Notre Dame Journal of Formal Logic
Volume 28, Number 1, January 1987

On the Brink of a Paradox?

I. SUSAN RUSSINOFF*

In his introduction to the translation of Frege's *Begriffsschrift* published in *From Frege to Gödel: A Source Book in Mathematical Logic*, Jean van Heijenoort writes ([5], p. 3):

> . . . If we also observe that in the derivation of formula (91) he substitutes $\mathfrak{F}$ for f, we see that [Frege] is on the brink of a paradox.

What this means may, at the moment, be obscure, but it is my aim to illuminate this passage and as I continue its meaning will become clear. van Heijenoort's claim that Frege is "on the brink of a paradox" is, of course, metaphorical, for Frege's system either leads to a paradox or it does not. Terrell Bynum, in his edition of the *Begriffsschrift*, maintains that no paradox can be generated in Frege's system. He writes ([1], p. 182):

> Van Heijenoort is in error in supposing that any paradox can arise from the deductive procedure Frege uses here.

In this paper I attempt to resolve this dispute. In the first section I reconstruct the system of Frege's *Begriffsschrift* and show that the reconstructed system is equivalent to a standard second-order predicate calculus, and then demonstrate the consistency of the reconstructed system.[1] I conclude, then, that if van Heijenoort is claiming that the system leads to a paradox or inconsistency, the dispute is settled on the side of Bynum. In the second part I consider the interpretation of the system of the *Begriffsschrift.* Frege is not clear about how the system is to be interpreted. In light of Frege's later writings on the distinction between function and object, the interpretation of second-order quantifications presents some difficulties for Frege. These difficulties may be seen as an anticipation of the well-known problem with the concept *horse*. The nature of

*I am grateful to George Boolos, Richard Cartwright, and an anonymous referee for helpful comments and discussion.

Received March 3, 1986; revised August 11, 1986

the anticipation will be detailed later. Moreover, certain claims implicit in Frege's way of interpreting his system lead to Russell's paradox. Hence, I shall argue that while the system is demonstrably consistent, the problems that arise in its interpretation may provide a basis for van Heijenoort's concern.

Frege's *Begriffsschrift* presents an axiomatized system of logic whose primitive connectives are the conditional and negation and whose rules of inference are modus ponens and substitution. Frege later draws an important and celebrated distinction between function and object, but in the *Begriffsschrift* the distinction is between function and argument. He writes ([5], p. 11):

> If in an expression, whose content need not be capable of becoming a judgment, a simple or compound sign has one or more occurrences and if we regard that sign as replaceable in all or some of these occurrences by something else (but everywhere by the same thing), then we call the part that remains invariant in the expression of a *function*, and the replaceable part the argument of the function.

Frege's way of drawing the distinction is syntactic. It should be noted that his terminology is not standard. At this point he takes sentence parts to *be* functions, yet later he takes sentence parts to *stand for* functions. In the *Begriffsschrift* Frege regards sentences as functions of the expressions occurring in them. He goes on to distinguish one-place from two-place functions and introduces appropriate notation. Predicates are treated as one-place functions and relational expressions as two-place functions ([5], p. 23):

> In order to express an indeterminate function of the argument A, we write A, enclosed in parentheses, to the right of a letter, for example
>
> $$\Phi(A)$$
>
> likewise,
>
> $$\Psi(A, B)$$
>
> means a function of the two arguments A and B that is not determined any further. . . . In general
>
> $$\Psi(A, B)$$
>
> differs from
>
> $$\Psi(B, A).$$

Frege's distinction between function and argument leads him to treat "functions" as arguments. As we shall see shortly, this seems to provide the basis for Frege's use of second-order quantification, and may be responsible for van Heijenoort's worry.

Following his discussion of the function, Frege introduces his notion of generality as follows ([5], p. 24):

> In the expression of a judgment we can always regard the combination of signs to the right of $\vdash$ as a function of the signs occurring in it.

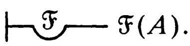
> If we replace this argument by a German letter and if in the content stroke we introduce a concavity with this German letter in it, as in
>
> 
>
> this stands for the judgment that whatever we may take for its argument, the function is a fact.

Frege goes on to introduce what is, in effect, the second-order quantifier. In the same place he writes:

> Since a letter used as a sign for a function, such as Φ in $\Phi(A)$ can itself be regarded as the argument of a function, its place can be taken in the manner just specified by a German letter.

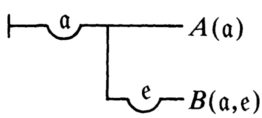
Thus $\Phi(A)$ yields both

and

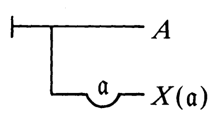


Frege defines the existential quantifier in terms of the universal quantifier and negation in the usual manner.

