Frege's model theory

Brice Halimi

First IUSS/IHPST Joint Workshop in the Philosophy of Mathematics and Logic

Main claim

Frege's universalism is compatible with contextualism.

More specifically:

- The defense of the universality of logic is compatible with the admission of variable *contexts* of discourse.
- A properly Fregean notion of model can be devised.
- An understanding of non-logical constants is not beyond Frege's logical horizon.

Working hypothesis: Some insights and devices of Frege's logic still ought to be reactivated, because otherwise Frege appears as the poor relation of Tarskian metatheory.

Outline

I will:

- compare the respective ways in which Kant and Frege combine two different features commonly ascribed to logic;
- criticize the label "logical universalism;"
- show how Frege's perspective countenances the existence of several contexts of discourse;
- draw on a suggestion made by Frege in his discussion of Hilbert, and construct a Fregean notion of model;
- set out a formal framework to illutrate that point and compare Frege with Tarski;
- take up a suggestion from Wilfrid Hodges and William Demopoulos that non-logical constants of a formal language can be compared to indexicals, and show, *contra* Hodges and Demopoulos, that Frege can, as well as Tarski, make sense of non-logical constants.

Two features commonly ascribed to logic

Logic is a theory that precedes any other theory. The juridiction of logic knows no bound.
This traditional characterization of logic is cashed out in terms of two distinct features:

- the universality of logic
- the radicality of logic

The universality of logic consists in its being about absolutely everything: a single unrestricted range of values for entity variables. Nothing can escape logic.

The radicality of logic corresponds to there being the one and the same logic that any reasoning must comply with. The principles of logic are laws that one cannot but presuppose them, even to challenge them. Nobody can escape logic.

Kant's *Logic*

The universality and the radicality of logic have been associated since at least Kant:

> *If [. . .] we put aside all cognition that we have to borrow from objects and merely reflect on the use just of the understanding, we discover those of its rules which are necessary without qualification, for every purpose and without regard to any particular objects of thought, because without them we would not think at all.*
> *("The Jäsche logic", Ak.* ix, *12)*

Logic is "a science a priori of the necessary laws of thought [radicality], not in regard to particular objects, however, but to all objects in general [universality]."

John Macfarlane, "Kant, Frege, and the Logic in Logicism" (2002)

Even though Kant and Frege obviously do not conceive of the generality of logic in the same way, they concur, according to Macfarlane in characterizing logic through its "normative generality."

Logic provides one with the norms with which any thought whatsoever must comply as such, "if one is to think at all," as opposed to the way in which one ought to think in some particular domain.

However, MacFarlane's analysis somewhat conceals the asymmetry of Kant's and Frege's respective agendas.

Asymmetry of Kant's and Frege's respective agendas

Kant – Formal logic is in a sense more general than transcendantal logic, because it does not presuppose the possibility of experience. Yet the domain ruled by formal logic cannot be seen as extending the domain ruled by trancendental logic (possible experience). The generality of formal logic has to be reckoned through something else than its universality, namely: It states conditions of possibility of thought as such.
To Kant, radicality is a way to account for the universality of formal logic from within a nonformalist conception of logic.

Frege – Frege's move goes the other way around.
To Frege, universality is a way to account for the radicality of logic from within an anti-psychologistic conception of logic, i.e., without committing oneself to any (slippery) notion of necessity.

To sum up

Both Kant and Frege bestow on logic two distinctive features: universality and radicality.

And both agree to consider those features as inseparable.

But Kant puts the stress on radicality, whereas in Frege logical universality prevails over logical radicality.

Hidden agenda

Logical universality (“absolute generality”) is incompatible with ZFC set theory.

Logical radicality (“the language of logic as a universal medium”) is incompatible with any semantic metatheory.

Model theory is the combination of the idea of a general semantic metatheory with the framework of ZFC.

In Frege’s universalism, logical universality only prevails.
As a result, Frege’s conception of logic is certainly hardly compatible with the usual ZFC-based model theory, but not with a general semantic metatheory implemented differently (not based on set-theoretic “universes of discourse”).

Frege's "logical universalism"

Logical universalism, a label that has been pinned on to Frege, designates:

- the sheer conflation of logical universality and logical radicality
- the further claim that "one cannot consider one's language from the outside".

Although a label that has been pinned on to Frege, it is not epitomized by Frege. Not any more by Russell, Wittgenstein, Carnap or even the young Tarski.

