Studies in Logic, Vol. 13, No. 3 (2020): 62–81
PII: 1674-3202(2020)-03-0062-20

Logicality and the Logicism of Frege Arithmetic and Simple Type Theory *

Weijun Shi

Abstract. The logicism of Frege and Russell consists of two-fold components: the provability thesis and the definability thesis. It is safe to say that the provability thesis cannot be completely upheld. However, to justify or dismiss them, in particular, the definability thesis, one needs a criterion for logicality to determine whether the constants for the concept "the number of" and the membership relation, among others, are logical. The criteria that I shall adopt are logicality as isomorphism-invariance on the part of Tarski and Sher and logicality as homomorphism invariance on the part of Feferman. Tarski and Sher have pointed out that Russell's constant for the membership relation is isomorphism-invariance on different occasions. I shall demonstrate the following conclusions in the article: First, this constant is also homomorphism invariance; Second, the constant for the concept "the number of" is neither isomorphism invariance nor homomorphism-invariance; Third, if logicality is isomorphism invariance or homomorphism invariance, then the definability thesis of Frege 's logicism (here Frege arithmetic) does not get justified; Forth, if logicality is isomorphism invariance, the thesis of Russell's logicism (here simple type theory) is fully justified, whereas it is not so provided that logicality is homomorphism invariance.

1 Introduction

The logicism of Frege and Russell tries to reduce mathematics, i.e., the fundamental axioms and theorems of number theory, to logic. But how should we make sense of such reduction? According to Carnap, the reduction consists of two theses ([2], p. 41):

∗ The concepts of mathematics can be derived from logical concepts by means of explicit definitions.

∗ The theorems of mathematics can be derived from logical truths by means of logical inferences.

Following Boolos, let us call the first the *definability thesis* and the second the *provability thesis*. These two theses are independent: neither one is weaker or stronger than the other. The reason is that a logical truth can involve extra-logical concepts,

Received 2019-11-12

Weijun Shi Department of Philosophy, Renmin University of China
Department of Philosophy, Humboldt-Universität zu Berlin
2015000867@ruc.edu.cn

*An earlier version of this paper was presented in Professor Michael Beaney's colloquium, held at Humboldt-Universität zu Berlin, in winter semester 2019. I would like to thank the audience for discussion. I am grateful to Michael Beaney for his comments and suggestion.

while a truth involving only logical concepts may not be a logical truth. The examples are not far away to find.[1] It is not controversial that Frege subscribes to both theses. Unlike Frege, Russell, however, is committed himself to the definability thesis, but not to the provability thesis, because, as Boolos points out, he explicitly rejects the axiom of infinity as a logical truth in *Principia Mathematica* ([4], pp. 270–272). Which thesis Frege and Russell subscribe to is of historical significance. But it is of greater significance whether these two theses are on their own justifiable.

What is indispensable to the justification or rejection of the theses is a criterion for logicality and a notion of logical truth. In Frege's or Russell's system, according to the definability thesis, the definiens, in terms of which the constants of mathematics are defined, are logical in the sense that it involves only logical vocabulary thereof. According to the provability thesis, the axioms of these two systems are logical truths. Thus, logicism will stand and fall with different criteria for logicality and notions of logical truth. In this paper, we will make use of Tarski's notion of logical truth, which is the standard one in logical textbooks.[2] It is a precondition for this notion that the constant of languages is divided into logical and extra-logical ones. In other words, the criterion for logicality is a precondition for Tarski's conception of logical truth. Let us see why it is so.

According to Tarski, any sentence ϕ of a language $\mathcal{L}$ is a logical truth if its sentential function ϕ^* is true in all structures of the sentential function. The notion of logicality comes in when we transform a sentence into its sentential function. Suppose all constants (except variables) of the language have type symbols in $\mathcal{T}$, which is defined as follows:

Definition 1

1. $e \in \mathcal{T}$;
2. If $\tau_1, \ldots, \tau_n \in \mathcal{T}$, then $(\tau_1, \ldots, \tau_n) \in \mathcal{T}$.

Given a domain M of objects of type e, with each type $\tau \in \mathcal{T}$ is associated a set M_τ which is defined as follows:

Definition 2

1. If $\tau = e$, then $M_\tau = M$;
2. If $\tau = (\tau_1, \ldots, \tau_n)$, then $M_\tau = P(M_{\tau_1} \times \ldots \times M_{\tau_n})$.

where P is the subset operation in set theory.

Let $M_{\mathcal{L}}$ be the built-in domain of the language $\mathcal{L}$. Suppose the extra-logical con-

[1] '$\exists x \exists y(x \neq y)$' does not contain extra-logical constants in the usual sense, but it is not a logical truth. '$\forall x(x = \mathit{Russell} \vee \neg(x \neq \mathit{Russell}))$' is a logical truth, even though it contains the extra-logical constant 'Russell'.

[2] Besides Tarski's notion of logical truth, there are others, say, Quine's substitutional conception of logical truth ([17], ch. 4).

stants occurring in ϕ are $E_1,\ldots, E_n$, which are of types τ_i ($1 \leqslant i \leqslant n$), respectively. The sentential function ϕ^* of ϕ is obtained by substituting a variable X_i of type τ_i for each occurrence of E_i in ϕ. ϕ is a logical truth if ϕ^* is true in all structures.[3]

Tarski's notion of logical truth is sensitive to the notion of logicality, as Tarski himself comes to recognize.[4] In order to see this sensitivity, let us look at an example. Suppose $\mathcal{L}$ contains 'Russell', 'philosopher', '$\exists$' and '$=$' as its only primitive constants, and 'Russell' and 'philosopher' denote Betrand Russell and the property *being a philosopher*, respectively. So the built-in domain $M_{\mathcal{L}}$ of $\mathcal{L}$ contains at least Russell, and the property is a subset of the domain which contains all philosophers. Consider the sentences :

(1) Russell is a philosopher.
(2) $\exists x(x = x)$.

They are true in the structure with the domain $M_{\mathcal{L}}$. Usually, 'Russell' and 'philosopher' are seen as extra-logical, while '$\exists$' and '$=$' are seen as logical. As a result, (1) is not seen as a logical truth, while (2) is seen as a logical truth, which accords with our usual or intuitive conception of logical truth.

(1) and (2) will be logical truths, if all of these primitive constants are logical. In this case,the sentential functions of (1) and (2) are themselves, because there are no extra-logical constants occurring in them that are to be substituted for by variables. Consequently, (1) and (2) are true in all structures with any domain. One might contest this by saying that (1) is false in the structure with the domain, say, $\{1, 2\}$, in which 'Russell' and 'philosopher' are interpreted as 1 and $\{2\}$, respectively. This is,

---

[3] It is controversial whether Tarski takes all structures to share the same domain $M_{\mathcal{L}}$, i.e., the built-in domain of the language $\mathcal{L}$ in [20]. Bays argues that Tarski intends all structures to share the same domain$M_{\mathcal{L}}$ ([1]), while others like Sher argues for the exactly opposite ([18], p. 41). Whether all structures share the same domain $M_{\mathcal{L}}$ is of great significance to which sentence of the language is valid. Suppose the built-in domain $M_{\mathcal{L}}$ of $\mathcal{L}$ has at least two objects. So the sentence '$\exists x \exists y(x \neq y)$' of $\mathcal{L}$ says that there are at least two objects in $M_{\mathcal{L}}$. Given that all constants in the sentence are logical, the sentence is a logical truth if all structures share the same domain $M_{\mathcal{L}}$. In contrast, if all structures do not necessarily share the same domain, then the sentence is not valid, because its sentential function, i.e., itself, would be false in all structures whose domains contain only one object. It is not up to me to determine which reading is right here. I just take it that the domain can vary. For the point that the notion of logical truth is sensitive to logicality, which I want to emphasize, the controversy is not that much important.

