in *PM* that a variable for a higher-type is best not treated as a single letter, but as something as structurally complex as its values (Landini's simplifications notwithstanding). Thus, a pure type (o) variable should appear as "$\phi!(\hat{x})$", and one of type ((o)) as "$\chi!(\hat{\phi}!(\hat{x}))$". The definition that I think would have been offered by Whitehead and Russell for a pure variable might read thus:

A pure variable of type τ is defined recursively as follows:

1. A variable letter with type-symbol o is a pure variable of type o.
2. If $\mu_1^{\tau_1}!(\nu_{1.1}, \ldots, \nu_{1.n_1}), \ldots, \mu_m^{\tau_m}!(\nu_{m.1}, \ldots, \nu_{m.n_m})$ are each pure variables of types $\tau_1, \ldots, \tau_m$, respectively, where $\mu_i^{\tau_i}$ (where $1 \leq i \leq m$) is the variable letter with which each begins, and $\kappa^{(\tau_1,\ldots,\tau_m)}$ is a variable letter with the indicated type-superscript, then $\kappa^{(\tau_1,\ldots,\tau_m)}!(\hat{\mu}_1^{\tau_1}!(\nu_{1.1}, \ldots, \nu_{1.n_1}), \ldots, \hat{\mu}_m^{\tau_m}!(\nu_{m.1}, \ldots, \nu_{m.n_m}))$ is a pure variable of type $(\tau_1, \ldots, \tau_m)$.
3. Nothing else is a pure variable.

Moving to the use of λ-operators instead, however, this could be changed to the following:

A pure variable of type τ is defined recursively as follows:

1. A variable letter with type-symbol o is a pure variable of type o.
2. If $\mu_1^{\tau_1}, \ldots, \mu_m^{\tau_m}$ are each pure variables of the types given by their superscripts, and $\kappa^{(\tau_1,\ldots,\tau_m)}$ is a variable letter with superscript $(\tau_1, \ldots, \tau_m)$, then $\lambda\mu_1^{\tau_1}, \ldots, \mu_m^{\tau_m} \,.\, \kappa^{(\tau_1,\ldots,\tau_m)}(\mu_1^{\tau_1}, \ldots, \mu_m^{\tau_m})$ is a pure variable of type $(\tau_1, \ldots, \tau_m)$.
3. Nothing else is a pure variable.

An individual variable is a pure variable of type o.

Strictly speaking these last two definitions are not needed for the definitions below, but are conceptually interesting.

In the approach of *9, quantifiers always appear at the very beginning of a formula, and they are applied to what Russell calls a "matrix" (*PM* p. 163). Vacuous variable binding is not allowed, so a matrix always contains free variables. I use the phrase "matrix-formula" as a generic term for a quantifier-free expression that may or may not contain free variables. We define this notion alongside the notion of a complex expression of a given type (a circumflex or λ-term) as follows:

A matrix-formula and *expression of type* τ are defined together, recursively, as follows:

1. A constant is an expression of type *o*.
2. A pure variable of type τ is an expression of type τ.[14]
3. If ρ^n is a universal constant with superscript *n*, and $\eta_1, \ldots, \eta_n$ are each either constants or individual variables, then $\langle \eta_1, \ldots, \eta_n \rangle \rho^n$ is a matrix-formula.
4. If $\mu^{(\tau_1, \ldots, \tau_n)}$ is a variable letter with the indicated superscript, and $\nu^{\tau_1}, \ldots, \nu^{\tau_n}$ are expressions of the given types, then $\mu^{(\tau_1, \ldots, \tau_n)}!(\nu^{\tau_1}, \ldots, \nu^{\tau_n})$ is a matrix-formula.
5. If α is a matrix-formula containing *distinct* free (uncircumflected) variable letters $\mu_1^{\tau_1}, \ldots, \mu_n^{\tau_n}$ with the indicated superscripts, and α' differs from α in placing circumflexes over every occurrence of each of $\mu_1^{\tau_1}, \ldots, \mu_n^{\tau_n}$ in α, then α' is an expression of type $(\tau_1, \ldots, \tau_n)$.
   (In the reconstructed syntax, this reads instead: if α is a matrix-formula containing *distinct* free variable letters $\mu_1^{\tau_1}, \ldots, \mu_n^{\tau_n}$ with the indicated superscripts, then $\lambda\mu_1^{\tau_1}, \ldots, \mu_n^{\tau_n} . \alpha$ is an expression of type $(\tau_1, \ldots, \tau_n)$.)
6. If α and β are matrix-formulæ then $(\alpha \vee \beta)$ is a matrix-formula.
7. If α is a matrix-formula, then $\sim\alpha$ is a matrix-formula.
8. Nothing else is either a matrix-formula or an expression of type τ.