The concavity in the content stroke indicates the scope of the German letter in it. For instance, the scope of '$\mathfrak{a}$' in



is the entire content of the judgment, while in



the scope is confined to the antecedent of the conditional. Italic Latin letters are used when the scope of the generality is the entire content. Given this, Frege introduces a principle of universal generalization which I shall discuss a bit later.

In Part II of the *Begriffsschrift* Frege begins to formulate his axiomatic system and derive some of its theorems. He draws a distinction between his rules or "modes" of inference and his axioms, or "rules of pure thought". He presents a total of nine axioms and derives more complex judgments in accordance with his rules of inference. At the start of Part II ([5], pp. 28–29) Frege writes:

> Merely to know the laws is obviously not the same as to know them together with the connection that some have to others. In this way we arrive at a small number of laws in which, if we add those contained in the rules, the content of all the laws is included, albeit in an undeveloped state.

The judgments of Parts II and III are numbered. The nine laws, or axioms, are (1), (2), (8), (28), (31), (41), (52), (54), and (58).

Axioms (1), (2), (8), (31), and (41) together with modus ponens and Frege's unstated substitution rule constitute a complete axiomatization of the propositional calculus (as proved by Łukasiewicz in 1931). Kneale showed that the nine axioms together with modus ponens, universal generalization, and confinement constitute a complete first-order predicate calculus. I shall show that Frege's nine axioms together with modus ponens, universal generalization, confinement, and his rules for substitution constitute a consistent second-order predicate calculus. Frege's substitution rules are not explicitly stated. I shall provide a formulation of these rules later.

The system to which I shall now show Frege's to be equivalent is Joel Robbin's as presented in [4], chapter six. The symbols of Robbin's system are individual variables and an infinity of *n*-place relation variables of degree *n*. His system contains a standard complete set of axioms for the propositional calculus to which is adjoined:

(A4) $\forall vA \supset S^v_a|$ *where v is any variable and a is a variable or constant of the same type (that is, an individual variable or constant if v is an individual variable and an n-place relation variable or constant if v is an n-place variable) and a is free for v in A.*
(A5) $\forall v(A \supset B) \supset (A \supset \forall vB)$ *where v has no free occurrence in A.*
(A6) $\exists F \forall x_1 \forall x_2 \ldots \forall x_n\ (F(x_1, x_2, \ldots x_n) \supset A)$ *where F is any n-place relation variable not occurring free in A and* $x_1, x_2, \ldots, x_n$ *are individual variables.*

(A6) is the comprehension axiom schema and says that every well-formed formula defines a relation of objects. The two rules of inference are modus ponens and universal generalization:

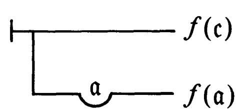
(R1) *B can be inferred from A and* $A \supset B$.
(R2) $\forall vA$ *can be inferred from A if v is any variable.*

Robbin's first three axioms are Frege's formulas (1), (2), and (31), respectively. To show that Frege's system is equivalent to Robbin's involves showing that Frege's substitution rules are equivalent to Robbin's comprehension axiom. But first I shall discuss Robbin's rule of universal generalization and his (A4) and (A5) which are his principles of universal instantiation and confinement.

Robbin's principle of universal instantiation is given by the schema:

(A4) $\forall vA \supset S^v_a|$, *where v is any variable of the same type as a.*

This formulation allows instantiation of both first-order and second-order quantifications. Frege's principle of universal instantiation is not stated in the same way. His formula (58) is a first-order principle of instantiation:

(58) 

which together with substitution gives rise to an infinite number of instances. Although Frege does not enunciate a second-order principle of universal instantiation, it is clear from the text that he intends his system to include one.

Frege treats first-order and second-order quantification in a parallel fashion. At the time of the writing of the *Begriffsschrift* Frege had not yet distinguished the different levels of functions. Later he would say that the first-order quantifier stands for a second-level function and the second-order quantifier stands for a third-level function. In the *Begriffsschrift*, however, after introducing his notation for the first-order quantifier, he draws on his notions of function and argument to introduce second-order quantification. Since '$F(A)$' can be seen as a function of the argument 'F', 'F' can be replaced by a German letter.

At this point in the text Frege first introduces a principle of universal instantiation. He says ([5], p. 24):

> From such a judgment, therefore, we can always derive an arbitrary number of judgments of less general content by substituting each time something else for the German letter and then removing the concavity in the content stroke.

