Arbitrariness and Frege Arithmetic

Ludovica Conti
University of Vienna
Wien, Österreich
ludovica.conti@univie.ac.at

Abstract

This paper focuses on a less-known version of Abstractionism, that we'll call Arbitrary Frege Arithmetic. This system aims to restore Fregean Logicist project (of a logical derivation of second-order Peano Arithmetic) by adopting, as much as possible, its original tools, namely a logical second-order system augmented with a (consistent) version of Basic Law V. The major differences from Frege's Grundgesetze will be a recasted presentation of the abstractionist vocabulary and a weakening of first-order classical logic in a negative free logic one. Both these choices are supported by philosophical and formal motivations that allow us to consider them completely compatible with the abstractionist spirit.

Keywords: Abstraction principles, Logicism, Free logic, Arbitrariness, Frege Arithmetic.

1 Introduction

Abstractionist theories in philosophy of mathematics are systems composed by a logical theory augmented with an abstraction principle (AP), of form: $\forall X \forall Y(@X = @Y) \leftrightarrow E(X, Y)$[1] – that introduces, namely rules and implicitly defines, a term-forming operator @ by means of a formula satisfying the conditions of an equivalence relation E. As is well-known, the seminal abstractionist program, Frege's Logicism, failed[2]: Russell's Paradox proved its inconsistency and, *a fortiori*, its non-logicality. In the last century, both the issue of consistency and the issue of logicality have been resumed in the abstractionist debate (cf. Tarski 1956, Fine 2002, Antonelli 2010, Cook 2016, Boccuni and Woods 2020). More precisely, on the one side, different revisions of Frege's original system have been proposed in order to avoid Russell's Paradox and to obtain a consistent system that is strong enough to derive (at least, a relevant portion of) Peano Arithmetic. On the other side, given a semantical definition of logicality as permutation invariance, some abstraction principles have been proved to be logical (Antonelli 2010, Cook 2016).

Nevertheless, many concerns are still open. Particularly, regarding the preliminary condition of consistency, the ways out of Russell's Paradox proposed so far do not precisely mirror a corresponding explanation of the origin of the contradiction and often imply a weakening of the hoped strength of the theory (cf. Heck 1996, Wehmeier 1999, Ferreira and Wehmeier 2002)[3]; regarding the issue of logicality, an undesired dilemma overshadows the abovementioned results: precisely in case of logical (i.e. permutation invariant) abstraction principles, their implicit *definienda* turn out to be non logical (Antonelli 2010) – so preventing a full achievement of the Logicist goal.

[1] In the rest of the paper, I'll adopt this axiomatic version of AP. Given full Comprehension Axiom Schema (that will be assumed in the systems that we'll investigate), it is provably equivalent to the schematic form: $@x.\alpha(x) = @x.\beta(x) \leftrightarrow E(\alpha(x), \beta(x))$. Cf. Linnebo 2016

[2] It was proposed with the foundational purpose to derive arithmetical laws as logical theorems and to define arithmetical expressions by logical terms.

[3] In Ferreira 2018 and Boccuni 2022, second-order Peano Axioms are recovered but by appealing to stronger logical resources – i.e. double-sorted variables

My preliminary hypothesis is that these – apparently unrelated –problems have a common source in some unquestioned assumptions of Frege's project (inherited also by the following abstractionist programs). I argue that such assumptions are part of what we can call the traditional view of the abstraction, that includes the choice of classical logic as the base theory, with the related semantical consequence of full referentiality of the vocabulary, and the choice of a so-called canonical interpretation function for all the (both primitive and defined) expressions of the language.