[4] According to Tarski, a sentence ϕ is the logical consequence of the sentences of the set $\mathfrak{K}$ iff any model of the sentential functions of the sentences of $\mathfrak{K}$ is also a model of the sentential function of ϕ. An extreme case of the notion of logical consequence is such that $\mathfrak{K}$ is empty. In such case, we have the notion of logical truth. Now Tarski writes:

> The extreme would be the case in which we treated all terms of the language as logical: the concept of following formally[5] would then coincide with the concept of following materially-the sentence X would follow from the sentences of the class $\mathfrak{K}$ if and only if either the sentence X would be true or at least one sentence of the class $\mathfrak{K}$ were false. ([20], p. 188)

however, not true, because ‘Russell’ and ‘philosopher’ are logical, so that they are to denote Russell and the property *being a philosopher*. So it is absolutely wrong to assign 1 and $\{2\}$ to them, respectively. In other words, any structure with the domain $\{1, 2\}$ is just not a structure of the sentential functions of the language $\mathcal{L}$ at all.

In contrast, if all of these primitive constants are extra-logical, then both (1) and (2) are not logical-truths. ‘Russell’, ‘philosopher’, ‘$\exists$’ and ‘$=$’ are of types e, (e), $((e))$ and (e, e), respectively. Let ‘X_1’, ‘X_2’, ‘X_3’ and ‘X_4’ be variables of these types, respectively. So ‘$X_2(X_1)$’ can serve as the sentential function of (1), while ‘$X_3x(xX_4x)$’ is the sentential function of (2). Then it is easy to find structures in which these two sentential functions are not satisfied. Consequently, (1) and (2) are not logical truths.

What distinguishes logical constants from extra-logical ones is that the former are to be interpreted uniquely or in the same way on all structures, while the latter do not have to be so. In Tarski’s original formulation of the concept of logical truth, it is required that each sentence be transformed into its sentential function. This requirement, however, is not necessary. For to assess the validity of a sentence, we can interpret all of its logical constants as their default denotations or meanings on all structures, and all of its extra-logical constants arbitrarily with their types being respected; there is no need to replace those extra-logical constants by variables of the same types, and then assign objects of the same types to these variables. For example, given that ‘$\exists$’ is logical, it is to be interpreted as $\{A : A \subseteq M, A \neq \emptyset\}$ on each structure with any domain M, in order to determine whether (say) ‘$\exists x(x = x)$’ is valid. It cannot be assigned other denotations, say, $\{M\}$. Similarly, if ‘Russell’ is logical, it is to be interpreted as the person Betrand Russell on each structure. (In this connection, Russell must be in the domain of the structure.) If a constant is extra-logical, then besides its default denotation, it can be interpreted as any objects of the same type. E.g. if ‘$\exists$’ is extra-logical, then it can be interpreted as objects of the type $((e))$, say, $\{A : A \subseteq M\}$ and $\{M\}$, on any structure with a domain M. By the same token, if ‘Russell’ is extra-logical, then it can be interpreted as any object of type e on all structures, even if Russell is not in their domains.[6]

[5]Here, by ‘following formally’ is meant what we call ‘logical consequence’ today.

[6]When I write these three paragraphs, I fully keep in mind Sher’s interpretation of what distinguishes extra-logical constants from logical ones, which does not accord with my reading. Sher writes:

> It has been said that to be a logical constant in a Tarskian logic is to have *the same* interpretation in all models. Thus for “red” to be a logical constant in logic $\mathcal{L}$, it has to have a constant interpretation in all the models for $\mathcal{L}$. I think this characterisation is faulty because it is vague. ([18], p. 45)

I do not hold that from the perspective of Tarski, a logical constant such as ‘$\exists$’ should have *the same* interpretation in all models. Clearly, given any two nonidentical domains (sets), the interpretations of this quantifier cannot be the same in the set-theoretical sense. Rather, I hold that logical constants should be interpreted uniquely or in the same way. What the words ‘uniquely’ or ‘in the same way’ means is already glossed by taking the example of ‘$\exists$’ as logical in the paragraph. Now Sher asks how one can interpret ‘red’ (assume that the predicate denotes the set of all red objects) in the same way in all models

2 Frege's logicism

Frege's original system in [11], via. second-order logic plus the axiom V, is inconsistent; but the system can be rendered as consistent in different ways. For example, predicative second-order logic plus the axiom V is consistent ([13]). But it is too weak to derive Peano arithmetic (PA)[7] as its theorems ([13])[8]. In this paper, I will not consider the system for two reasons. The first is its weakness; the second is that from the perspective of ZFC, which we are going to make use of as the metatheory, the axiom V is just false. Instead, I will consider another version of Frege's original system, *Frege arithmetic* (FA), which is consistent and strong enough for PA ([4], pp. 183–201).

The language of FA is a second-order language, which contains the following as its primitive vocabulary:

(i.1) Variables: variables for concepts and relations and objects.
(i.2) Constants: first and second-order quantifiers, propositional connectives, identity.
(i.3) Constants: $\sharp$.

Terms, formulas and sentences are defined in a canonical way. '$\sharp$' is a functional constant of type $((e), e)$, which is informally read as 'The number of'. Syntactically, attaching a variable 'F' for concepts to '$\sharp$' forms the term '$\sharp F$' for objects.

Besides all axioms of second-order logic, FA has two more axioms :

(HP) $\forall F \forall G(\sharp F = \sharp G \leftrightarrow F\, eq\, G)$.[9]
(3) $\forall F \exists ! x(x = \sharp F)$.

What (HP) expresses is that the numbers of two concepts are equal iff the concepts are equinumerous. (3) says that every concept has a unique number.

Just as the first-order PA has a standard interpretation or standard model ([16],

if it is a logical constant. My answer is that if 'red' is to be counted as logical, then a model whose domain does not contain all red objects just fails to qualify as a model of the logical language in which 'red' is identified as logical. In the case, not all sets can serve as the domain of models of the language. For Sher's point of view on the question of the roles of logical constants, see [18, pp. 46–52].

[7]PA has five axioms :

PA.1 Zero is a natural number.
PA.2 If x is a natural number, its successor is a natural number.
PA.3 If x is a natural number, its successor is not zero.
PA.4 If x and y are natural numbers and their successors are equal, then x is y.
PA.5 Induction axiom.

[8]In the paper, Heck shows that the induction axiom is not provable from the system. But what is known as Robinson arithmetic is interpretable in the system.

[9]Here '$F\, eq\, G$' is an abbreviation of a second-order formula which says that the concepts F and G are equinumerous.

p. 160), FA also has its standard interpretation. The denotation of the constant '$\sharp$' in the standard interpretation is its default denotation. Give a domain M, the domains of object variables, concept variables and binary relation variables are M, $P(M)$, and $P(M \times M)$, respectively. The constants of (i.2) are interpreted in the canonical way given that they are logical. As for '$\sharp$', its denotation is a function from $P(M)$ to M. But what function is the default denotation of the constant?

For Frege, (HP) is true, if not a logical principle. Frege sees the axiom V as 'a fundamental law of logic' ([10], p. 142) and derives (HP) as a theorem of it (plus necessary definitions) in second-order logic. Hence Frege takes (HP) to be true. (Admittedly, the axiom V is inconsistent in second-order logic. But what concerns us here is not the problem of consistency, but whether or not Frege thinks of (HP) as true.) Therefore, the default denotation of '$\sharp$' is such a function that renders (HP) true. There is no such function in any structure with a finite domain ([4], pp. 213–214, pp. 305–306), because if there are n (any finite number) objects in the domain, there will be $n+1$ equinumerous concepts. Hence its default denotation is any function $\boldsymbol{f} : P(M) \to M$ that satisfies the condition :

(C) For any $F, G \in P(M)$, if $\boldsymbol{f}(F) = \boldsymbol{f}(G)$ iff $|F| = |G|$.[10]
M is infinite.