It is clear that Frege intends this principle to apply to both first-order and second-order quantifications. When he says that we can substitute "something else for the German letter", this German letter may be either upper case or lower case; that is, it may be either an individual bound variable or a functional bound variable. Thus from

$$\vdash\!\!-\!\!\overset{\mathfrak{a}}{\smile}\!\!-\!\!-\,\Phi(\mathfrak{a})$$

we can derive each of

$$\Phi(a)$$
$$\Phi(b)$$

and

$$\Phi(c)$$

and from

$$\vdash\!\!-\!\!\overset{\mathfrak{F}}{\smile}\!\!-\!\!-\,\mathfrak{F}(a)$$

we are supposed to be able to derive

$$\vdash\!\!-\, f(a)$$
$$\vdash\!\!-\, g(a)$$

and

$$\vdash\!\!-\, h(a).$$

The same considerations apply to Frege's rules of universal generalization and confinement.

Frege gives his rule of universal generalization as follows ([5], p. 25):

> A Latin letter may always be replaced by a German one that does not yet occur in the judgment; then the concavity must be introduced immediately after the judgment stroke. For example, instead of
>
> $$\vdash\!\!\!-\!\!-\!\!-\; X(a)$$
>
> we can write
>
> $$\vdash\!\smile^{\mathfrak{a}}\!\!-\; X(\mathfrak{a}).$$

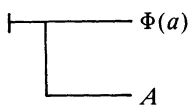
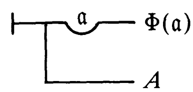
Given Frege's remarks concerning the replacement of a function letter by a German letter, it is clear that he intends his rule of universal generalization to hold for second-order as well as first-order variables. Thus from

$$\vdash\!\!\!-\!\!-\!\!-\; X(a)$$

we can derive

$$\vdash\!\smile^{\mathfrak{F}}\!\!-\; \mathfrak{F}(a)$$

The same can be said of Frege's principle of confinement, which Frege states as follows ([5], p. 26):

> It is clear that from
>
> 
>
> we can derive
>
> 
>
> if A is an expression in which $\mathfrak{a}$ does not occur.

The justification Frege provides for this principle can easily be adapted for the second-order case. Furthermore, it is clear from the text that Frege intends a second-order principle of confinement, for he uses a second-order principle in several derivations (cf. formulas (91), (93), and (95)).

Thus while Frege does intend his principles of universal instantiation, universal generalization, and confinement to be general and not restricted to first-order quantification, some modification of his formulations is required. This is not a serious difficulty. Frege does act as if these principles are general ones, and at this stage he saw no need to distinguish different orders of quantification. This may, in part, be what concerns van Heijenoort. Yet an appropriate modification is consistent with Frege's remarks about the quantifier and generality.

One possible modification is suggested by Bynum in his introduction to the *Begriffsschrift*. Bynum points out that Frege occasionally cites a first-order principle in a derivation when, strictly, he needs a corresponding second-order principle. In fact, it is these slips that worry van Heijenoort. In particular, van Heijenoort is concerned about the derivation of formulas (77) and (91). He writes ([5], p. 3):

> In the derivation of (77) [Frege] substitutes $\mathfrak{F}$ for a in $f(a)$, at least as an intermediate step. If we also observe that in the derivation of formula (91) he substitutes $\mathfrak{F}$ for f, we see that he is on the brink of a paradox.

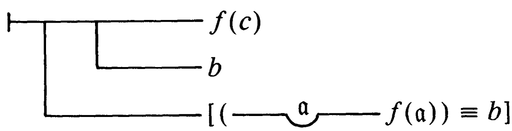
Bynum agrees that there is a problem with Frege's derivations of (77) and (91) but he offers a solution. He does not view the difficulty as symptomatic of any inconsistency in Frege's system. His solution is to provide, in each case, an appropriate corresponding second-order principle.

In the derivation of (77), for instance, Frege cites (68) and indicates the substitutions by a table. The table does indicate that $\mathfrak{F}$ is to be substituted for a in (68), yet Bynum points out that Frege really should have cited an analogous second-order principle that involves quantification over functions. Bynum provides the appropriate second-order principle that Frege should have cited instead of the first-order principle that he does cite. Bynum treats the derivation of (91) in the same way. In a footnote ([1], p. 175) to formula (77) he says:

> The idea of treating $F(y)$ as a function of the function F is in no way contrary to Frege's later thought. To state his thought precisely, however, required notational machinery (which he had not yet devised) to distinguish first- from second-level functions. With that available the difficulty can easily be resolved. van Heijenoort is in error in supposing that any paradox can arise in the system. In the *Begriffsschrift* Frege never confuses first- and second-level functions, though he does not yet have separate terms for them.