In the rest of the paper, I show that by renouncing one or both of these problematic assumptions we can recover consistency and/or logicality. More precisely, I propose a double revision of Frege's Logicist program: on the one side, weakening the canonical interpretation function for the implicitly defined (abstract) expressions of the vocabulary (cf. Boccuni and Woods 2020), I prove that any consistent revision of BLV turns out to be logical (i.e. permutation invariant); on the other side, I show that such an arbitrary interpretation, on a (negative) free logic background, allows us to identify a restriction of BLV, able to precisely exclude the paradoxical concepts, namely to avoid Russell's Paradox, but, at the same time, to preserve the derivational strength necessary to derive second-order Peano axioms. This means that this system, precisely renouncing to the Traditional assumptions mentioned above, is able to recover both Frege's goals of logicality and consistency.

2 Non-canonical Abstraction

By a non-canonical account of the abstraction I mean the result of renouncing the meta-semantic assumption of a uniform interpretation function for all the expressions of a same syntactic category. Such an account emphasises the semantic difference between undefined expressions (such as non-logical constants) and defined ones, whose meaning should be fixed by the theory itself. Accordingly, in a non-canonical perspective, the semantic indeterminacy of the abstractionist vocabulary (as exhibited by mixed identity statements) is not a problem but a feature to be accounted for. A formulation of this attitude has been characterised by Antonelli (**Antonelli:2010**)[4] as a deflationary reading of the abstraction. We can summarise this approach as a negative thesis: abstractionist theories are categorical, and abstraction principles do not provide sufficient information to identify an intended model – understood, in the Neologicist framework, as the model in which abstract objects are identified as the singular and determined interpretation of the abstract terms and are provably distinct from all other non-abstract objects. On the contrary, abstraction principles are silent about the particular function that their operator selects – among all those that are able to map equivalent concepts in a same object and non-equivalent concepts in different objects; accordingly, abstraction functions only index classes of their arguments by objects of the codomain – they impose a lower bound on the cardinality of the domain, but are neutral with respect to the identity of their values; thus, abstract terms should only denote possible indexes of equivalence classes – indifferent to their specific nature.

As a negative thesis, deflationism is still compatible with different notions of indeterminacy and different implementations of it. In what follows, I suggest that a fruitful way to spell out such an indeterminacy is through the notion of arbitrary reference, which in turn can be further spelled out through "quantificational" and "parametrical" meanings (according to their prototypes of arbitrary expressions). In the "quantificational" approach (cf. Woods 2014), classical interpretation function from sub-sentential expressions to their singular denotations must be substituted by a generalised interpretation, (i.e. a function which sends some domains

[4]"We characterize this view of abstraction as deflationary in that the main role it ascribes to abstraction is to provide such a lower bound, while denying the objects delivered by abstraction any special status.

D to a set of objects of the same type in the type-hierarchy over D) that assigns a whole range of candidate ones (made available in the different models). In this perspective, abstraction principles are able to single out a whole class of (compatible) knowable and standard functions as a range of candidate denotations of the abstractionist operator. Accordingly, the abstract terms exhibit a "quantificational behaviour", similar to that of the variables, i.e. are plurally referring expressions, "ranging over" a class of admissible elements (functions/objects) that could serve as denotation. Correspondingly, the notions of truth and logical consequence need to be rephrased by renouncing the classical valuation function (bivalence) and split in two branches: the sub-truth evaluated on each singular model (namely, relative to any possible precisification of the interpretation function) and the super-truth evaluated by quantifying over all the candidate models and relative to all the possible precisifications of the interpretation function.

On the contrary, in the "parametrical" approach (cf. Andreas and Schiemer 2016), the classical interpretation function from sub-sentential expressions to their singular denotations must be substituted by a choice interpretation, that picks up an indeterminate item into the whole range of candidate ones (made available in the different models). In this perspective, arbitrary terms are singularly referring expressions, denoting particular, though indeterminate, elements (functions/objects). Correspondingly, the notions of truth and logical consequence need to be rephrased by renouncing the classical valuation function (bivalence) and split in two branches: the local truth evaluated respect to the existence of at least a possible choice interpretation function (namely, relative to any possible value) and the general truth evaluated by quantifying over all the possible choice interpretation functions, namely relative to all the candidate models in which they insist[5].

