Fregean Quantification Theory

Saul A. Kripke

This is a post-peer-review, pre-copyedit version of an article published in the *Journal of Philosophical Logic*. The final authenticated version is available at:

https://doi.org/10.1007/s10992-013-9299-x

Fregean Quantification Theory

The present note is a side outgrowth of my study of Frege, much of it contained in Kripke (2008). It occurred to me that, given Frege's idea that sentences are names of truth-values (which are objects of the domain of the quantifiers like any other), negation need not be taken as a primitive; it can be defined in terms of the (material) conditional and universal quantification.[1] This is true as long as the conditional has its usual table of values when applied to truth-values. But then it also occurred to me that a change in Frege's definition of the values of the conditional (according to him arbitrary, but necessary to the system), when the arguments are not both truth values, enables one to define identity as well. Thus, Frege's unusual semantics and syntax leads to a formulation of first-order logic with identity whose economy of primitives has not, as far as I know, been noted before. In what follows, the semantics and syntax of the language will be presented in contemporary terms.

As we said, for Frege, formulae are simply terms of a special sort, denoting either the true or the false. The syntax of Fregean first-order logic is given by a list of terms. There is an infinite list of (individual) variables x_1, x_2, … For each $n \geq 0$, $i \geq 1$, there is an infinite list of n-ary function letters $\{f^n_i\}$. (If $n = 0$, the "function letter" is called a "constant"). In addition to these, there is a special binary function symbol, *cond*, and for each variable x_i a second level function symbol (x_i).

[1] In contemporary logic, of course, it is standard that the negation of A is definable as A $\supset \perp$ (i. e., in the present notation, *cond* A F). The referee has pointed out that the definition of F as $(x_1)x_1$ is already given by Church (1951), p. 9, as I had forgotten. (One [inessential] difference is that for Church, the two truth-values form a type of their own, as the range of the universal quantifier, and are not included in D.)

The class of terms is defined inductively as the least class closed under the following conditions.[2]

(a) Constants and variables are terms.

(b) If f^n_i ($n \geq 1$) is a function letter and $t_1, \ldots, t_n$ are terms, $f^n_i t_1 \ldots t_n$ is a term.

(c) If t_1 and t_2 are terms, so is $condt_1t_2$.

(d) If x_i is a variable and t is a term, so is $(x_i)t$.

A Fregean domain is a triple <D, **t**, **f**>, where D is a set, **t** ε D, **f** ε D, and **t** ≠ **f**. (Intuitively, **t** and **f** are "the True" and "the False", respectively.)

An *interpretation* or *valuation* v of a Fregean term in D is an assignment of an element of D to each variable or constant, of an n-place function from $D^n \rightarrow D$ to each n-place function letter, and obeys the following special rules for *cond* and (x_i), respectively:

(1) If $v(t_1)$ = **t** and $v(t_2)$ = **f**, then $v(condt_1t_2)$ = **f**. Otherwise, if $v(t_1)$ and $v_2(t_2)$ are both truth values, $v(condt_1t_2)$ = **t**. If either $v(t_1)$ or $v(t_2)$ is *not* a truth-value, then $v(condt_1t_2)$ = **t** if $v(t_1) = v(t_2)$ and = **f** otherwise.

[2] The referee has emphasized that the notation is Polish and that parentheses are not needed as a bracketing device. (The use of parentheses in the universal quantifier is different and of course eliminable by using another notation.) The Polish character of the notation is an adventitious byproduct of the fact that the conditional is merely a binary function on the domain, written like other such function.

I shall therefore omit parentheses in the initial official clauses defining the syntax and semantics. Later, I shall freely use them when I think they improve readability.

Of course, Frege's own original notation also did not require bracketing but has never been used again because of its peculiar two-dimensional character. (I am not personally fond of Polish notation, in spite of its advantages for certain purposes.)

(Intuitively, *cond* represents the material conditional. As in Frege, its values are arbitrary when the arguments are not truth-values, but they must be defined. We make a different choice from Frege for this case, since it reduces the number of primitives.)