Logical universalism is an invention of its enemies.

The "no-metatheory" interpretation

A tradition upheld by van Heijenoort and Hintikka put in the same "universalist" bag four independent claims:

- logic resorts to absolutely unrestricted (unspecified) variables
- logic refers to one single universe of discourse
- the language of logic is a universal medium in which one can say everything.
- meta-theoretical studies, and semantics to begin with, are impracticable and, for that reason, banned.

This obviously does not square with the distinction that Frege emphasizes in the *Grundgesetze* between the "expository language" (*Darlegungssprache*) and the "auxiliary language" [*Hilfessprache*] of his logical system. The latter is certainly not universal in the sense of the ordinary language, and the former eludes in some respects the rules set out in the formal system, as the concept 'horse' paradox shows.

Frege, "On the Foundations of Geometry" (1903)

The universality of logic does not preclude one from considering local universes of discourse, in particular for metatheoretical purposes.

Frege (1903):

> *[...] Euclidean geometry presents itself as a special case of a more inclusive system which allows for innumerable other special cases — innumerable geometries, if that word is still admissible. [...] If one wanted to use the word "point" in each of these geometries, it would become equivocal. To avoid this, we should have to add the name of the geometry, e.g. "point of the A-geometry," "point of the B-geometry," etc. Something similar will hold for the words "straight line" and "plane."*

Obviously, all special geometries are not subdomains of a single geometrical domain, but different specifications of the meaning of fundamental geometrical concepts, and thus correspond to different theoretical contexts.

Logical contextuality

So Frege was not insensitive to the idea of context variation.

Many scholars (Landini, Tappenden, Heck, Antonelli, Sullivan, Blanchete, Korhonen) already showed that there was no objection of principle in Frege against "meta-theoretic" considerations. Most of them refer to Frege's 1906 "On the Foundations of Geometry."

However, the focus of the reconsideration of Frege has remained mostly truth: Frege's metatheoretic samples pertain to semantics understood as the theory of reference and truth.
The now very substantial material adduced by the no-"no-metatheory" interpretation calls for a stronger claim:
Not only Frege's conception of logic countenances metatheoretic considerations, but Frege's suggestions in his 1906 essay hint at a genuine model-theoretic framework, although of course not of the Tarskian sort.

Frege's 1906 "Foundations of Geometry"

Frege introduces variable reinterpretations as internal translations called vocabularies:

> *Imagine a vocabulary: not, however, one in which words of one language are opposed to corresponding ones of another, but where on both sides there stand words of the same language but having different senses. [...] We may say in general that words with the same grammatical function are to stand opposite one another. Each word occurring on the left has its determinate sense —at least we assume this— and likewise for each one occurring on the right. Now by means of this opposition the senses of the words on the left are also correlated with the senses of the words on the right. Let this correlation be one-to-one, so that on neither the left nor the right is the same thing expressed twice. We can now translate; not, however, from one language to another, whereby the same sense is retained; but into the very same language, whereby the sense is changed.*

Tappenden, "Frege on axioms, indirect proofs, and independence arguments in geometry" (2000)

> *Crucially, for Frege, thoughts are evaluated as having the truth-values they actually have. No true thought is treated as false or projected into counterfactual circumstances in which it is false. (Axioms will be paired with other thoughts some of which may be actually false.)*

Important difference with Tarski: A Fregean vocabulary $e \to e'$ relates expressions e and e' belonging to some interpreted language (interpreted from the outset).

Still, Frege's perspective is neither in principle, *nor in fact* incompatible with the recognition of various contexts of discourse.

Two objections

Admittedly, such admission of various contexts, in Frege's writings, pertains to geometry, rather than to arithmetic and logic.

Still, nothing precludes one from extending Frege's suggestion to a general framework, beyond the strict discussion of geometry. (This has nothing to do with the fact that geometry is not as general as arithmetic and logic.)

Admittedly, Frege rejects any vocabulary which would change the sense of a word standing for a logical constant. Logical terms should be left invariant by any vocabulary.

But that is precisely what happens in modern semantics as well.

Two obstacles

Actually, Frege mentions two obstacles in following the strategy based on vocabularies:

> *We will find that this final basic law [the formality of logical laws according to which each object or first-level concept can be replaced, as far as logic is concerned, with any other] which I have attempted to elucidate by means of the above-mentioned vocabulary still needs more precise formulation, and that to give this will not be easy. Furthermore, it will have to be determined what counts as a logical inference and what is proper to logic.*

The first obstacle is the fact that no syntactic result is able to establish the kind of *absolute* logical dependencies that Frege has in mind.