Besides successfully handling semantical indeterminacy (thereby dissolving the cross-sortal identity problem) and inheriting the above mentioned advantages of deflationism in general (i.e. full compatibility with abstractionist results and enhancement of the language priority thesis), an arbitrary interpretation of the abstractionist vocabulary exhibits further pleasant aspects (at least in a Fregean and Neologicist perspective), such as a more pronounced analyticity, a new kind of non-arrogance and an epistemically safe impredicativity. In the next sections, I aim to emphasise that, given an arbitrary interpretation – or, more precisely, a corresponding paraphrase – of the abstractionist vocabulary, the criterion of permutation invariance (and then the desirable feature of logicality) is satisfied by any abstraction operator defined by consistent revisions of BLV; on the contrary, the extensional operator defined by the original version of BLV fails precisely for the arguments that determine its inconsistency. Accordingly, an alternative, consistent and logical, extensional operator will be (partially) defined by means of a corresponding restriction of the axiom.

3 Logicality

Arbitrary interpretation of the abstractionist vocabulary allows us to provide an arbitrary rephrase of the criterion of Weak Objectual Invariance: an abstraction operator $@_R$ – understood as denoting any of the range of its candidate denotation – is weakly objectually invariant if and only if, for any model $M =< \Delta, I >$ and permutation π of Δ, $\pi(@) = \{< \pi(X), \pi(y) > \mid < X, y > \in @\} = @$, i.e. $\forall X, \forall y\ @(X) = y$ if and only if $@(\pi[X]) = \pi[y]$[6].

[5] To be more precise, there are at least two (equivalent) ways to model this choice-functional meaning of arbitrariness. On the one side, we can build choice functions on each (fully determined model) – cf. Andreas and Schiemer 2016; on the other side, we could build choice functions as part of the different models themselves.

[6] This criterion has been further strengthened (cf. Boccuni and Woods 2020) as Weak Objectual Invariance under isomorphism: an abstraction operator $@_R$ is weakly objectually invariant under isomorphism if and only if, for any model $M =< \Delta, I >$ and bijection i of $\Delta \to \Delta'$, $i(@^{\Delta}) = \{< i(X), i(y) > \mid < X, y > \in @\} = @^{\Delta'}$, i.e. $\forall X, \forall y\ @(X) = y$ if and only if $@(i(X)) = i(y)$.

By extending this result about the cardinal operator, we can prove that many consistent revisions of Fregean BLV implicitly define extensional operators that are weakly objectually invariant. The most obvious example is constituted by what we can call Boolosean restrictions of BLVb, cf. Boolos 1986, Shapiro 2003, (BLV^B): $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow (\phi(X) \wedge / \vee \phi(Y) \rightarrow \forall x(Xx \leftrightarrow Yx)))$. As is well known, principles instantiating this schema rule abstraction functions that perform as an injective function when they take arguments satisfying the condition ϕ but are free to map the other concepts in the same object. No matter the specific condition ϕ and the unruled behaviour of the function for the arguments that do not satisfy it, the function (arbitrarily interpreted) is weakly objectual invariant (under permutation and also under isomorphism).

Theorem 3.1. *The extensional operator ϵ^* defined by BLV^B is weakly objectually invariant.*[7]

A fortiori ϵ^*, arbitrarily interpreted, is weakly objectually invariant under permutation. Intuitively, given any permutation π of Δ, every concept, both satisfying and not satisfying ϕ, is co-extensional only with itself. Then, the partition of the second-order domain – both of the portion that constitutes the domain of the injection and of its complement – is invariant under permutation (and isomorphism). Given an arbitrary interpretation of the abstractionist vocabulary, permutation does not affect the set of candidate denotations of the abstract terms. Thus $\pi(\epsilon^*) = \epsilon^*$.