A matrix is a matrix-formula containing one or more free (uncircumflected)[15] variable-letters.

The order of a matrix is the number one greater than the largest of the orders of the type-symbols of the (uncircumflected)[16] variable letters it contains.

An elementary proposition is a matrix-formula that does not contain any variable-letters.

An atomic proposition is an elementary proposition not containing "$\sim$" or "$\vee$" (i.e., one of the form $\langle \eta_1, \ldots, \eta_n \rangle \rho^n$).

A formula is defined recursively as follows:

1. A matrix-formula is a formula. (All uncircumflected variable letters it contains, if any, are free.)
2. If α is a formula containing variable letter μ free, then $(\mu)\,\alpha$ is a formula. (Its free variable letters are those of α minus μ.)
3. If α is a formula containing variable letter μ free, then $(\exists\mu)\,\alpha$ is a formula. (Its free variable letters are those of α minus μ.)
4. Nothing else is a formula.

A closed formula or proposition is a formula without free variables.

The order of a proposition is 0, if it is an elementary proposition, or else it is the order of the matrix that occurs after all its quantifiers.

B. Semantics of PM

I here offer a preliminary outline of what I take the intended semantics of *PM* to be based on the introduction to *PM*'s brief discussion of a hierarchy of truth and falsity for formulæ of different orders. It is intended as a rough sketch only, and in particular, significant elaborations would be necessary to do full justice to the expanded quantificational language adopted in *10. Due to an unfortunate mismatch in Russell's own description, "first truth" was used for elementary propositions, "second truth" for first-order quantified propositions, "third third" for second-order quantified propositions (*PM* p. 42). To avoid confusion, I shift the ordinals on senses of truth so they match the ordinals for the orders.

The meaning of each constant is given by fiat.

- The meaning of a universal constant with superscript *n* is some *n*-place relation between individuals. (Russell still at this time believed that relations-in-intension, etc., are themselves individuals.[17])
- The meaning of a particular constant is some particular.

Atomic truth: An atomic proposition $\langle \eta_1, \ldots, \eta_n \rangle \rho^n$ (where $\eta_1, \ldots, \eta_n$ are universal or particular constants and ρ^n is a universal constant) has *atomic truth* if and only if the individuals meant by $\eta_1, \ldots, \eta_n$ stand in the relation meant by ρ^n.

Elementary truth or 0^{th} *truth* is defined recursively as follows:

1. An atomic proposition has elementary truth just in case it has atomic truth.

2. $\sim\alpha$ has elementary truth just in case α does not have elementary truth.
3. $(\alpha \vee \beta)$ has elementary truth just in case either (or both) of α and β have elementary truth.

Truth conditions for quantified propositions can only be rigorously stated when the reconstructed syntax with λ-operators is used. In the metalanguage, I use the notation "$\mathcal{Reduce}[\lambda\xi.\,\alpha|\beta]$" for the λ-calculus expression in normal form reached by applying λ-contraction to $\lambda\xi.\,\alpha(\beta)$ as many times as is required to reach normal form. Notice that while $\lambda\xi.\,\alpha(\beta)$ may not even be a well-formed formula of *PM*'s reconstructed syntax, $\mathcal{Reduce}[\lambda\xi.\,\alpha|\beta]$ will be (presuming $\lambda\xi.\,\alpha$ and β are themselves expressions of the appropriate types).

$(n+1)^{st}$ *truth* is defined in terms of $n^{\mathbf{th}}$ **truth**, with an additional level of recursion for the number of quantifiers the formula contains counting right to left beginning with the first one using a variable letter with order n. We shall annotate this as $(n+1)^{\text{th}}/k$ for order $(n+1)^{\text{th}}$ truth for a proposition of order $n+1$ with at most k such quantifiers. (A formula is said to have $(n+1)^{\text{th}}$ truth *simpliciter* just in case it has $(n+1)^{\text{th}}/k$ truth, where k is the number of such quantifiers it contains.)