The second obstacle is the problem of delineating the logical constants.

First obstacle

Each syntactic representation of logical relations between thoughts by a system of derivations between sentences corresponds only to a certain decomposition of logically complex notions, linked to a particular choice of primitives.

Answer: Establishing facts about derivability w.r.t. a certain syntactic representation of logic is not a second-best option, even in Frege's view,

Second obstacle

Definition 1. Given two terms a, b with the same syntactic type, v_a^b is the vocabulary (i.e., the reinterpretation of the language) induced by the replacement of 'a' with 'b'. For any sentence p of the language, $v_a^b(p)$ is the (possibly) new sentence generated from p by this vocabulary.

A thought expressed by a sentence p is valid (resp. contravalid) iff there exists some constituent u of p such that, for any term x of the same syntactic type as u, the thought expressed by $v_u^x(p)$ is true (resp. false).

A thought is (logically) determinate if it is either valid or contravalid.

Definition 2 (simplified). The class of logical terms is the largest class of terms, the composition of which can elicit but logically determined thoughts.

(Adaptation of the definition given in Carnap's *Logical Syntax of Language*.)

Outcome: Both obstacles that stand in the way of implementing Frege's suggestion can be overcome.

Analysis of Frege's suggestion

Each "vocabulary" performs a context change, the replacement of certain semantic values with others: The meaning of a word is assigned as the new meaning of some other word.

In that framework, each context variation induces a systematic reinterpretation of language.

Indeed, an interpreted language is not a mere collection of words, but a system of meanings and references, and thus any change about one of them is always a simultaneous change of several interconnected meanings and references.

This is what complex names embody in the *Grundgesetze*, were they object-names or function-names. Each complex name, considered as an autonomous syntactic unit, has a structure.

Each context change has to reflect the dependencies that follow from the way complex names are formed (*Grundgesetze* §§29-31).

Formalization

Any technical apparatus fit to represent Fregean context change should possess the following features:

- Each context appears as one item within some space of all possible (partial) reinterpretations of the language.
- To each context corresponds a certain system of interconnected semantic values and each context change respects these connections.
- The relationships between any two contexts are mirrored by the relationships existing between the respective references of complex names as they are interpreted in both contexts.

Fibrations

A general technical apparatus suggests itself: that of fibrations.

Suppose that, to each object b of some category B, is assigned a category E_b, in such a way that any arrow $u : b \to b'$ in B gives rise to a functor $u^* : E_{b'} \to E_b$, called a "reindexing functor."
B is the "base category", E_b is the "fiber" above b and the disjoint union E of all fibers is the"total space."

The corresponding fibration over B is the functor $\pi : E \to B$ sending each object in E_b to b, and each arrow in E_b to 1_b (identity on b).

Fibrations generalize the set-theoretic notions of surjection and pre-image: $E_b = \pi^{-1}(b)$.
The fundamental extra-ingredient is the fact that the base category B is endowed with some structure (as embodied by its arrows), and that the existence of a fibration requires a systematic connection between the relations between any two objects in the base space, and the relations between the corresponding fibers in the total space. Base category = control space.

The picture extracted from Frege's 1906 suggestion naturally fits the definition of a fibration.

- Each Fregean vocabulary is defined as a variant of the regular use of (certain) names, and as such belongs to some base variation space. This is the horizontal dimension of the picture.
- On top of each vocabulary stands a fiber: the complete referential system induced by the reinterpretation of certain terms. This is the vertical dimension of the picture.
- Both dimensions are correlated, since each horizontal variation is matched by a correlative variation of what stands above each vocabulary. Hence a fundamentally two-dimensional picture.

Each vocabulary change is a partial shift of all semantic values inducing a transformation (reindexing functor) between indexed structured systems (fibers).
This remains to be formalized.

Let L be a fixed language, a sublanguage of "the" language L_0, and let σ be the signature of L.
For instance, if L is the language of group theory, then $\sigma = \{e, \cdot, (-)^{-1}\}$.