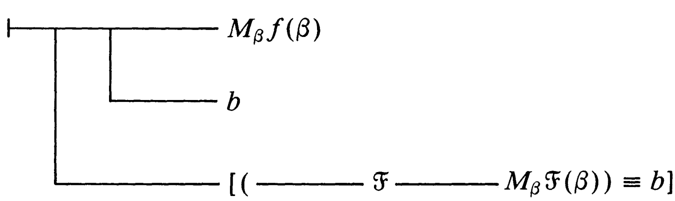
I agree with Bynum here, yet his treatment has certain drawbacks. Frege's formula (68) is:

(68)



The formula Bynum supplies to be used in its stead in the derivation of (77) is:

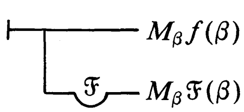
(68′)



Together with the substitution table he provides, (68′) does yield (77). While Frege's (68) is a first-order formula, Bynum's (68′) is a second-order formula. Yet the notation of (68′) is not the notation of the *Begriffsschrift*. In formula (68), 'f' is a first-level function letter, and 'c' stands in its argument-place in '$f(c)$'. In formula (68′), 'M' is a second-level function letter and 'f' is a first-level function letter that stands in its argument place in '$M_\beta f(\beta)$'. Thus '$f(\xi)$' is a first-level function name where 'ξ' indicates the argument place and '$M_\beta \Phi(\ )$' is a second-level function name where '$\Phi(\ )$' indicates the argument place. The first-order quantifier of (68) is a second-level function name whose argument place is filled by a first-level function name. The second-order quantifier of (68′) is a third-level function name whose argument place is filled by a second-level function name.

Bynum's solution invokes a distinction Frege later makes between different levels of functions and accordingly involves expanding the syntax of the *Begriffsschrift* to include higher level function names. Bynum's treatment of (68) suggests introducing second-order formulas which, together with Frege's substitution rules, will provide second-order principles of universal instantiation, universal generalization, and confinement. In particular, a second-order principle of universal instantiation would be:

(58′)



With substitution (58′) gives rise to all the desired instances of universal instantiation. Universal generalization and confinement may be treated in a parallel manner.

Yet there is a problem with this solution, for, as we have seen, it entails expanding the notation of the *Begriffsschrift* to include second-level function letters as well as first-level function letters. If the notation is so expanded, the original problem is merely removed to a higher level. If we allow second-level function letters and the replacement of these letters by German letters, then third-order quantification is introduced. Hence third-order principles of universal instantiation, universal generalization, and confinement must be formulated. Yet if these third-order principles are introduced in a parallel way, they will involve third-level function letters, and the original problem recurs on a different level.

A modification that avoids the defect of Bynum's treatment is to offer schematic formulations of second-order principles of universal instantiations, universal generalization, and confinement. In the case of universal instantiation, a schema is needed that will allow the derivation of

$$\ldots F \ldots$$

from

$$\text{—}\overset{\mathfrak{F}}{\smile}\text{—}\ \ldots \mathfrak{F} \ldots .$$

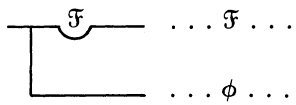
For universal generalization a schematic rule is needed that allows the derivation of

$$\text{—}\overset{\mathfrak{F}}{\smile}\text{—}\ \ldots \mathfrak{F} \ldots$$

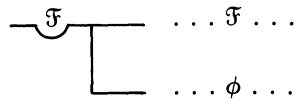
from

$$\ldots F \ldots,$$

and for confinement the needed schema will allow the derivation of



from



where

$$\ldots \Phi \ldots$$

does not contain free occurrences of $\mathfrak{F}$. The principles may be stated after those of Robbin as follows:

(UI) $\forall vA \supset S^v_a A|$ where v and a are function variables
(UG) $\forall vA$ can be inferred from A, if v is a function variable
(Conf) $\forall v(A \supset B) \supset (A \supset \forall vB)$ where v is a function variable not free in A.

In this way, Frege's principles are equivalent to Robbin's.

The remaining discrepancy between Robbin's and Frege's systems can be eliminated by showing that Frege's rules for substitution are equivalent to the comprehension axiom schema. As I have said, Frege does not clearly enunciate his rules for substitution, but by observing his practice the rules can be formulated.

Frege does state a principle for the alphabetic change of variables ([5], p. 51):

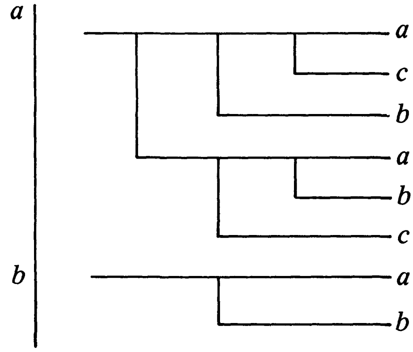
> Replacing a German letter everywhere in its scope by some other one is, of course, permitted, so long as in places where different letters initially stood different ones also stand afterward. This has no effect on content.