A similar reasoning can be rephrased for any consistent version of BLV, obtained both by weakening the logical background (e.g., in case of predicative subsystems of *Grundgesetze*, cf. Heck 1996, Wehmeier 1999, Ferreira and Wehmeier 2002) or by different restrictions of the principle itself (e.g. restrictions of BLVa on a free logic background (BLV^F): $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow (\phi(X) \wedge / \vee \phi(Y) \wedge \forall x(Xx \leftrightarrow Yx))))$.

This kind of results extends the solution of the logicality dilemma from Neologicism to all the consistent revisions of Frege's program. Not only the cardinal operator, but also the extensional operator, implicitly defined by means of the logical (weakly invariant) relation of co-extensionality, is logical (weakly objectually invariant) as well.

4 Consistency

The logical part of the language of Arbitrary Logicism, L_F, includes denumerably many first-order variables $(x, y, z, ...)$, denumerably many second-order variables $(X, Y, Z, ...)$, logical connectives $(\neg, \rightarrow)$ and a first-order existential quantifier $(\exists)$[8]. We can also usefully define a predicative monadic constant (E!), whose extension is equal to the range of identity: E!a $=_{def} \exists x(x = a)$. The only non-logical primitive symbol is the term-forming operator ϵ which applies to monadic second-order variables to produce complex singular terms $(\epsilon(X))$[9].

[7] Given D, D' and an isomorphism $\zeta : D \rightarrow D'$, we have to prove that $\zeta(\epsilon^{*D}) = \epsilon^{*D'}$. (Given an arbitrary interpretation, ϵ^{*D} and $\epsilon^{*D'}$ respectively denote the set of candidate canonical denotations, namely the set of abstraction functions, respectively, $f : \wp(D) \rightarrow D$ and $f' : \wp(D') \rightarrow D'$, that satisfy BLV^B).
$\forall f \in \epsilon^{*D}, \zeta(f) = \zeta \circ f \circ \zeta^{-1} = f' : \wp(D') \rightarrow D'$ satisfying BLV^B, i.e. $f' \in \epsilon^{*D'}$. Then, $\zeta(\epsilon^{*D}) \subseteq \epsilon^{*D'}$.
$\forall f' \in \epsilon^{*D'}, \zeta^{-1} \circ f' \circ \zeta = f : \wp(D) \rightarrow D$ satisfying BLV^B, i.e. $f \in \epsilon^{*D}$ such that $\zeta(f) = f' \in \zeta(\epsilon^{*D'})$. Then, $\epsilon^{*D'} \subseteq \zeta(\epsilon^{*D})$.

[8] We can also define the other connectives and the universal quantifier $\forall x Ax =_{def} \neg\exists x \neg Ax$.

[9] Let D be the full first-order domain (then, the second-order domain is constituted by its power-set $\wp(D)$). The satisfaction clauses for the formulas of L_F are defined in terms of an evaluation function V and an assignment function I that ascribes elements of D to the first-order terms and elements of $\wp(D)$ to the second-order terms:

- $V(Pt_1, ..., t_n) = 1 \leftrightarrow I(t_1), ..., I(t_n) \in D \wedge < I(t_1), ..., I(t_n) > \in I(P)$; 0 otherwise;
- $V((s) = (t)) = 1 \leftrightarrow I(s), I(t) \in D \wedge I(s) = I(t)$; 0 otherwise;
- $V(E!t) = 1 \leftrightarrow I(t) \in D$; 0 otherwise;
- $V(\neg\alpha) = 1 \leftrightarrow V(\alpha) = 0$; 0 otherwise;

The theory involves, as its logical part, the axioms and inference rules of non-inclusive negative free logic with identity ($NF^{=}$):

NF1) $\forall v\alpha \rightarrow (E!t \rightarrow \alpha(t/v))$;

NF2) $\exists vE!v$;

NF3) $s = t \rightarrow (\alpha \rightarrow \alpha(t//s))$[10];

NF4) $\forall v(v = v)$;

NF5) $P\tau_1, ..., \tau_n \rightarrow E!\tau_i$ (with $1 \leq i \leq n$);

$\forall$I): $E!a...\phi(a/x) \vdash \forall x\phi$;

$\exists$E): $\phi(a/x), E!a...\psi, \exists x\phi \vdash \psi$, where a is a new individual constant which does not occur in ϕ and ψ.