A Fregean σ-structure S consists of a set Ω of expressions, together with a "vocabulary" $f : \Omega \to \Omega_0$ such that $\sigma \subseteq \mathrm{Im}(f)$, where Ω_0 is the set of all expressions of the language L_0.
Each such σ-structure $S = \langle \Omega, f \rangle$ can be understood as a context (in the broad sense) for the use of all the elements of σ.

$f(e) =$ expression (of L_0) that e replaces in L: In L, e means what $f(e)$ means usually (in L_0).

In case $f(c) = c$, S simply adopts the intended interpretation of 'c'.

Example: shift from Euclidean to projective plane geometry

The signature σ of the language at stake contains the words ‘plane’, ‘line’ and ‘incident’.

The vocabulary f_p that is part of the σ-structure $S_p = \langle \Omega_p, f_p \rangle$ representing projective geometry is specified by:

- f_p(‘plane’) = ‘surface of a sphere’
- f_p(‘line’) = ‘great circle (except the equator)’
- f_p(‘incident’) = ‘incident’.

Then the thought expressed in S_0 by ‘Any two lines in the plane are incident’ is false (because of the case of parallel lines), whereas the same sentence, as reinterpreted in S_p, expresses a thought that comes out true, namely the thought *Any two great circles on the sphere, both distinct from the equator, are incident*.

A Fregean σ-structure S consists of a set Ω of expressions, together with a "vocabulary" $f : \Omega \to \Omega_0$ such that $\sigma \subseteq \mathrm{Im}(f)$, where Ω_0 is the set of all expressions of the language L_0.

A morphism $S = \langle \Omega, f : \Omega \to \Omega_0 \rangle \to S' = \langle \Omega', f' : \Omega' \to \Omega_0 \rangle$ of Fregean σ-structures is simply a vocabulary $g : \Omega' \to \Omega$ such that $f' = f \circ g$.

Let $\mathcal{F}$ be the category of Fregean σ-structures.

A σ-structure S is σ-congruent with a σ-structure S', written $S \equiv_\sigma S'$, iff there is a morphism $S \to S'$ of σ-structures such that $\sigma \subseteq \mathrm{Fix}(g)$ —which means that S and S' reinterpret all the expressions of σ in the same way.

The relation of σ-congruence (or its symmetrical closure) is an equivalence relation.

A context (in the narrow sense) is a σ-congruence class of σ-structures, for some implicit signature σ: It is a complete class of σ-structures sharing the same reinterpretation of all the members of σ.

All σ-structures of the same context as S can be regarded as alternative possible worlds in that context.

It is natural to consider, above each σ-structure S, the system R_S mentioning the reference of all simple and complex proper names, as well as the value-ranges of all function-names (of all levels), as these names are reinterpreted by S, together with all the dependencies and connections between them.

This system R_S can easily be given the structure of a category:

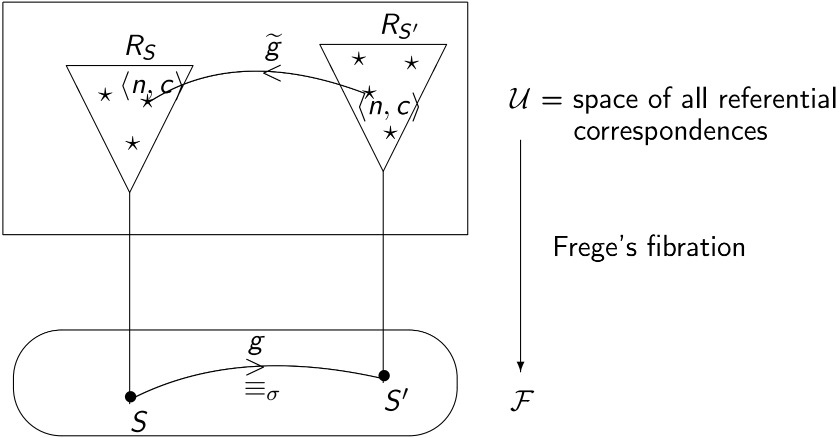
- The objects of R_S are all referential pairs $\langle n, r \rangle$, where either 'n' is a proper name and r its reference, or 'n' is a function-name and r its value-range.
- The arrows of R_S are determined by the following rule: There is an arrow $\langle n, r \rangle \to \langle n', r' \rangle$ each time 'n' forms part of 'n'' or coincides with it. (*Grundgesetze*, §26).