Of course Frege must mean that a bound variable may be replaced by another variable of the same type. In this way we avoid the substitution of a function letter for an individual variable and *vice versa*. Frege's remark that this "has no effect on the content" hints at an important restriction that must be placed on substitution. Roughly, we want to ensure that variables that are free or bound in the original formula remain free or bound in the formula that results from substitution. Frege's substitution rules are to preserve validity, implication, and equivalence. There are two kinds of substitution that must be regulated: substitution for propositional letters and substitution for predicate and function letters.

Frege does permit substitution of formulas for propositional letters. He derives a formula or demonstrates its validity and uses substitution instances of the formula to derive other formulas. He treats propositional letters as variables and substitutes more complicated propositional and quantificational formulas for them. He says that expressions that give way to their various substitution instances "contain" those instances as "special cases" ([5], p. 16).

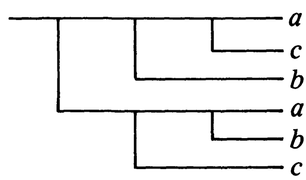
In most derivations Frege cites the number of a judgment to be used in the derivation and indicates the substitutions to be made by a table under the citation. It is left to the reader to construct the appropriate substitution instance of the cited formula. The table serves as an abbreviation. For example,

(1):

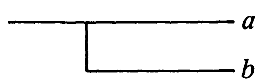


is an abbreviation for:

in judgment (1) put



for '*a*', and



for '*b*'.

The rule that governs the derivation of quantificational formulas from truth functional formulas, and more complex formulas from simpler formulas can be stated as follows[2]:

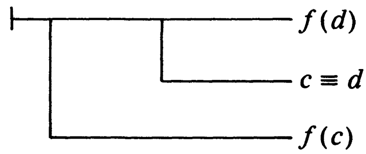
(S1) *A formula A may be substituted for a propositional letter in a formula B provided that no free variable of A is captured by a quantifier of B.*

Frege also allows substitution for functional letters or predicate letters. For example, consider the table:

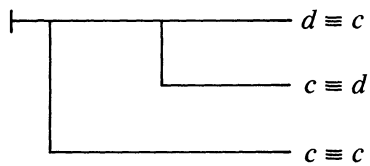
$$(53): \quad F(A) \,|\, (A \equiv C).$$

Judgment (53) is the following:

(53)



Hence the substitution instance constructed from the table is:



Frege's use of the capital Greek letters 'A', 'B', and 'Γ' in such tables is similar to Quine's use of circled numerals '①', '②', '③', . . . as place holders in schematic predicates. A schematic predicate, for Quine, is like an open sentence except that it contains place-holders instead of some or all of the free variables. For example,

'F ①', 'G ① ②', and '$(\exists x)$ $(F$ ① x & Gx ①$)$'

are schematic predicates. Monadic schematic predicates may be substituted for monadic predicates and, in general, n-place schematic predicates may be substituted for n-place predicates. In the table

$$f(A)|(A \equiv C)$$

'A' acts like Quine's '①' ([5], p. 24). The table indicates that '(① $\equiv c$)' is to be substituted for the monadic 'f'. Appropriate restrictions must be placed on the substitution of schematic predicates for predicates or function letters. It must be ensured that no variable of the substituted schematic predicate becomes captured by a quantifier of the formula in which it is substituted, and no variable of the formula is captured by a quantifier of the predicate. The rule may be stated as follows:

(S2) *A schematic predicate may be substituted for a function letter or predicate in a formula A provided that no free variable of the predicate is captured by a quantifier of A and no free variable of A is captured by a quantifier of the predicate.*

It remains to show that the rules of substitution and the comprehension axiom schema of Robbin's system are equivalent. In the proofs that follow I take for granted the rules of existential generalization, existential instantiation, and other basic logical rules that are common to first-order and second-order quantification. The justifications of these rules do not involve either substitution or the comprehension axiom.

First, to show that substitution implies the comprehension axiom schema, note that

(1) $(\forall x)(G(x) \leftrightarrow G(x))$

is a theorem. By existential generalization we can derive

(2) $(\exists F)(\forall x)(F(x) \leftrightarrow G(x))$.

By substituting 'Φ' for 'G' in (2) we get

(3) $(\exists F)(\forall x)(F(x) \leftrightarrow \Phi(x))$

which is the comprehension axiom.