1. α has $(n+1)^{\text{th}}/0$ truth just in case it has n^{th} truth.
2. $(\xi^{\tau})\,\alpha$, where ξ^{τ} has type symbol τ (of at most order n), has $(n+1)^{\text{th}}/(k+1)$ truth just in case for every closed expression β of type τ, $\mathcal{Reduce}[\lambda\xi.\,\alpha|\beta]$ has $(n+1)^{\text{th}}/k$ truth.
3. $(\exists\xi^{\tau})\,\alpha$, where ξ^{τ} has type symbol τ (of at most order n), has $(n+1)^{\text{th}}/(k+1)$ truth just in case for at least one closed expression β of type τ, $\mathcal{Reduce}[\lambda\xi.\,\alpha|\beta]$ has $(n+1)^{\text{th}}/k$ truth.

Logical truth characterizes those propositions that would have the kind of truth appropriate for their order regardless of which atomic propositions are or true and which false (and possibly, regardless of the make-up of the list of constants).

Notes

[1]See e.g., van Heijenoort 1967 and Goldfarb 1979; for responses, see Landini 1998, and Proops 2007.

[2]Here, and throughout, I substitute Russellian notation (also used by Church) for quantification and truth functional operators for Frege's to facilitate comparison. One should not forget the other differences, however, such as Frege's use of a function calculus rather than a predicate calculus.

[3]The closest I know of of an attempt to state rigorously a full-blown higher-order logic using such a notation is found in Bostock 1974, although it does not fully live up to modern standards of precision.

[4]I am intentionally modifying the type notation of Church's work (Church 1940), again to facilitate comparison.

[5]This is often called β-reduction. It seemed better in this context to make do with more intuitive, if less precise, vocabulary.

[6]For what it's worth, Church called Frege's doctrine of the incompleteness of functions "problematical" (Church 1951b, p. 101), and he explored both intensional logics in which (6), (7) and (9) would be regarded as intensional synonyms and those in which they wouldn't (Alternatives (1) and (0), respectively, in Church 1951a).

[7]Indeed, there are historical links between Church's initial work on the λ-calculus and the early attempts at exactly stating the replacement rule for higher-order logic; see Klement 2003.

[8]His views changed often during this period; I am generalizing for the sake of concision. He did consider other views, particularly in 1903.

[9]It is also incompatible with a view that equates propositional functions with simple concepts, predicates or universals, as is still too often assumed in discussions of Russell.

[10]Thanks to Bernard Linsky, who is preparing these notes for publication, for bringing the passage to my attention and providing the precise wording.

[11]For further discussion, see Klement 2010 and Klement 2012.

[12]This is explicit at (Russell 1956c, p. 205), and also in an unpublished letter to Moore dated 2 Oct. 1922. For discussion see Klement 2004.

[13]I here omit the letters which are given constant meanings; cf. *PM* p. 5. In my exposition, Greek letters not on this list are used in the metalanguage for quantifying over object-language expressions, as with the use of "τ" for arbitrary type symbols in the preceding definition.

[14]Apart from individual variables, this step is actually redundant with the other rules, but is worth stating on its own for clarity.

[15]The qualification is not technically necessary, since a matrix-formula cannot contain circumflected variable letters without containing uncircumflected variables as well. Indeed, it must contain an uncircumflected variable letter of a higher order than any circumflected variable it contains.

[16]Again, the qualification is not necessary. See note 15.

[17]I have argued in favor of this elsewhere (Klement 2004).

References

Bostock, David, 1974. *Numbers and Arithmetic*. Oxford: Clarendon Press.

Church, Alonzo, 1940. "A Formulation of the Simple Theory of Types." *Journal of Symbolic Logic* 5: 56–68.

——, 1951a. "A Formulation of the Logic of Sense and Denotation." In *Structure, Method and Meaning: Essays in Honor of Henry M. Sheffer*, edited by P. Henle, H. Kallen, and S. Langer, pp. 3–24. New York: Liberal Arts Press.

——, 1951b. "The Need for Abstract Entities in Semantic Analysis." *Proceedings of the American Academy of Arts and Sciences (Daedalus)* 80: 100–112.