(2) If $v_i(t)$ = **t** for every valuation v_i different from v at most in its assignment to x_i, then v assigns **t** to $(x_i)t$; otherwise it assigns **f**.

A term is *valid in a domain* D if its value is always **t** for any valuation in D; it is *valid*, *simpliciter*, iff it is valid in every domain. Similarly, it is *satisfiable in a domain* D if for some valuation it takes the value **t**; *satisfiable*, *simpliciter*, if it is satisfiable in some domain D.

Note that the extension of validity would not change if we allowed **t** = **f**. On the other hand, satisfiability would change radically (see below).

To get the usual notions, F is defined as $(x_1)x_1$. For any term t, *neg* (t) is *cond* (t, F). For any terms t_1, t_2, *conj*(t_1, t_2) is *neg*(*cond*(t_1, neg(t_2)), and *id*(t_1, t_2) is *conj*(*cond*(t_1, t_2) *cond*(t_2, t_1). This gives us all the usual notions of first-order logic with identity. (As Parsons 1987 has also noted, from Frege's point of view the material biconditional is a special case of identity!)

Also, we could define T as *neg*(F) and the Fregean horizontal —t as *id*(t, T).

As I said above, the extension of validity would not change if we allowed **t** = **f**. So if the system were formalized to prove the valid formulae it would remain the same. But satisfiability would change: *conj*(t,(*neg*(t))) would be satisfiable. One could add to the system a description operator: $\iota x_i t$. If there a unique element a ε D such that $v(t)$ = **t**, when x_i is assigned a as value, the values of the other variables remaining fixed, then $v(\iota x_i t)=a$. Otherwise, $v(\iota x_i t)$=**f**. Terry Parsons has shown that one could even include an operator corresponding to Frege's course-of-values operator (Parsons 1987). But to do so would be

to abandon uniqueness of the semantical conditions for the definitions, as well as the spirit of a formalization of first-order logic. Moreover, here we do not require an infinite domain, which is needed for Parson's construction.[3]

I have not tried to find an elegant axiomatization of the system. But it is obvious that everything done here can be formalized in conventional first-order logic, and hence, by Gödel's completeness theorem, the valid formulae are recursively enumerable.

References

Church, A. (1951). A Formulation of the Logic of Sense and Denotation. *Structure, Methods, and Meaning: Essays in Honor of Henry M. Sheffer*, P. Henle, H. M. Kallen, and S. K. Langer (eds.), New York: The Liberal Arts Press.

Hilbert, D. and Ackermann, W. (1938). *Grundzüge der theoretischen Logik*, second edition. Berlin: Springer. (First edition, 1928.) Translated as *Principles of Mathematical Logic*, R. E. Luce (ed.) and L. Hammond, G. Leckie, and F. Standhardt (trans.). New York: Chelsea Publishing Company, 1958.

Kripke, S. (2008). Frege's Theory of Sense and Reference: Some Exegetical Notes. *Theoria*, 74: 181-218. Reprinted in *Philosophical Troubles. Collected Papers Volume I*. New York: Oxford University Press.

[3] We have differed from Frege in other ways. Following much (but not all) modern usage, we don't distinguished in style of letters between free and bound variables, nor, more importantly, on the occurrence of free and bound variables. (I personally think that there is something to be said for some of these restrictions. Restrictions of this sort are in Hilbert and Ackermann 1928 as well as in Frege.)

Another difference from Frege is the use of variables over an arbitrary domain D, albeit with two distinguished elements. Frege wanted his first order variables to range over all objects. We haven't done this here, following contemporary conceptions (why courses-of-values are not included in the domains is explained in the text. Also, we don't define the definite description operator in terms of the course-of-values function for the same reasons.)

Parsons, T. (1987). On the Consistency of the First-Order Portion of Frege's Logical System. *Notre Dame Journal of Formal Logic*, 28: 161-68.