The other direction is the derivation of substitution from the comprehension axiom. Given some theorem

(1) $\ldots G(\ ) \ldots$

where 'G' is a functional variable, we want to show

(2) $\ldots \Phi(\ ) \ldots$

to be a theorem as well, where 'Φ' is a formula substituted for 'G' in (1). Frege's substitution table would look like:

$$(1): \qquad G(\Gamma) | \Phi(\Gamma).$$

In a broad outline, the proof goes as follows. First we show that

(3) $(\forall x)(G(x) \leftrightarrow \Phi(x)) \supset (\ldots G(\ ) \ldots \leftrightarrow \ldots \Phi(\ ) \ldots)$

holds for all formula contexts

$$\ldots \underline{\qquad} \ldots .$$

The proof is by mathematical induction on the complexity of the formula context in which 'G' occurs.[3] Since (1) is a theorem, we derive

(4) $(\forall x)(G(x) \leftrightarrow \Phi(x)) \supset \ldots \Phi(\ ). \ldots$

An application of universal quantification to (4) yields

(5) $(\forall G)((\forall x)(G(x) \leftrightarrow \Phi(x)) \supset \ldots \Phi(\ ) \ldots).$

Since 'G' does not occur free in Φ, (5) is equivalent to

(6) $(\exists G)(\forall x)(G(x) \leftrightarrow \Phi(x)) \supset \ldots \Phi(\ ). \ldots$

Then by the comprehension axiom, from (6) we derive

(7) $\ldots \Phi(\,) \ldots$

as desired. Hence substitution follows from the comprehension axiom.

I have shown that Frege's substitution rules are equivalent to Robbin's comprehension axiom schema. The reconstruction of the *Begriffsschrift* is now complete. Frege's system is equivalent to Robbin's second-order predicate calculus. Thus in order to show that Frege's system is consistent, all that remains to show is that Robbin's second-order logic is consistent.

The function e is defined inductively as follows:

(i) If A is an atomic formula, $e(A)$ is the result of erasing all the individual variables and constants of A.
(ii) If A is a quantification $\forall vB$, $e(A)$ is $e(B)$
(iii) If A is a negation $-B$, $e(A)$ is $-e(B)$
(iv) If A is a conditional $B \supset C$, $e(A)$ is $e(B) \supset e(C)$
(v) If A is a quantification $(\forall V)B$, $e(A)$ is the conjunction of the result of substituting 'T' for 'V' in $e(B)$ and the result of substituting '$\perp$' for 'V' in $e(B)$.
(vi) If A is a quantification $(\exists V)B$, $e(A)$ is the disjunction of the result of substituting 'T' for 'V' in $e(B)$ and the result of substituting '$\perp$' for 'V' in $e(B)$.

We want to show that for each axiom A, $e(A)$ is a tautology and furthermore that the rules of inference preserve tautologousness. It will follow, then, that if A is a theorem, $e(A)$ is tautology.

For the first three axioms, (A1)–(A3), of Robbin's system, it is clear that an application of e yields a tautology.

(A4) is universal instantiation:
$e[(\forall v)(A(v) \supset A(a))]$ *is a tautology of the form*: $p \supset p$.

(A5) is confinement:
$e[(\forall v)(A \supset B) \supset (A \supset (\forall v)B)]$ *is a tautology of the form*: $(p \supset q) \supset (p \supset q)$.

(A6) is the comprehension axiom schema:
$e[(\exists F)(\forall x)(F(x) \leftrightarrow \Phi(x))]$ is $T \leftrightarrow e(\Phi) \vee \bot \leftrightarrow e(\Phi)$ *is a tautology*.

(R1) is modus ponens:
If $e(A)$ *and* $e(A \supset B)$ *are tautologies, then so is* $e(B)$

(R2) is universal generalization:
If $e(A)$ *is a tautology, then so is* $e[(\forall v)A]$, *which is just* $e(A)$.

Hence for all formulas A, if A is a theorem, $e(A)$ is a tautology. If B and $-B$ were both theorems, then $e(B)$ and $-e(B)$ would both be tautologies. But this is impossible. Therefore the theory is consistent. The application of the function e is tantamount to interpreting the system in a universe of discourse with a single element. All of the theorems of the system are true under such an interpretation, but no formula and its negation can be true.

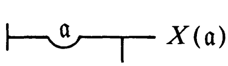
Frege's *Begriffsschrift* is consistent. It does not lead to a paradox. What, then, is behind van Heijenoort's worry? In an attempt to answer this question, I turn now to the interpretation of Frege's theory.

Frege says very little concerning the interpretation of the system presented in the *Begriffsschrift*. In particular, he says little about how we are to interpret the second-order formulas of the third part of the work. It is these formulas that concern van Heijenoort. While the inclusion of second-order quantifications does not render the theory inconsistent, there are problems associated with the interpretation of these formulas that can be seen as anticipating problems that arise in Frege's later writings. In particular, Frege seems dangerously close to Russell's paradox and the problem with the concept *horse*.