To assess if Frege's logicism succeeds, it is necessary to determine the logicality of the constant '$\sharp$' with its default interpretation $\boldsymbol{f}$, as well as of the constants in (i.2). Boolos holds that:

(A) The constants in (i.2) are logical.[11]
(B) The constant '$\sharp$' is extra-logical ([4], fn. 3, p. 186).[12]

To survey whether (HP) is a logical truth, '$\sharp$' is to be assigned any function of type $((e), e)$ in the structure with any domain, thanks to (B). (HP) is false in the structure with the domain M, either finite or infinite, in which '$\sharp$' is assigned the function g that for any $A, B \in P(M)$, $g(A) = g(B)$. Therefore, (HP) is not a logical truth. Besides, the rest axioms of FA except (HP) are logical truths.

Boolos's assessment of Frege's logicism is, therefore, as follows:

Premise (A) and (B) hold.
Conclusion The definability thesis and the provability thesis of Frege's logicism are not justified.

The reason that the provability thesis is not justified is that (HP) is not a logical truth.[13]

[10] $|F|$ is the cardinality of the set F.

[11] Boolos does not explicitly hold (A). But (A) is usually taken for granted.

[12] Here Boolos explicitly says that he sees the constant as extra-logical, while Frege sees it as logical.

[13] (PA. 1), (PA. 2) and (PA. 5) are theorems of second-order logic (with additional definitions) plus (3). The derivation of (PA. 3) and (PA. 4), which requires the existence of infinite objects, needs (HP). Consequently, (PA. 1), (PA. 2) and (PA. 5) are logical truths, while (PA. 3) and (PA. 4) are not.

But why is the definability thesis is said to be unjustifiable? It seems that Carnap and Boolos both subscribe to such an inference: If the definition of any arithmetical constant involves extra-logical constants, then the former is extra-logical; since the constant '$\sharp$' is extra-logical, so is any arithmetical constant whose definition involves it. But the claim (+) that if the definition of any arithmetical constant involves extra-logical constants, then the former is extra-logical is not guaranteed at all. Whether the claim holds up depends upon what it is for a constant to be logical. As will be seen later, although all constants in the language of Russell's simple type theory are logical in the sense that all of them are homomorphism-invariance, almost all arithmetical constants for numbers are not homomorphism-invariance.

Before introducing Russell's system, let us pause for a while to look at Carnap's assessment of the definability thesis. Carnap, as opposed to Boolos, thinks that the definability thesis is fully justified. Carnap says that Frege and Russell arrive at the same conclusion about the logical status of natural numbers independently: natural numbers are 'logical attributes which belong, not to things, but to concepts. ([2], p. 42) To be precise, to say that the number of a concept F is n just means that there are exactly n objects which fall under F. 'The number of the concept F' is repressed by Carnap as '$n_m(F)$'. According to Carnap, '$n_m(F)$' is defined as:

$$\exists x_1 \dots \exists x_n \begin{bmatrix} x_1 \neq x_2 \wedge \dots \wedge x_{n-1} \neq x_n \wedge F(x_1) \wedge \dots \wedge F(x_n) \wedge \\ \forall y(F(y) \to y = x_1 \vee \dots \vee y = x_n) \end{bmatrix}$$

Carnap regards all familiar logical constants of first-order logic as logical, without saying anything on second-order quantifiers. ([2], p. 42) (Of course, Carnap owes us a criterion for logicality.) Since the definiens is a formula of first-order logic, Carnap asserts that the notion 'the number of the concept F' is logical. In addition, he holds that other arithmetical concepts such as addition and natural numbers also can be defined by formulae which involve only logical vocabulary. But I think that Carnap's assertions are not justified for two reasons.

First, the aforementioned definiens involves not only constants which Carnap sees as logical but also the constant 'F', which is a constant rather than a variable. The constant is not on Carnap's list of logical terms and is usually seen as extra-logical. Second, Carnap asserts that other arithmetical concepts are definable in terms of the logical terms of first-order logic. But it is hardly clear how this can be done.[14] Take

[14]Hodges ([14]) forcefully argues that the arithmetic truths, say,

(4) $7 + 5 = 12$.
(5) The successor of 1 equals to 2.
express the same thought (Gedanken) as
(6) $\forall X \forall Y[\exists_7 x Fx \wedge \exists_5 x Gx \wedge \neg \exists x(Fx \wedge Gx) \to \exists_{12} x(Fx \vee Gx)]$.
(7) $\forall F(\exists_1 x Fx \wedge \exists y \neg Fy \to \exists G \exists_2 x Gx)$.

(Hodges says clearly that (4) expresses the same thought as (6). That (5) expresses the same thought the concept 'natural numbers' as an example. Taking into consideration that Carnap has defined '$0_m(F)$', '$1_m(F)$' and so on, one might suggest defining the concept as '$x = 0_m(F) \vee \ldots \vee n_m(F) \vee \ldots$'. However, the definiens is not a sentence of first-order logic, because it contains infinitely many disjuncts.

3 Russell's logicism

Simple type theory (STT), which is a simplification of the ramified system in *Principia Mathematica*, is consistent. Moreover, it is so strong that not only PA but also set theory can be devolved out of it. STT is a higher-order system. For deriving PA, STT can be seen as a forth-order system. The language of STT contains the following vocabulary:

(ii.1) Variables 'x^n' ($0 \leqslant n \leqslant 3$) for objects of type n.
(ii.2) First-order quantifiers ('$\forall x^0$'), second-order quantifiers ('$\forall x^1$'), third-order quantifiers ('$\forall x^2$'), forth-order universal quantifiers ('$\forall x^3$') and propositional connectives.
(ii.3) The constant 'ϵ' (for membership).

The atomic formula of the language is of the form '$x^n \epsilon x^{n+1}$' (informally, x^n is a member of the set x^{n+1}).

STT has three axioms, i.e., the axiom of extension and the axiom of comprehension and the axiom of infinity ([16], pp. 290–291):

$$\text{(Inf)} \;\; \neg\forall x^3[0^2 \epsilon x^3 \wedge \forall x^2(x^2 \epsilon x^3 \rightarrow sx^2 \in x^3) \rightarrow \emptyset^2 \epsilon x^3].$$ [15]

(Inf) says that $\emptyset^2$, the empty set of type 2, is not a number, which is equivalent to say that there are infinitely many objects of type 0.

STT has its standard interpretations. Suppose $M_0 = M$ is the domain (any non-empty set) of a structure. Let $M_{n+1} = P(M_n)$. Then each variable 'x^n' ranges over the set M_n. Since the constant 'ϵ' is of type $(n, n+1)$, its denotation must be some relation that is a subset of $M_n \times M_{n+1}$. The standard or default denotation of the constant is the relation $R = \{(a, b) : a \in M_n, b \in M_{n+1}, a \in b\}$. Among all

as (7) is not mentioned by Hodges. But it is clear that this assertion is made in the spirit of Hodges.) (6) and (7) are sentences of second-order logic. So no appeal can be made to these examples to back up Carnap's assertion that all arithmetical concepts can be defined in terms of formulae of first-order logic.

[15] The original formulation of the axiom of infinity in [16, pp. 290–291] is not (Inf). But it is equivalent to (Inf). (Inf) contains three constants '0' (zero of type 2), 's' (successor) and '$\emptyset^2$' (empty set of type 2). The occurrence of these constants can be eliminated from (Inf) by means of the definitions of them as follows:

1. $x^n \epsilon \emptyset^{n+1}$ iff $x^n \neq x^n$.
2. $x^2 = 0^2$ iff $x^2 = \{\emptyset^1\}$.
3. $s(x^2)$ iff$\{y^1 : \exists z^0[z^0 \epsilon y^1 \wedge y^1 \setminus \{z^0\} \epsilon x^2\}]$.

For more details, see [16, pp. 290–293].

relations R is the relation $\{(a, b) : a \in M_n, b \in M_{n+1}, a \in b, M \text{ is infinite}\}$, which should be the default interpretation of the constant 'ϵ' if Russell sees (Inf) as a truth. However, Russell sees (Inf) only as 'an arithmetical hypothesis' ([4], p. 268), so that he does not assert that it is a truth, let alone a logical truth. So the default denotation of the constant is any R.

In the same vein, to assess whether the definability thesis and the provability thesis of Russell's logicism succeed, we have to determine the logicality of the constants in (ii.2) and (ii.3). In this connection, it is generally held that:

(A*) The constants in (ii.2) are logical.