Any morphism $g : S \to S'$ in $\mathcal{F}$ induces a natural mapping $\widetilde{g}$ from $R_{S'}$ to R_S: $\widetilde{g}(\langle n'(x_i'), r' \rangle) = \langle n'(g(x_i')), r \rangle$.

In case two structures S and S' are σ-congruent, each non-logical constant c in the signature σ is preserved by $\widetilde{g}$.

Since the correspondence $S \mapsto R_S$ is "functorial," one gets a fibration over $\mathcal{F}$, Frege's fibration $p_F : \mathcal{U} \to \mathcal{F}$.

Frege's fibration



Vocabulary shift (in Frege's sense) is thus interpreted as reindexing in Frege's fibration.

By way of comparison: Tarski's semantics

Natural category attached to L = the category $\mathcal{T}$ whose objects are all L-structures, and whose arrows $g : M \to M'$ are all L-elementary embeddings $M \prec M'$.

Natural system atached to each L-structure M: the hierarchy D_M of all definable subsets of M, i.e., all the subsets of $|M|^n$ of the form $\phi^M := \{(a_1, \ldots, a_n) : M \vDash \phi[a_1, \ldots a_n]\}$ for some L-formula $\phi(x_1, \ldots, x_n)$.
Each individual constant c is represented in M by the definable subset $(x = c)^M = \{c^M\}$.

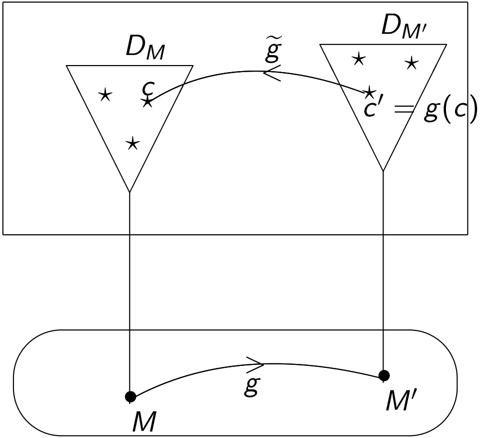
This hierarchy D_M can naturally be turned into a category:

- The objects of D_M are all definable subsets of M.
- The arrows of D_M are all definable maps between them.

Any elementary embedding $g : M \to M'$ induces a mapping $\widetilde{g} : \phi^{M'} \mapsto g^{-1}(\phi^{M'})$ from $D_{M'}$ to D_M.

So, again, the correspondence $M \mapsto D_M$ allows one to define a fibration over $\mathcal{T}$, Tarski's fibration $p_T : \mathcal{V} \to \mathcal{T}$.

Tarski's fibration



$\mathcal{V}$ = space of all definable hierarchies

Tarski's fibration

$\mathcal{T}$

Remarks

1) In Tarski's semantics, each σ-structure M is a context of its own. But M is the unique possible world that can be considered in its own context (in context M).

In two-dimentional parlance, only "diagonal" or "primary intentions" can be represented.

2) Tarski's semantics should not be confused with Tarski's fibration.

In Tarski's semantics, one is always entirely free to shift from an L-structure M to another L-structure M', regardless of the connections between M and M'.
In Tarski's fibration, specific accessibility paths (elementary chains) are prescribed in the base category, and, accordingly, specific connections between fibers are distinguished.

This is what Frege's fibration puts to the fore as well: Each particular signature σ induces a collection of contexts, which correspond to admissible paths from one structure to another.

Non-logical constants as indexicals

Wilfrid Hodges suggested that non-logical constants of a formal language can be seen as indexicals within the space of structures for the language.

For example, the symbol '$\cdot$' of the group operation, in the language L of group theory, is interpreted, in the *context* of each L-structure, by a specific operation in that structure: in any group G, "[...] the symbol '$\cdot$' automatically refers to the operation labelled '$\cdot$' in the group G."

Non-logical constants are given a reference by being applied to a particular structure, exactly as the word 'yesterday' can be given a meaning by being used in a particular temporal context.

Frege's alleged blindness to contextuality

Claim made by Wilfrid Hodges and William Demopoulos: Frege would be unable to account for indexicals (in natural language) and for non-logical constants (of a formalized language).

According to Demopoulos, Frege's blindness to non-logical constants would stem from his blindness to indexicals, which itself would stem from his conception of the reference of designators as being determined in an "absolute" way.

I now would like to challenge this kind of assessment.

Fregean non-logical constants

In Frege's fibration as well as in Tarski's, a non-logical constant becomes nothing but a section of the fibration.