Additionally, the theory involves an axiom-schema of universal instantiation for second-order variables ($\forall X\phi(X) \rightarrow \phi(Y)$), a rule of universal generalisation (GEN), a second-order comprehension axiom schema (CA: $\exists X\forall x(Xx \leftrightarrow \alpha)$) and *modus ponens* (MP)[11].

The abstraction principle that characterizes this theory is obtained by weakening the right-to-left conditional of Basic Law V (BLV: $\forall F\forall G(\epsilon F = \epsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx))$, i.e. BLVa (arbitrarily interpreted), by means of the condition of Permutation Invariance (cf. Antonelli 2010, Boccuni and Woods 2020).

W-BLV: $\forall F\forall G(\epsilon F = \epsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx) \wedge \epsilon(\pi(F)) = \pi(\epsilon F))$[12]

As is well known, the ϵ operator (as defined by standard BLV), also arbitrarily interpreted, is not Permutation Invariant – because, roughly speaking, by being inconsistent it is unable to define or rule any function. We can emphasize that, given an arbitrary interpretation, Permutation Invariance fails precisely for the argument that determines its inconsistency. In other words, as can be pointed out for other consistent revisions of BLV, in any case in which it is safely restricted, ϵ turns out to satisfy Permutation Invariance, namely it is such that $\pi(\epsilon) = \epsilon$, i.e. $\forall X\forall y(\epsilon X = y \leftrightarrow \epsilon(\pi(X)) = \pi(y))$. Then, the second conjunct of the right-hand side of W-BLV requires that – no matter which object y is identical to ϵF – ϵ satisfies Permutation Invariance for the considered arguments[13].

- $V(\alpha \wedge \beta) = 1 \leftrightarrow \alpha = 1 \wedge \beta = 1$; 0 otherwise;
- $V(\alpha \vee \beta) = 1 \leftrightarrow \alpha = 1 \vee \beta = 1$; 0 otherwise;
- $V(\forall v\alpha) = 1 \leftrightarrow \forall s \in D, V_{(t,s)}(\alpha(t/v)) = 1$ – where t is not in α and $V_{(t,s)}$ is the valuation function on the model $< D, I^* >$ such that $I^* = I$, except that $I^*(t) = s$.
- $V(\forall V\alpha) = 1 \leftrightarrow \forall S \subseteq D, V_{(T,S)}(\alpha(T/V)) = 1$ – where T is not in α and $V_{(T,S)}$ is the valuation function on the model $< D, I^* >$ such that $I^* = I$, except that $I^*(T) = S$.

[10]Where $\alpha(t//s))$ is the result of replacing one or more occurrences of s in A by t.

[11]From these axioms we can also derive the following theorems: T1) $\forall xE!x$; T2) $t = t \leftrightarrow E!t$; T3) $(\neg E!s \wedge \neg E!t) \rightarrow (\alpha \rightarrow \alpha(t//s))$.

[12]This abstraction principle is clearly circular because the extensional operator occurs on both the sides of the biconditional. The idea that circularity defeats the definitional role of such principle (or, in general, of implicit definitions) is controversial. Anyway, in this framework, what we need is a principle that rules the behavior of a new symbol of the language and W-BLV carries out this task.

[13]This revision of BLV (particularly of BLVa) is featured by a restriction that, with respect to many other (syntactical ones), is expressible into the language. Indeed, the permutation π of the operator or of the concepts mentioned in the right-hand side of the bi-conditional can be defined as abbreviation of the effects of any first-order bi-jective function f: $D_1 \rightarrow D_1$ on the entities (sets, relations or functions) further up in the type hierarchy.