(B*) The constant 'ϵ' is logical.

Now it might be problematic to say that 'ϵ' is generally seen as extra-logical. For in logical books, say, [12, ch. 4] and [16, pp. 289–293] in which the theory of types is discussed in detail, the authors do not explicitly mention whether this constant is logical. But I suggest that it is at least seen as logical by some authors. For example, Boolos remarks that 'one who accepts the theory of types will almost surely regard Infin ax[16] as true and in logical vocabulary'. ([4], p. 271) Another case that can be made for (B*) has something to do with the interpretation of the constant on structures. According to Hatcher, the constant is assigned the relation $R = \{(a, b) : a \in M_n, b \in M_{n+1}, a \in b\}$ on each structure with the domain M. This indicates that Hatcher sees it as logical, otherwise it should be interpreted as any relation of type $((e), e)$.

Because of (A*) and (B*), the constants in (ii.2) are interpreted on each structure in a canonical way, and 'ϵ' is interpreted as $R = \{(a, b) : a \in M_n, b \in M_{n+1}, a \in b\}$ on each structure. As a result, it is easy to show that the axioms of STT except (Inf) are logical truths. (Inf) is not a logical truth, because it is false in any structure with a finite domain M.

Concerning Russell's logicism, it is generally held that

(Premise*) (A*) and (B*) hold.

(Conclusion*) The provability thesis of Russell's logicism is not justified, while the definability thesis thereof is justified.

The first part of (Conclusion*) holds, because (Inf) is not a logical truth.[17] When it comes to the second part of (Conclusion*), the claim (+) on p. 68 is taken for granted. As is already pointed out, whether the claim holds is completely determined by the conception of logicality. Anyone who wants to defend (Conclusion*) and (Conclusion) has to offer a criterion for logicality which can justify (Premise*) and (Premise). Now I will turn to such criterion next.

[16]By Infix ax is meant the axiom of infinity.

[17]All axioms of PA except (PA. 4) are theorems of the axioms of extension and comprehension. So the former is logical truths. The derivation of (PA. 4) from STT needs (Inf). Since (Inf) is not a logical truth, so is (PA. 4).

4 Logicality as invariance under morphisms

There are different approaches to logicality. Among them is Tarski's approach, according to which logicality is to be characterised as invariance under isomorphisms. This approach has the advantage of mathematical precision. For this reason, I will look at if it can be invoked to justify (Premise) and (Premise*).

The motivation underlying Tarski's approach is that logical notions are formal in the sense that they do not distinguish objects, i.e., they are insensitive to the switching of objects. Suppose M is a set of objects of type e. let f be any bijection on M. Then, for any object O of type $\tau \in \mathcal{T}$, O is mapped under f to another object $f(O)$. Tarski suggests that O is a logical object if it is invariant under any transformation f, i.e., $O = f(O)$ ([22]). A constant is logical if its reference is a logical object in Tarski's sense. However, unlike Tarski, logicians and philosophers after him tend to speak of logicality of operations across domains, rather logicality of objects on a single domain. But these two different modes of dealing with logicality are more or less the same in their spirits. I will follow the mainstream in the paper.

There are three variants of the approach at issue: Tarski and Sher take logicality to be isomorphism-invariance; Feferman takes it to be homomorphism-invariance; Bonnay takes it to be potential isomorphism-invariance. For any constant whose type level[18] is $\leqslant 2$, all of these variants fit for deciding their logicality. But for any constant whose type level is larger than 2, the third variant is not applicable without being generalised. The reason is related directly to the definition of potential isomorphisms, as is to be explained as follows. For this reason, I shall not consider the third variant, because some constants of the language of STT and FA have type levels larger than 2.

To formulate the content of the first two variants, we need the definitions of the concepts 'the denotation (or operator) of a constant associated with each domain', 'argument-structures', 'isomorphisms' and 'homomorphisms'. These definitions, or more exactly, the ways of formulating these definitions, which I will make use of below, are due to Sher.

Given a domain M(infinite on some occasion), each constant $\mathcal{C}$ of type τ has a unique denotation $f_{\mathcal{C}}(M)$ on M, which is a member of the set M_τ.[19] By a domain M is always meant a non-empty set of objects of type e. For simplicity, we say that $f_{\mathcal{C}}$ is the denotation of $\mathcal{C}$ on M. If the constant $\mathcal{C}$ is of type τ, we also say that its denotation $f_{\mathcal{C}(M)}$ is of type τ. Let us give some examples.

---

[18]The level of types are defined as follows. $\mathit{level}(e) = 0$. For any type $\tau = (\tau_1, \ldots, \tau_n) \in \mathcal{T}$, $\mathit{level}(\tau) = \max(\mathit{level}(\tau_1), \ldots, \mathit{level}(\tau_n)) + 1$.

[19]Sher requires that each constant should have a unique denotation on *every domain*. As a result, the predicate 'red' is to have a denotation on a domain containing no red objects at all. I just require that it have a denotation on a domain containing all red objects (and possibly others). So the *any isomorphic (or homomorphic) structures* in (Tarski–Sher Thesis) and (Feferman Thesis) below are just such that their domains contain the default denotation of the constant $\mathcal{C}$. I deviate from Sher's way of assigning objects to constants.

- $f_{Russell}(M) = Russell$, if $Russell \in M$.
- $f_{=}(M) = \{(a, b) : a \in M, b \in M, a = b\}$.
- $f_{\forall}(M) = \{M\}$.[20]
- $f_{\exists}(M) = \{A : A \subseteq M, A \neq \emptyset\}$.
- $f_{\exists^2_k}(M) = \{A : A \subseteq P(M^k), A \neq \emptyset\}$.[21]
- $f_{\forall^2_k}(M) = \{P(M^k)\}$.[22]
- $f_{\sharp}(M) = \boldsymbol{f}$, where $\boldsymbol{f}$ satisfies the condition:
  (C) For any $A, B \in P(M)$ and $a, b \in M$, if$(A, a) \in f_{\sharp} \wedge (B, b) \in f_{\sharp}$, then $|A| = |B|$ if and only if $a = b$.
- $f_{\epsilon}(M) = \{(a, b) : a \in M_{\tau}, b \in M_{(\tau)}, a \in b\}$, where $\tau = (\ldots(e)\ldots)$ (n pairs of parenthesis, $n \geqslant 0$).[23]

Given any set M (maybe infinite on some occasions), with each vocabulary $\mathcal{C}$ of type $\tau = (\tau_1, \ldots, \tau_n)$ is associated a $n{+}1$-tuples $\langle M,\ a_1 \ldots a_n\rangle$, where each a_i is of type τ_i and therefore is an element of the set M_{τ_i}. Each a_i is the i-th argument of the denotation $f_{\mathcal{C}}(M)$ of $\mathcal{C}$ on M. Following Sher, let us call the tuple $\langle M,\ a_1 \ldots a_n\rangle$ *the argument-structure of the constant* $\mathcal{C}$ *on* M. For example:

- The argument-structure of 'Russell' is $\langle M, a\rangle$, where Russell is a member of M and $a \in M$.
- The argument-structure of '$=$' is $\langle M, a, b\rangle$, where $a, b \in M$.
- The argument-structure of '$\forall$' is $\langle M, A\rangle$, where $A \in P(M)$.
- The argument-structure of '$\sharp$' is $\langle M, A, a\rangle$, where $A \in P(M)$, $a \in M$ and M is infinite.
- The argument-structure of 'ϵ' is $\langle M, a, A\rangle$, where $a \in M_{\tau}$ and $A \in M_{(\tau)}$.

With the notion 'argument-structures' at our disposal, we can define the concepts of isomorphic- and homomorphic argument-structures. Suppose M and N are two universes and f is a bijection from M to N. Then, the bijection will induce a bijection from the set M_{τ} to the set N_{τ} for any type $\tau = (\tau_0, \ldots, \tau_n)$.

---

[20]The first-order universal and existential quantifiers are of type $((e))$.