A section of p_F is a functor $s : \mathcal{F} \to \mathcal{U}$ such that $p_F \circ s = \mathrm{id}_{\mathcal{F}}$.

This amounts to the choice, above each object S of the base category $\mathcal{F}$, of a certain item $s(S)$ in the corresponding fiber $\mathcal{U}_S = H_S$, in such a way that, for any two objects S and S', $s(S) = \widetilde{g}(s(S'))$ for some morphism $g : S \to S'$ in $\mathcal{F}$.

The choice made in each fiber is constrained insofar as all the choices thus made should comply with the conditions indicated by the morphisms of the base category.

The definition of a section of Tarski's fibration is analogous.

A section of Frege's fibration consists of a collection of items (referential pairs), one for each σ-structure, so that any two σ-congruent σ-structures share the same item as soon as this item is intended to interpret an expression of σ.

Similarly, a section of Tarski's fibration consists of a collection of items (definable subsets), one for each L-structure, so that any two L-structures with a morphism between them have matching items as soon as these items interpret the same expression of σ.

This view of non-logical constants as sections is almost tautological, since the signature σ is precisely built into those morphisms.

Still, it allows to understand non-logical constants as indexicals in Tarski as well as in Frege.

The role (or character) of a demonstrative, says Perry, is a "rule that takes us from each context of utterance to a certain object."

In the case of a non-logical constant (seen as an indexical), the rule at stake corresponds to the constraint set upon the admissible choices for a section of the fibration.

This stands in sharp contrast with what happens in Tarski's *semantics*: The respective interpretations c^M and $c^{M'}$ of the same constant (or the respective interpretations φ^M and $\varphi^{M'}$ of the same formula) in two different structures M and M'can be considered independently of each other.

A non-logical constant, understood as a section, does not work in the same way in Frege and in Tarski.

A change of context, in Frege, is not a change of place, but an internal reorganization of the language.

By comparison:
Each non-logical constant, in Tarskian *semantics* (as opposed to Tarski's *fibration*), can be reinterpreted independently of the rest of the language.

Accordingly, Tarski's *semantics* conflates contexts with domains and merges both into the mixed notion of "universe of discourse."

This conflation is exactly the target of Frege's criticism of the invocation of "circumstances" of truth:

> *A proposition that holds only under certain circumstances is not a real proposition. [. . .] If we suppose that a proposition can hold under certain circumstances but not under others, then we allow ourselves to be led by the nose by self-induced inexactitudes of expression.*
>
> *("On the Foundations of Geometry", 1906)*

To Frege, the confusion of pseudo-propositions (formal sentences) with real (interpreted) propositions, and the confusion of contexts with circumstances are two facets of the same fundamental confusion.

A change of context is irreducible to a change of domain and to a change of circumstances as well.

To put it another way: In the context of consistency and independence proofs, Hilbert put forward the general concept of reinterpretation of a formal theory.

Tarski's semantics actually constitutes only a particular implementation of Hilbert's concept of reinterpretation.

The Fregean avenue opened by the notion of vocabulary is precisely another one, that the apparatus of fibrations is geared to bring out and to develop.

Back to Hintikka

1) The universality of the language does not preclude one at all from speaking "of how its semantical relations to the world could be changed."
Only, instead of changing the references of the words, one interchanges the meanings of some words.

2) The universality of language does not block at all "a systematic variation of the interpretation of a language."
Only, language is interpreted, and reinterpreting the language is not understood according to a referentialist conception any more, but as an internal operation.

3) The universality of language does not make at all "model theory impossible."
Only, a different kind of model theory is proposed, that has been upstaged by Tarskian model theory even before it could be recognized and developed.

Conclusion

1. Logical universality is perfectly compatible with contextuality.
   Believing that logical universality excludes contextuality amounts to confusing a context variation with a domain variation.
   But a context variation is not a domain variation, unless the principle of Tarskian semantics is presupposed.
2. It is not true that that Frege could make sense neither of non-logical constants nor of the idea of context.
   It is the other way around: Frege's suggestion (in his 1906 "Foundations of Geometry") makes it possible to analyze contextuality in a more satisfying way than Tarski's semantics does.
3. The technical apparatus used to decide what is possible and what is not counts. The apparatus of fibrations allows one both to free interpretation of Frege from a technical artefact and to enrich Tarski's semantics into Tarski'fibration.