Frege was not concerned with truth in models, nor did he ask whether his system is complete. It is not clear how Frege chose his axioms, but I shall describe the semantics I believe Frege did have in mind for the system of his *Begriffsschrift*.

The universe of discourse for Frege's individual variables is not restricted, it is the entire universe. Although Frege never explicitly states this, it is implicit in the text of the *Begriffsschrift*. Frege translates the quantification:

as "the judgment that whatever we may take for its argument, the function is a fact" ([5], p. 24), and



as “whatever $\mathfrak{a}$ may be, $X(\mathfrak{a})$ must always be denied” ([5], p. 27). In another place Frege writes, “$\overset{\mathfrak{a}}{\smile} f(\mathfrak{a})$ means that $f(\mathfrak{a})$ takes place whatever we may understand by $\mathfrak{a}$” ([5], *p.* 51). It is clear that Frege does not restrict the range of the individual variables; they range over “*whatever* we may take as value”, that is, everything that there is. As a result, any formula

$$\Phi(x)$$

expresses a generality about everything in the universe. The notation of the universal quantifier is introduced both as a way to confine the scope of the generality to the antecedent of a formula and, in conjunction with negation, to express particular existential claims. We can assume Frege to have a certain interpretation in mind. I see no reason to think that functional and predicate letters do not range over Fregean functions and concepts. Since the range of the individual variables is the entire universe, the range of the functional variables is included in the range of the individual variables. Frege does not distinguish different universes of discourse for different types of variables. As we have seen, Frege makes no formal distinction between first-order and second-order quantification. Second-order quantification is introduced for the same considerations as is first-order quantification. Both follow naturally from his discussion of the function and argument.

It should be noted that although Frege’s notions of function and argument are presented syntactically in the *Begriffsschrift*, in view of his later characterization of functions and concepts, function variables cannot range over sets of objects or individuals. Sets are “complete”, or “saturated”, and hence lack the requisite incompleteness of the function.

How, then, are we to interpret the second-order formulas of the *Begriffsschrift*? While Frege says little about quantificational formulas under interpretation, the text does provide a few clues. In his section on functions ([5], pp. 23–24), Frege writes that the formula

$$\vdash\!\!\!-\!\!-\!\!-\!\!-\ \Phi(A)$$

can be read as ‘A has the property Φ’, and the formula

$$\vdash\!\!\!-\!\!-\!\!-\!\!-\ \Psi(A,B)$$

as ‘B stands in the relation Ψ to A’ ([5], p. 24). Thus one-place functional letters stand for properties and two-place functional letters for two-place relations. Frege translates

$$\vdash\!\!\top\!\!\overset{\mathfrak{a}}{\smile}\!\!\top\!\!-\ \Lambda(\mathfrak{a})$$

as ‘there are Λ’, where we call something that has the property Λ a Λ. Hence when functional letters occur in quantifiers we can presume that they range over properties and relations. For Frege, an object has a property Φ just in case that object falls under the concept Φ. In “On Concept and Object” ([2], p. 51), Frege writes:

> I call the concepts under which an object falls its properties; thus ‘to be Φ is a property of Γ’ is just another way of saying ‘Γ falls under the concept of a Φ’.

Frege does talk of *the* relation Ψ and *the* property Φ, yet we may assume that properties are Fregean concepts and that a suitable interpretation of the system must take care to avoid running into the problem with the concept *horse*. In light of Frege's doctrines concerning concepts and objects, it is clear that the standard interpretation that allows functional variables to range over sets of objects is inadequate for Frege's theory.

In Part III of the *Begriffsschrift* Frege begins to use second-order quantification to develop his definition of the ancestral of a relation. Frege translates certain of the formulas of Part III into words. On examination of these translations it becomes evident that the functional letters are intended to range over properties and relations. In particular, formula (76), which is Frege's definition of the proper ancestral of a relation, contains a universal second-order quantifier. Frege says ([5], p. 60) that (76) "can be rendered into words somewhat as follows":

> If from the two propositions that every result of an application of the procedure f to x has property F and that F is hereditary in the f-sequence, it can be inferred, whatever F may be, that y has the property F, then I say: 'y follows x in the f-sequence'.

Here Frege translates the universal quantifier as "whatever F may be" and talks of "property F" and "the procedure f". Procedures are just relations and properties are concepts. Frege does not, however, give us any clear characterization of properties and procedures. He does, as I have pointed out, lapse into using the definite article, he talks of "the property . . ." and "the procedure . . .". Are these properties and relations saturated objects? Frege does say that the translation is a rough one, and I shall assume that his use of the definite article is due to an "awkwardness of language".