[21]This kind of identification of k-ary second-order existential quantifier $\exists^2_k$ is due to J. Kontinen ([15]). Originally, given a domain M, J. Kontinen identifies the denotation of $\exists^2_k$ as the set $\{(M, A) : A \subseteq P(M), A \neq \emptyset\}$. In order to keep the uniformity, we change Kontinen's original way of identifying the denotation of constants into the current one. Such change is not essential, though. Kontinen says that 'the first example is the familiar k-ary second-order existential quantifiers.' This remarks shows that with respect to second-order quantifiers, we should not speak of their denotations generally; rather we should speak of the denotation of a second-order quantifier binding a variable for k-aries relations. The reason is that in (say) '$\exists x$' x denotes any object of type e, while in the second-order quantification '$\exists X$' X denotes any k-ary relations of type $(e, \ldots, e)$ with k e's.

[22]Here I follow Kontinen in his spirit in identifying the second-order universal quantifier in the framework of relational types.

[23]The type symbol n is equivalent to $(\ldots(e)\ldots)$ in the system of relational type symbols.

Definition 3

1. If $\tau = e$, then $f_\tau = f$.
2. If $\tau \neq e$, then $f_\tau(A) = \{(f_{\tau_0}(a_0), \ldots, f_{\tau_n}(a_n)) : (a_0, \ldots, a_n) \in A)\}$, where $A \in M_\tau$.

Definition 4 Any two argument-structures $\langle M, a_1, \ldots, a_m\rangle$ and $\langle N, b_1, \ldots, b_n\rangle$, where a_i and b_i are of the same type $\tau = (\tau_0, \ldots, \tau_n)$, are isomorphic, if and only if:

1. $m = n$;
2. There is a bijection f from M to N;
3. For each $1 \leqslant i \leqslant m$, $f_\tau(a_i) = b_i$, where f_τ is defined in Definition 3.

We can drop the subscript for types, writing $f(A)$ instead of $f_\tau(A)$, because the type of A makes it clear which function is at issue.[24] Suppose M and N are two universes and f is a surjection from M to N. Then, the bijection will also induce a surjection from the set M_τ to the set N_τ for any type $\tau = (\tau_0, \ldots, \tau_n)$.

Definition 5

1. If $\tau = e$, then $f_\tau = f$.
2. If $\tau \neq e$, then the domain of f_τ, in notation, *Dom*(f_τ), is defined as follows:
   - $Dom(f_\tau) = \{A \subseteq M_{\tau_1} \times \ldots \times M_{\tau_n} : \forall a_1, b_1 \in Dom(f_{\tau_1}) \ldots \forall a_n, b_n \in Dom(f_{\tau_n})[f_{\tau_1}(a_1) = f_{\tau_1}(b_1) \wedge \ldots \wedge f_{\tau_n}(a_n) = f_{\tau_n}(b_n) \Rightarrow (a_1, \ldots, a_n) \in A \leftrightarrow (b_1, \ldots, b_n) \in A]\}$.
   - $f_\tau(A) = \{(f_{\tau_1}(a_1), \ldots, f_{\tau_n}(a_n)) : (a_1, \ldots, a_n) \in A \cap Dom(f_{\tau_1}) \times \ldots \times Dom(f_{\tau_n})\}$.[25]

Often we do not care what the induced function on M_τ with a very complex type τ is. Rather, what concerns us is simple types. For example, for the type $(e, \ldots, e)$, $f_{(e,\ldots,e)}$ is defined as follows:
$f_{(e)}(A) = \{(f(a_1), \ldots, f(a_n)) : (a_1, \ldots, a_n) \in A\}$, and its domain is $\{A \subseteq M : \forall a_i, b_i \in M, f(a_i) = f(b_i) \Rightarrow (a_1, \ldots, a_n) \in A \leftrightarrow (b_1, \ldots, b_n) \in A\}$.

Definition 6 Any two structures $\langle M, a_1, \ldots, a_m\rangle$ and $\langle N, b_1, \ldots, b_n\rangle$, where a_i and b_i are of the same type $\tau = (\tau_0, \ldots, \tau_n)$, are homomorphic, if and only if:

1. $m = n$;
2. There is a surjection f from M to N.
3. For each $1 \leqslant i \leqslant m$, $f_\tau(a_i) = b_i$, where f_τ is defined in Definition 5.

[24] For example, suppose that $M = \{1, 2\}$, $N = \{a, b\}$ and $f(1) = a$, $f(2) = b$. Then, each set of type (e), say, the set $\{1, 2\}$, will be mapped to the set $\{a, b\}$. Each set of type $((e))$, say, the set $\{\{1\}\}$, will be mapped to the set$\{\{a\}\}$, etc.

[25] By '$\Rightarrow$' and '$\leftrightarrow$' are meant 'if, then' and 'if and only if'.

Definition 6 is very different from Sher's formulation of homomorphic argument-structures in [19, p. 334]. The difference lies in that the induced function that Sher makes use of in her formulation is not the one defined in Definition 5. Rather, it is the one that results from Definition 3 by replacing the bijection f therein with a surjection. I shall refer to this resultant definition as 'Definition*' later. Definition* is not loyal to Feferman's concept of homomorphism-invariance which is defined by using functional types ([8]).[26]

Suppose M and N are any two domains, $\mathcal{C}$ is a constant of type $\tau = (\tau_1, \ldots, \tau_n)$, and the denotations of $\mathcal{C}$ on M and N are $f_{\mathcal{C}}(M)$ and $f_{\mathcal{C}}(N)$, respectively. $\mathcal{M} = \langle M, a_1, \ldots, a_m\rangle$ and $\mathcal{N} = \langle N, b_1, \ldots, b_n\rangle$ are the argument-structures of $\mathcal{C}$ on M and N, respectively. According to Tarski and Sher ([18, 19, 22]), $\mathcal{C}$ is a logical constant iff $\mathcal{C}$ is invariant under isomorphisms. To be precise,

> <u>Tarski–Sher Thesis</u>
> $\mathcal{C}$ is *logical*$_{TS}$ if and only if
> For <u>any</u> two isomorphic argument-structures $\mathcal{M} = \langle M, a_1, \ldots, a_m\rangle$ and $\mathcal{N} = \langle N, b_1, \ldots, b_n\rangle$, $(a_1, \ldots, a_n) \in f_{\mathcal{C}}^M$ if and only if $(b_1, \ldots, b_n) \in f_{\mathcal{C}}^N$.

It is worthy of mentioning that Sher imposes such a restriction on the constant $\mathcal{C}$ ([18], p. 54) that its level is $\leqslant 2$. We will not respect the restriction, because some of the constants of the language of FA and STT have levels larger than 2.

Feferman rejects the thesis, because, among others, it overgeneralises the domain of logical notions by admitting many mathematical and set-theoretical notions such as alephs. In order to narrow the expanded field, Feferman uses homomorphism-invariance to characterize logicality ([7, 8, 9]) :

> <u>Feferman Thesis</u>
> Any monadic $\mathcal{C}$ is *logical*$_F$ if and only if
> For <u>any</u> two homomorphic argument-structures $\mathcal{M} = \langle M, a_1, \ldots, a_m\rangle$ and $\mathcal{N} = \langle N, b_1, \ldots, b_n\rangle$, $(a_1, \ldots, a_n) \in f_{\mathcal{C}}(M)$ if and only if $(b_1, \ldots, b_n) \in f_{\mathcal{C}}(N)$.

Compared with (Tarski–Sher Thesis), Feferman's thesis narrows the field of logic notions dramatically. For example, all cardinal quantifiers '$\exists_n$', either $n < \aleph_0$ or $n \geqslant \aleph_0$, fail to pass the test of the thesis. In addition,the constant for identity, which is usually seen as a logical vocabulary, is also excluded by the thesis from the field of logic. Having said that, there are still certain constants for set-theoretical notions, say, 'being a well-ordering' and 'being well-founded', that survive the thesis.