Accordingly, W-BLV, as a bi-conditional, turns out to be satisfied by any concept instantiating the universal quantifier. On the one side, given an arbitrary interpretation of the abstraction operator, for any concept different from Russellian concept (R), $\pi(\epsilon) = \epsilon$. On the other side, we can consider Russell's Paradox as a *reductio ad absurdum* of the alleged truth of both the sides of the bi-conditional for the concept R: the contradiction proves that ϵR – as legitimately admitted on a free logical background – does not exist, namely it is a term devoid of denotation; accordingly, it is not identical to itself (so, falsifying the left-hand side of W-BLV) and, even if R, as any other concept, is co-extensional with itself, it falsifies Permutation Invariance of the operator[14]. Accordingly, also the right-hand side of W-BLV is false and also the instance of the bi-conditional for the concept R is verified.

Such a restricted version of W-BLV allows us to derive a corresponding restricted version of Hume's Principle (W-HP). However, as already proved (Boccuni and Woods 2020), cardinal operator (in our choice-functional translation or in an arbitrary interpretation) satisfies permutation invariance without restrictions. For this reason, even if we are only able to prove W-HP, it is equivalent to full HP, so it allows us to derive the main arithmetical results, including Frege's Theorem.

References

Andreas, Holger and Georg Schiemer (2016). "A choice-semantical approach to theoretical truth". In: *Studies in History and Philosophy of Science Part A* 58, pp. 1–8.

Antonelli, G Aldo (2010). "Notions of invariance for abstraction principles". In: *Philosophia Mathematica* 18.3, pp. 276–292.

Boccuni, Francesca (2022). "Plural Frege Arithmetic". In: *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences* 26-1, pp. 189–206.

Boccuni, Francesca and Jack Woods (2020). "Structuralist neologicism". In: *Philosophia Mathematica* 28.3, pp. 296–316.

Boolos, George (1986). "Saving Frege from contradiction". In: *Proceedings of the Aristotelian Society*. Vol. 87. JSTOR, pp. 137–151.

Cook, Roy T (2016). "Abstraction and four kinds of invariance (or: What's so logical about counting)". In: *Philosophia Mathematica* 25.1, pp. 3–25.

Ferreira, Fernando (2018). "Zigzag and Fregean arithmetic". In: *The philosophers and mathematics: Festschrift for Roshdi Rashed*, pp. 81–100.

Ferreira, Fernando and Kai F Wehmeier (2002). "On the consistency of the Δ 1 1-CA fragment of Frege's Grundgesetze". In: *Journal of Philosophical Logic* 31, pp. 301–311.

Fine, Kit (2002). *The limits of abstraction*. Clarendon Press.

Heck, Richard G (1996). "The consistency of predicative fragments of Frege's Grundgesetze der Arithmetik". In: *History and Philosophy of Logic* 17.1-2, pp. 209–220.

Linnebo, Øystein (2016). "Impredicativity in the neo-Fregean program". In: *[Ebert and Rossberg, 2016]*, pp. 247–268.

Shapiro, Stewart (2003). "Prolegomenon To Any Future Neo-Logicist Set Theory: Abstraction And Indefinite Extensibility". In: *The British journal for the philosophy of science* 54.1, pp. 59–91.

Tarski, Alfred (1956). "The concept of truth in formalized languages". In: *[Tarski, 1983]*, pp. 152–278.

[14]This last claim follows from the definition of π and the result of non-existence of ϵR: on the one side, $\epsilon(\pi(R)) = \epsilon(\{\pi(x)|x \in R\}) = \epsilon(X)$ – where X is any other concept (based on π); on the other side, $\pi(\epsilon R)$, given that ϵR is not denoting, is another well-formed term without denotation; then, the identity between ϵX (for any X that is obtained by means of a permutation of R) and the empty term $\pi(\epsilon R)$ is false.

Wehmeier, Kai F (1999). “Consistent fragments of Grundgesetze and the existence of non-logical objects”. In: *Synthese* 121, pp. 309–328.
Woods, Jack (2014). “Logical indefinites”. In: *Logique et Analyse* 227, pp. 277–307.