——, 1976. "Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski." *Journal of Symbolic Logic* 41(4): 747–760.

Frege, Gottlob, 1892. "On Concept and Object." In *Collected Papers on Mathematics, Logic and Philosophy*, edited by Brian McGuinness, pp. 182–94. New York: Basil Blackwell.

——, 1964. *Basic Laws of Arithmetic*. Berkeley: University of California Press. Translated by M. Furth; originally published in 1893–1902 as *Grundgesetze der Arithmetik*, 2 vols. (Jena: H. Pohle).

——, 1980. *Philosophical and Mathematical Correspondence*. Chicago: University of Chicago Press. Edited by Hans Kaal.

Goldfarb, Warren, 1979. "Logic in the Twenties: The Nature of the Quantifier." *Journal of Symbolic Logic* 44: 351–68.

Hindley, J. Roger and Jonathan P. Seldin, 1986. *Introduction to Combinators and λ-Calculus*. Cambridge: Cambridge University Press.

Klement, Kevin C., 2003. "Russell's 1903-05 Anticipation of the Lambda Calculus." *History and Philosophy of Logic* 24: 15–37.

——, 2004. "Putting Form Before Function: Logical Grammar in Frege, Russell and Wittgenstein." *Philosopher's Imprint* 4: 1–47.

——, 2005. "The Origins of the Propositional Functions Version of Russell's Paradox." *Russell* 24: 101–32.

——, 2010. "The Functions of Russell's No Class Theory." *Review of Symbolic Logic* 3: 633–664.

——, 2012. “Review of G. Landini, *Russell*.” *Journal of the History of Analytical Philosophy* 1: 21–28.

Landini, Gregory, 1998. *Russell’s Hidden Substitutional Theory*. Oxford: Oxford University Press.

Linsky, Bernard, 1999. *Russell’s Metaphysical Logic*. Stanford: CSLI Publications.

Moore, G. E., forthcoming. “Russell on Philosophy of Mathematics: Notes 1911–12.” Edited by James Levine.

Proops, Ian, 2007. “Russell and the Universalist Conception of Logic.” *Noûs* 41: 1–32.

Russell, Bertrand, 1906. “Types.” In *Collected Papers of Bertrand Russell*, vol. 5, edited by Gregory Moore, pp. 498–514. London: Routledge, 2014.

——, 1919. *Introduction to Mathematical Philosophy*. London: George Allen & Unwin.

——, 1931. *The Principles of Mathematics*. Cambridge: Cambridge University Press. (First edition 1903).

——, 1940. *Inquiry into Meaning and Truth*. London: Allen and Unwin.

——, 1950. “Logical Positivism.” In Russell (1956a), pp. 365–382. (Written 1950).

——, 1956a. *Logic and Knowledge*. London: Allen and Unwin. Edited by R. C. Marsh.

——, 1956b. “Mathematical Logic as Based on the Theory of Types.” In Russell (1956a), pp. 57–102. (First published 1908).

——, 1956c. “The Philosophy of Logical Atomism.” In Russell (1956a), pp. 175–281. (First published 1918).

——, 1958. *My Philosophical Development*. London: Allen and Unwin.

——, 1973. “Is Mathematics Purely Linguistic?” In *Essays in Analysis*, ed. D. Lackey, pp. 295–306. New York: George Braziller. (Written 1951–52).

Stevens, Graham, 2005. *The Russellian Origins of Analytical Philosophy*. London: Routledge.

Urquhart, Alasdair, ed., 1994. *The Collected Papers of Bertrand Russell*, vol. 4. London: Routledge.

Urquhart, Alasdair, 2003. “The Theory of Types.” In *The Cambridge Companion to Bertrand Russell*, edited by Nicholas Griffin, pp. 286–309. Cambridge: Cambidge University Press.

van Heijenoort, Jean, 1967. “Logic as Calculus and Logic as Language.” *Synthese* 17: 324–330.

Whitehead, A. N. and Bertrand Russell, 1925–1927. *Principia Mathematica,* 3 vols. Cambridge: Cambridge University Press. (First edition 1910–1913).

Wittgenstein, Ludwig, 1979. *Notebooks 1914–1916*. Chicago: University of Chicago Press, 2nd ed. Ed. G. H. von Wright and G. E. M. Anscombe.