Faced with these two shortcomings of (Feferman Thesis), i.e., the exclusion of identity and the inclusion of the aforementioned set-theoretical notions, Feferman

[26] The concept of homomorphism invariance can be specified in terms of relational types or functional types. For more on the relation between these two methods of specifying the concept, see [5, p. 18, p. 21].

turns to Quine's way of handling identity in a language with standard grammar ([17], ch. 5) and the notion of polyadic quantifiers, respectively. Bonnay, while agreeing with Feferman that (Tarksi-Sher thesis) overgeneralises logic, rejects the expedient of appealing to the notion of polyadic quantifiers. Bonnay suggests characterizing logicality in terms of potential isomorphism-invariance ([3]). However, Bonnay's approach is only applicable to constants whose type level is $\leqslant 2$, because in his definition of the concept 'partial isomorphism' [27] between two structures $\langle M,\ a_1, \ldots, a_m\rangle$ and $\langle N,\ b_1, \ldots, b_n\rangle$, each a_i and b_i are at most of level 1. As is already pointed out, the concept of potential isomorphism-invariance can be generalised to any constant with any type. But I will not do the generalisation here; rather I will focus on the two theses mentioned above.

Now let me turn to set up a criterion of logicality for propositional connectives, which denote truth-functions. Suppose $\mathcal{S}$ is a propositional connective and $G_{\mathcal{S}}$ is the truth-function it denotes. According to Sher and Feferman, $\mathcal{S}$ is logical, if $G_{\mathcal{S}}$ is permutation-invariant on the domain $\{T, F\}$. To be precise,

> Truth-function Thesis
> $\mathcal{S}$ is logical if and only if
> For *any* identity function f on $\{T, F\}$ and any $x_1, \ldots, x_n \in \{T, F\}$, $(x_1, \ldots, x_n) \in G_{\mathcal{S}} = (f(x_1), \ldots, f(x_n)) \in G_{\mathcal{S}}$.

For constants with relational types, the thesis on logicality is that logicality is invariance under isomorphisms or homomorphisms. For propositional connectives, the thesis on logicality is that logicality is permutation-invariance. One might be wondering if there is any connection between these two theses. In other words, what is motivation for choosing the identity function rather than others in the latter thesis? There is, in fact, a nice connection.[28] Thanks to the connection, I do not distinguish between $logicality_{TS}$ and $logicality_F$ concerning propositional connectives. It stands to reason that all propositional connectives in (i.2) and (ii.2) are both *logical*$_{TS}$ and *logical*$_F$.

---

[27] See [6, pp. 181–183].

[28] There are different accounts for the connection, see [22, fn. 6]; [19]; [3, p. 11]. For any vocabulary $\mathcal{C}$ of the type $(\tau_1, \ldots, \tau_n)$ in $\mathcal{T}$, its denotation on domain M is the n-ary relation $R^M(\alpha_1, \ldots, \alpha_n) \subseteq M_{\tau_1} \times \ldots \times M_{\tau_n}$, where each n-th argument is a member of the set M_{τ_n}. Now following Bonnay, let us identify the denotation of $\mathcal{C}$ as the class of structures $f_{\mathcal{C}} = \{\langle M, \alpha_1, \ldots, \alpha_n\rangle : (\alpha_1, \ldots, \alpha_n) \in R^M$ and M is any domain$\}$. $\mathcal{C}$ is *logical*$_{TS}$(*logical*$_F$), if for any two isomorphic (or homomorphic) argument-structures $\langle M, a_1, \ldots, a_n\rangle$ and $\langle N, b_1, \ldots, b_n\rangle$, $\langle M, a_1, \ldots, a_n\rangle \in f_{\mathcal{C}}$ if and only if $\langle N, b_1, \ldots, b_n\rangle \in f_{\mathcal{C}}$. Let the True and the False be the sets $\emptyset$ and $\{\emptyset\}$, respectively. The propositional connective, say, '$\neg$', has as its denotation the class $f_{\neg} = \{\langle M, \emptyset, \{\emptyset\}\rangle, \langle M, \{\emptyset\}, \emptyset\rangle :$ M is any domain$\}$. This can be done for all propositional connectives. For any two argument-structures $\langle M, a_1, \ldots, a_n\rangle$ and $\langle M, b_1, \ldots, b_n\rangle$, where each a_i and b_i is the True or the False, whatever function $g : M \to N$ is, it is always the case that $f(a_i) = b_i$. So truth-values are not switched. Consequently, all propositional connectives are *logical*$_{TS}$ and *logical*$_F$.

5 Logicality of quantifiers

It can be easily shown that all quantifiers in (i.2) and (ii.2) are *logical*$_{TS}$. Now let us have a look at whether they are *logical*$_F$. Take universal quantifiers '$\forall$' and '$\forall^2_k$' as an example.

'$\forall$' is not homomorphism-invariant according to Definition*. The argument-structure of '$\forall$' on any domain M is $\langle M, A\rangle$, where $A \subseteq M$ (A is of type (e)). Consider the argument-structures $\langle M, A\rangle$ and $\langle N, B\rangle$, where $M = \{1, 2, 3\}$, $N = \{4, 5\}$, $A = \{1, 2\}$, $B = \{4, 5\}$, and $g : M \to N$ is such that $g(1) = 4$, $g(2) = g(3) = 5$. $\langle M, A\rangle$ and $\langle N, B\rangle$ are, then, homomorphic, because g is a surjection and $g(A) = B$. It is not hard to see that '$\forall$' is not invariance under $\langle M, A\rangle$ and $\langle N, B\rangle$, because $\{1, 2\} \notin f_\forall(M)$, i.e., $\{1, 2\} \neq M$, but obviously, $\{4, 5\} \in f_\forall(N)$, i.e., $\{4, 5\} = N$. However, Sher writes:

> (a) isomorphism-invariant operators that are also homomorphism-invariant.
>
> (ii) The existential and universal quantifiers. ([19], p. 335)

If Sher makes use of Definition*, she has to admit that '$\forall$' is not homomorphism invariance.

As is said, Sher's Definition* does not accord with Feferman's own. It can be proven easily that by Definition 6, '$\forall$' is *logical*$_F$. The structures $\langle M, A\rangle$ and $\langle N, B\rangle$ mentioned above, which are homomorphic in the sense of Definition*, are not homomorphic any more. For the function $g_{(e)}$ is not defined for the set $A = \{1, 2\}$, that is, it is not a member of the domain of the function. This is so, because $g(2) = g(3) = 5$, but $2 \in A$ and $3 \notin A$.

We have to decide whether or not the higher-order universal quantifiers in (i.2) and (ii.2) are *logical*$_F$. As Feferman proves in the example 3 of [8], the second-order universal quantifier '$\forall^2_k$' is not *logical*$_F$[29]. Let us give another, simple proof within the framework of relational type symbols. Let $M = \{1, 2, 3\}$ and $N = \{4, 5\}$, and $g : M \to N$ such that $g(3) = 5$, $g(1) = g(2) = 4$. Consider $\langle M, A\rangle$ and $\langle N, B\rangle$, which are two argument-structures of '$\forall^2_1$', where $A = \{\{1, 2, 3\}, \{1, 2\}, \{3\}, \emptyset\}$ and $B = \{\{4, 5\}, \{4\}, \{5\}, \emptyset\}$. Each member of A and B are of type (e). $\langle M, A\rangle$ and $\langle N, B\rangle$ are homomorphic under g.[30] However, $A \notin f_{\forall^2_1}(M)$, because $A \neq P(M)$, while $B \in f_{\forall^2_1}(N)$, because $B = P(N)$.

[29]Feferman only proves the special case in which $k = 1$. According to Feferman, the operation associated with the quantifier '$\forall^2_1$' on a domain M is the function $f_{\forall^2_1}$ of type $((e \to b) \to b) \to b$ such that $f_{\forall^2_1}(f) = [T, \text{if } \forall p \in M_{e\to b} f(p) = T, \text{else } F]$, where f is of type $(e \to b)$. Feferman works with functional type symbols. With relational type symbols, '$\forall^2_k$' is of type $(((e)))$, and the operation associated with it on a domain M is, correspondingly, the set $\{P(M)\}$.