No matter how these properties and relations are characterized, if they exist then they must be contained in the range of the first-order variables; for the range of the first-order variables is the entire universe. However, if the range of the second-order variables is included in the range of the first-order variables, we find Frege very close to Russell's paradox. Consider:

(*) $(\exists F)(\forall x)(F(x) \leftrightarrow -A(x,x))$.

The formula (*) is an instance of the comprehension axiom schema and hence is a theorem for Frege's theory. Frege does not tell us how to interpret formulas for the form

$$(\exists F)(\ldots\ldots F(\ )\ldots\ldots)$$

yet from what he does say it seems that such formulas are to be read as

There is a property such that. . . .

Let us say that an object has a property Φ just in case Φ applies to that object. If we suppose that properties exist, it follows that they are in the range of the individual variables. If we let '$A(x,x)$' stand for 'x is a property that does not apply to itself' then (*) becomes:

> There is a property that applies to an object if and only if that object is a property that does not apply to itself.

On the supposition that properties are themselves things in the universe, this interpretation of (*) gives us Russell's paradox. The odious interpretation can be avoided if it is stipulated that the range of '*F*' is not included in the range of the individual variables. Yet then it would follow that some values of functional variables are not in the universe. This way out of the paradox is not suggested by anything Frege writes. He certainly talks as if properties and relations exist, and hence are in the universe. When introducing second-order quantifications he states no qualification about the range of their variables. Just as first-order quantifications express generalities about things in the universe, so second-order quantifications express generalities about properties and relations. Any restrictions on the range of the second-order variables must be seen as *ad hoc.*

Frege says nothing to exclude properties and relations from the range of the individual variables, and he never considers alternative interpretations or models for his system. All this lends support to van Heijenoort's claim that Frege is "on the brink of a paradox". Although the system is a consistent one, the interpretation of the system that is suggested by the text does lead to Russell's paradox.

Another problem that arises in connection with the interpretation of the system presented in the *Begriffsschrift* is one associated with the problem with the concept *horse.* If the values of the second-order variables are not saturated objects and have the required incompleteness of the Fregean function, then how do we accordingly interpret second-order existence claims? Note that the quantifier

There is such a thing as __________

has the peculiar feature that all its grammatical completions assert the existence of objects and not concepts or functions. The problem is: If properties and relations do exist, how can we name these values of second-order variables without naming objects?

Frege needs the means to express grammatically second-order existence claims. If properties and relations cannot be named by proper names then the natural language quantifier above is an inadequate translation of the second-order existential quantifier. Yet there seems to be no other way to interpret formulas such as

There is another related problem. Consider the formula:

$$(\forall x)(F(x) \vee -F(x)).$$

This would be interpreted as meaning:

Everything is such that either it has property *F* or it does not have property *F*.

If properties are in the universe then one instance of this formula is:

Property *F* either has property *F* or it does not.

However, given Frege's doctrines concerning concepts and objects, this is ungrammatical. We cannot grammatically predicate having property *F* or failing to have property *F* of property *F*. Thus in light of the Fregean hierarchy of objects, concepts, and functions, neither disjunct is grammatical. This would be avoided if the universe of discourse for the different types of variables were distinct. Yet they cannot be distinct if the individual variables range over everything that there is. Even if this problem were resolved, Frege would be left with the problem of grammatically asserting the existence of properties and relations.

Thus while van Heijenoort is incorrect in claiming that the system presented in the *Begriffsschrift* leads to an inconsistency or paradox, the interpretation suggested by Frege's text does anticipate some serious problems that arise in Frege's later writings.

NOTES

1. Other reconstructions of the system presented in the *Begriffsschrift* are possible, but the one offered in this paper is both a simple and natural explication. The major results of Frege's system go through by the same argumentation used by Frege.
2. See [3] for a discussion of what I label '(S1)' and '(S2)'.
3. The induction steps correspond to the logical operators. The steps are routine and therefore omitted.

REFERENCES

[1] Frege, Gottlob, *Conceptual Notation and Related Articles*, trans. and ed., Terrell Ward Bynum, Clarendon Press, Oxford, 1972.

[2] Geach, Peter and Max Black, eds., *Translations from the Philosophical Writings of Gottlob Frege*, Basil Blackwell, Oxford, 1970.

[3] Quine, W. V. O., *Methods of Logic*, 3rd Ed., Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1970.

[4] Robbin, Joel W., *Mathematical Logic*, W. A. Benjamin, Inc., New York, 1969.

[5] van Heijenoort, Jean, ed., *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Harvard University Press, Cambridge, Massachusetts, 1967.

Massachusetts Institute of Technology
Cambridge MA 02139