[30]To see this, we have to show that A is in the domain of the function $g_{((e))}$. The domain of $g_{(e)}$ is $\{X \subseteq M : \forall x, y \in M[g(x) = g(y) \Rightarrow x \in X \leftrightarrow y \in X]\}$. Therefore, domain of the function $g_{((e))}$ is $\{Y \subseteq P(M) : \forall x, y \in Dom(g_e)[g_e(x) = g_e(y) \Rightarrow x \in Y \leftrightarrow y \in Y]\}$. Now it is easy to verify that $A \in Dom(g_{((e))})$. Similarly, it is easy to show that each member of A is in the domain of $g_{(e)}$. Moreover, $g_{((e))}(A) = B$.

What is the implication of the fact that the quantifier '$\forall_k^2$' is not *logical*$_F$? Feferman writes:

> In any case, I count it as an argument in favor of the homomorphism invariance condition for logicality that it excludes second-order, and thence higher-order, quantification, by example 3 of the preceding section. ([8])

Feferman would be definitely right if he says that the second-order universal quantifier, as well as universal quantifiers of even much higher-order, is not *logical*$_F$. But he says 'it excludes second-order, and thence higher-order'. I cannot see how (Feferman Thesis) can exclude the second-order existential quantifier '$\exists_k^2$'. Its denotation on any non-empty domain M is $f_{\exists_k^2}(M) = \{A : A \subseteq P(M^k), A \neq \emptyset\}$. Given any two homomorphic argument-structures $\langle M, A\rangle$ and $\langle N, B\rangle$, where A and B are of type $((e), \ldots, (e))$ (k occurrences of (e)), it cannot hold that $A \in f_{\exists_k^2}(M)$ and $B \notin f_{\exists_k^2}(N)$. Similarly, it is also impossible that $A \notin f_{\exists_k^2}(M)$ and $B \in f_{\exists_k^2}(N)$. For no surjection can map a non-empty set (an empty set) onto an empty set (a non-empty set).

It seems abnormal that the second-order quantifier '$\forall_k^2$' is not *logical*$_F$, while the second-order '$\exists_k^2$' is so, because they are usually definable one another.[31] E.g., '$\forall_k^2 X$' is defined as'$\neg\exists_k^2 X\neg$'. Intuitively, given that '$\neg$' and '$\exists_k^2$' are *logical*$_F$, so would be '$\forall_k^2$'. But as is already shown, $\forall_k^2$ is not *logical*$_F$. So the intuition might be unreliable. The first-order universal and existential quantifiers are both *logical*$_F$. This seems to verdict the observation that if we take one of them to be a primitive logical constant and define the other in a canonical way, then the defined one would also qualify as logical. This observation is proved by Feferman's theorem 6 in [8], according to which, any operation definable from the operations of first-order logic without identity is also definable from monodic homomorphism-invariant operations in terms of negation, conjunction and first-order existential quantifier, and vice versa. This theorem has nothing to do second-order quantifiers, however. So the aforementioned intuition does not get supported by it.

6 Logicality of the constants for the concept "The number of" and the membership relation

Let us look at if the constant in (i.3), i.e., '$\sharp$', and the constant in (ii.3), i.e., 'ϵ', are *logical*$_{TS}$ and *logical*$_F$. If a constant is not *logical*$_{TS}$, then it is not *logical*$_F$, because any homomorphism invariance is isomorphism invariance. I will show that '$\sharp$' is not *logical*$_{TS}$. For this purpose, it suffices to find two isomorphic argument-structures

[31] One should not be concerned with this abnormality too much, because there is a version of invariance under homomorphism which even bans operations definable in terms of first-order logical operations from being logical, see [5].

$\langle M, A, a\rangle$ and $\langle N, B, b\rangle$, where $A \in P(M)$, $a \in M$, $B \in P(N)$, $b \in N$, such that it is not the case that $(A, a) \in f_\sharp(M)$ if and only if $(B, b) \in f_\sharp(N)$.

The required isomorphic argument-structures are constructed as follows. Let $M = N = \mathbb{N}$, i.e., the set of natural numbers. Let $F_n = \{F : F \in P(\mathbb{N}), |F| = n\}$ for any $0 < n \in \mathbb{N}$. So F_n has $\aleph_0$ elements, where $0 < n \in \mathbb{N}$. If $n = 0$, let $F_0 = \emptyset$. In addition, let $G = \{F \subseteq \mathbb{N} : |F| = \aleph_0\}$. It is clear that $\bigcup_{n\in\mathbb{N}} F_n \cup G = P(\mathbb{N})$. Let $\boldsymbol{f} : P(\mathbb{N}) \to \mathbb{N}$ be the following function: for each $F \in F_n$ and $n \geqslant 0$, $\boldsymbol{f}(F) = n+1$ and for each $F \in G$, $\boldsymbol{f}(F) = 0$. This function satisfies the condition **(C)** on p. 67. Obviously, $f_\sharp(M) = f_\sharp(N) = \boldsymbol{f}$. Suppose $g : M \to N$ be such a function that $g(0) = 1$, $g(1) = 0$ and $g(x) = x$ for all $x > 1$. Obviously, g is a bijection. Let $\langle M, A, a\rangle = \langle \mathbb{N}, F_0, 1\rangle$ and $\langle N, B, b\rangle = \langle \mathbb{N}, F_0, 0\rangle$. Obviously, these two argument-structures are isomorphic. However, $(A, a) \in f_\sharp(M)$ and $(B, b) \notin f_\sharp(N)$, because $f_\sharp(M)(F_0) = 1$ and $f_\sharp(N)(F_0) = 1$. Therefore, '$\sharp$' is not *logical*$_{TS}$. So it is not *logical*$_F$, either.

I claim that 'ϵ' is *logical*$_{ST}$.[32] 'ϵ' is of type $(\tau, (\tau))$, where $\tau = (\ldots(e)\ldots)$ (n pairs of parenthesis, $n \geqslant 0$). I shall show that it is *logical*$_{ST}$ in the case that $n = 1$. It can be shown in the similar way that it is *logical*$_{ST}$ when $n > 1$. Suppose there are two isomorphic argument-structures of 'ϵ' $\langle M, a, A\rangle$ and $\langle N, b, B\rangle$, where $a \in M$, $A \subseteq M$, $b \in N$, $B \subseteq N$, such that $(a, A) \in f_\epsilon(M)$ but $(b, B) \notin f_\epsilon(N)$. Since the structures are isomorphic, there is a bijection $g : M \to N$ such that $g(a) = b$, and $g(A) = B$. Since $(a, A) \in f_\epsilon(M)$, $a \in A$. Thus, $g(a) = b \in g(A) = B$. Therefore, the supposition cannot hold.

The constant 'ϵ' is *logical*$_F$. Otherwise, suppose $\langle M, a, A\rangle$ and $\langle N, b, B\rangle$ are two homomorphic argument-structures of 'ϵ' such that it is not the case that $(a, A) \in f_\epsilon(M)$ iff $(b, B) \in f_\epsilon(N)$, where f is a surjection from M to N. Suppose $a \in A$ but $b \in B$. Obviously, this cannot be the case. Suppose $a \notin A$ but $b \in B$. Since $b \in B$ and $g_{(e)}(A) = B$, there is some $c \in A$ that $g(c) = b$. Let us show that $a \in A$. Because A is in the domain of $g_{(e)}$, it is the case that for all $x, y \in M$, if $g(x) = g(y)$, then $x \in A$ iff $y \in A$. $g(a) = g(c)$, so $b \in A$ iff $a \in A$. Therefore, $a \in A$. So we arrive at a contradiction. By the same token, we can show that the constant 'ϵ' is also *logical*$_F$ for $n > 1$.

7 Conclusion

Undoubtedly, the provability thesis of Frege-Russellian logicism is certainly unjustifiable. So let us focus on the definability thesis. Insofar as logicality is either

[32] Sher already points out the conclusion about 'ϵ' But she only considers the case in which 'ϵ' is of type $(e, (e))$. ([19], p. 305) Similarly, Tarski also mentions that it is logical and claims, further, that 'using the method of *Principia Mathematica*, set theory is simply part of logic.' ([22], p. 152) So Tarski claims that the definability thesis of Russell's logicism succeeds, although he does not show in his paper that the quantifiers and propositional connectives are also logical.

isomorphism invariance or homomorphism invariance, the constant '$\sharp$' for the concept " the number of " is not logical, while all constants in (i.2) are so. Consequently, Boolos's assessment of Frege's logicism on p. 70 is right. It is easy to show that there are arithmetical concepts that are not isomorphism invariance. Because any homomorphism invariance is also isomorphism invariance, any arithmetical concept that is not isomorphism invariance is not homomorphism invariance.

Let us take 0 as an example of non-isomorphism invariance in Frege's system. The constant 0 is defined as $\sharp[x : x \neq x]$ in Frege's system. 0 is of type e and its argument-structures are $\langle M, a\rangle$, where $a \in M$ and M is infinite. It holds in any arbitrary domain M that $f_\sharp([x : x \neq x]) = f_\sharp(\emptyset) \in M$. Suppose $f_\sharp(\emptyset) = a \in M$. Therefore, $f_0(M) = m$, i.e., the numeral 0 denotes the object a in M. Let N be such a set that $f_0(N) = b \in N$, where $|N| = |M|$. There is a bijection $f : M \to N$ such that $f(a) = c \in N$ and $c \neq b$. Therefore, $\langle M, a\rangle$ and $\langle N, c\rangle$ are isomorphic. However, $a = f_0(M)$ and $c \neq f_0(N)$. Therefore, 0 is not isomorphism invariance.

If logicality is isomorphism invariance, then the definability thesis of Russell's logicism is fully justified, because in the case the claim (+) on p. 68 holds for the language of STT.[33]

If logicality is homomorphism invariance, then the definability thesis of Russell's logicism cannot be fully justified. Higher-order universal quantifiers are not *logical*$_F$, but its corresponding existential quantifiers are *logical*$_F$. Let us replace the higher-order universal quantifiers occurring in the axioms of STT by the corresponding existential quantifiers. Then, all constants in (ii.2) and (ii.3) are *logical*$_F$.[34] However, the claim (+) does not hold any more provided that logicality is homomorphism invariance. To see this, let us take the cardinal numeral n^2 (its denotation is the number n of type $((e)))$ in STT as an example.

The argument-structure of n^2 is (M, A), where M is any non-empty set and $A \subseteq M$. $f_{n^2}(M) = \{A \subseteq M : |A| = n\}$ if $|M| \geqslant n$; otherwise,$f_{n^2}(M) = \emptyset$. Let $M = \{u_1, \ldots, u_n\}$, $N = \{v_1, \ldots, v_{n-1}\}$ and $f : M \to N$ be a surjection that $f(u_1) = f(u_2) = v_1$ and for $2 < m \leqslant n$, $f(u_m) = v_{m-1}$. It is not hard to verify that $\langle M, M\rangle$ 和 $\langle N, N\rangle$ are homomorphic, because $M \in Dom(f_{(e)})$. Obviously, $M \in f_{n^2}(M)$ and $N \notin f_{n^2}(N)$. Therefore, n^2 is not homomorphism invariance.

References

[1] T. Bays, 2001, "On Tarski on models", *Journal of Symbolic Logic*, **66(4)**: 1701–1726.

[2] P. Benacerraf and H. Putnam, 1983, *Philosophy of Mathematics: Selected Readings*, Cambridge: Cambridge University Press.

[33] For more information on the proof of the claim, see [21, pp. 384–393].

[34] It is worthy of noticing that the constant for identity is also not *logical*$_F$. But the constant is not necessary to higher-order logics, because it can be defined in a canonical manner.

[3] D. Bonnay, 2008, “Logicality and invariance”, *Bulletin of Symbolic Logic*, **14(1)**: 29–68.

[4] G. Boolos, 1998, *Logic, Logic and Logic*, Cambridge, MA: Harvard University Press.

[5] E. Casanovas, 2007, “Logical operations and invariance”, *Journal of Philosophical Logic*, **36(1)**: 33–60.

[6] H. Ebbinghaus, 1984, *Mathematical Logic*, Heidelberg: Springer.

[7] S. Feferman, 2004, “Tarski’s conception of logic”, *Annals of Pure and Applied Logic*, **126(1–3)**: 5–13.

[8] S. Feferman et al., 1999, “Logic, logics, and logicism”, *Notre Dame Journal of Formal Logic*, **40(1)**: 31–54.

[9] S. Feferman et al., 2010, “Set-theoretical invariance criteria for logicality”, *Notre Dame Journal of Formal Logic*, **51(1)**: 3–20.

[10] G. Frege, 1984, *Collected Papers on Mathematics, Logic, and Philosophy*, Oxford: Blackwell Publishers.

[11] G. Frege, 2013, *Gottlob Frege: Basic Laws of Arithmetic*, Oxford: Oxford University Press.

[12] W. S. Hatcher, 1982, *The Logical Foundations of Mathematics*, Oxford: Pergamon Press.

[13] R. G. Heck, 1996, “The consistency of predicative fragments of Frege’s *Grundgesetze der Arithmetik*”, *History and Philosophy of Logic*, **17(1–2)**: 209–220.

[14] H. T. Hodes, 1984, “Logicism and the ontological commitments of arithmetic”, *Journal of Philosophy*, **81(3)**: 123–149.

[15] J. Kontinen and J. Szymanik, 2011, “Characterizing definability of second-order generalized quantifiers”, in L. Beklemishev and R. de Queiroz (eds.), *Logic, Language, Information and Computation 18th International Workshop, WoLLIC 2011, Philadelphia, PA, USA. Proceedings*, pp. 187–200, Berlin/ Heidelberg: Springer.

[16] E. Mendelson, 2009, *Introduction to Mathematical Logic*, London: Chapman & Hall/ CRC.

[17] W. V. Quine, 1986, *Philosophy of Logic*, Cambridge, MA: Harvard University Press.

[18] G. Sher, 1991, *The Bounds of Logic*, Cambridge, MA/ London: Massachusetts Institute of Technology.

[19] G. Sher, 2008, “Tarski’s thesis”, in D. Patterson (ed.), *New Essays on Tarski and Philosophy*, pp. 300–339, Oxford: Oxford University Press.

[20] M. Stroinska and D. Hitchcock, 2002, “On the concept of following logically”, *History and Philosophy of Logic*, **23**: 155–196.

[21] A. Tarski, 1956, *Logic, Semantics, Metamathematics, Papers from 1923 to 1938*, Oxford: Clarendon Press.

[22] A. Tarski and J. Corcoran, 1986, “What are logical notions?”, *History and Philosophy of Logic*, **7(2)**: 143–154.

逻辑性和弗雷格算数与简单类型论的逻辑主义

石伟军

摘　　要

弗雷格和罗素的逻辑主义由两个部分构成：可证明性论题和可定义性论题。可以很确信地说，可证明性论题并不能得到完全的辩护。但是，为了向两者，特别是可定义性论题，提供辩护或者拒斥之，我们需要逻辑性的标准来决定，除了其它常元的逻辑性之外，表示数的常元和表示属于关系的常元的逻辑性。我将采用的逻辑性标准是塔尔斯基和谢尔提出的同构不变量标准和费弗曼提出的同态不变量标准。塔尔斯基和谢尔在不同的地方已经指出罗素的表示属于关系的常元是同构不变量。在本文中，我将证明如下结论：第一，表示属于关系的常元是同态不变量；第二，弗雷格的表示数的常元既不是同构不变量也不是同态不变量；第三，如果逻辑性是同构不变量或者同态不变量，弗雷格的逻辑主义（弗雷格算数）的可定义性论题不成立；第四，如果逻辑性是同构不变量，罗素的逻辑主义（简单类型论）的可定义论题成立，但若逻辑性是同态不变量，这个论题则不成立。

石伟军　中国人民大学哲学系
柏林洪堡大学哲学系
2015000867@ruc.edu.cn