Arithmetic, Abstraction, and the Frege Quantifier

G. Aldo Antonelli

University of California, Davis

2011 ASL Meeting
Berkeley, CA

Logicism and abstraction

Goal of the talk: to present a formalization of first-order arithmetic characterized by the following:

1. Natural numbers are identified with *abstracta* of the equinumerosity relation;
2. Abstraction itself receives a *deflationary* construal — abstracts have no special ontological status.
3. Logicism is articulated in a *non-reductionist* fashion: rather than reducing arithmetic to principles whose logical character is questionable, we take seriously Frege's idea that cardinality *is* a logical notion.
4. The formalization uses two main technical tools:
   - A *first-order* (binary) cardinality quantifier F expressing "For every A there is a (distinct) B's";
   - An abstraction operator Num assigning first-level objects to predicates.
5. The logicist banner is then carried by the quantifier, rather than by Hume's Principle.
6. Finally, the primary target of the formalization are the *cardinal* properties of the natural numbers, rather than the *structural* ones.

The main formalization strategies for first-order arithmetic:

- The *Peano-Dedekind approach*: numbers are primitive, their properties given by the usual axioms;
- The *Frege-Russell tradition*: natural numbers are identified with equinumerosity classes;
- The *Zermelo-von Neumann implementation*: natural numbers are identified with particular representatives of those equivalence classes, e.g.

$$\emptyset, \quad \{\emptyset\}, \quad \{\{\emptyset\}\}, \quad \{\{\{\emptyset\}\}\}, \ldots$$

or

$$\emptyset, \quad \{\emptyset\}, \quad \{\emptyset, \{\emptyset\}\}, \quad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

Note: Numbers are not always *members* of the equivalence classes they represent — e.g., the Zermelo numerals.

None of these are completely satisfactory:

- The Dedekind-Peano approach completely ignores the *cardinal* properties of numbers while only focusing on the *structural* ones.
- The Frege-Russell tradition is more general, correctly derives *structural* properties from *cardinal* ones, but it is higher-order.
- The Zermelo-von Neumann implementation can be carried out at the first-order but at the price of identifying the natural numbers with a particular kind of entities (Benacerraf problem). Cardinal properties are derived from structural ones, and then only thanks to embedding of $\mathbb{N}$ into a rich set-theoretic universe.

Note: In keeping with Benacerraf, on the present view of abstraction the issue of the "ultimate nature" of numbers is a pseudo-problem.

FREGE'S THEOREM

Peano Arithmetic is interpretable in second-order logic (including second-order comprehension) augmented by "Hume's Principle."

Hume's Principle (HP) asserts that:

$$\mathsf{Num}(F) = \mathsf{Num}(G) \Longleftrightarrow F \approx G,$$

where Num is an abstraction operator mapping second-order variables into objects, and $F \approx G$ abbreviates the second-order claim that there is a bijection between F and G.

The neo-logicists hail this result as a realization of Frege's program, based on the claimed *privileged status* of HP.

But not only is such a status debatable (more later), but the second-order nature of logical framework makes it intractable.

Abstraction Principles

The notion of a "classifier" is known from descriptive set theory:

Definition

If R is an equivalence relation over a set X, a *classifier* for R is a function $f : X \to Y$ such that $f(x) = f(y) \iff R(x,y)$.

An *abstraction operator* is a classifier f for the specific case in which $X = \mathscr{P}(Y)$, i.e., an assignment of first-order objects to "concepts" (predicates, subsets of the first-order domain), which is governed by the given equivalence relation.

An *abstraction principle* is a statement to the effect that the operator f assigns objects to concepts according to the given equivalence R:

$$\mathsf{Ab}_R : \qquad f(X) = f(Y) \iff R(X,Y).$$

Abstraction principles are often characterized as the preferred vehicle for the delivery of a special kind of objects — so-called *abstract entities* — whose somewhat mysterious nature includes such properties as non-spatio-temporal existence and causal inefficacy.

Digression: Hilbert's ε-calculus

The ε-calculus comprises the two principles:

(1) $\phi(x) \rightarrow \phi(\varepsilon x.\phi(x))$

(2) $\forall x(\phi(x) \leftrightarrow \psi(x)) \rightarrow \varepsilon x.\phi(x) = \varepsilon x.\psi(x)$

Addition of:

(3) $\varepsilon x.\phi(x) = \varepsilon x.\psi(x) \rightarrow \forall x(\phi(x) \leftrightarrow \psi(x))$.

would give an *abstraction principle* witnessed by a *choice function*. Can the above be consistently added to the ε-calculus?

- Principle (3) has no finite models: in a finite domain there is no injection of the (definable) concepts into the objects.
- Principles (1) and (3) are inconsistent: there is no injective choice function on the power-set of a set of size > 1.
- Principles (2) and (3) give the first-order fragment of Frege's *Grundgesetze* and are therefore consistent (T. Parsons).

The limitations of logicism

Contemporary neo-logicists pursue a *reductionist* version of logicism: arithmetic is reducible to a principle (HP) enjoying a logically privileged status.

But this version is subject to several objections:

1. The *Bad Company* objection: HP looks very much like other inconsistent principles (Boolos, Heck);
2. The *Embarassment of Riches* objection: there are pairwise inconsistent principles, each one of which is individually consistent (Weir);
3. The *Logical Invariance* objection: depending on how exactly one formulates invariance, HP might not be invariant under permutations, which is (at least) a necessary condition for logicality (Tarski, Feferman, McGee, Sher, Bonnay).

The first two are well known, so we focus on the last one.

Invariance under permutation was first identified by Tarski as a criterion demarcating logical notions, on the idea that such notions are independent of the subject matter.

A predicate P is invariant iff $\pi[P] = P$ for every permutation π, where $\pi[P]$ is the point-wise image of P under π.

The following are all invariant:

- One-place predicates: $\varnothing$, D;
- Two place predicates: $\varnothing$, D^2, $=$, $\neq$;
- Predicates *definable* (in FOL, infinitary logic, etc.) from invariant predicates.

(And conversely, invariant notions are all definable in a possibly higher-order or infinitary language [McGee]).

Notions of invariance are available for entities further up the type hierarchy, e.g., *quantifiers*. But there is no accepted notion of invariance for abstraction principles.

Notions of Invariance for Abstraction

There are, *prima facie*, three different ways in which invariance can be applied to abstraction. Let R be an equivalence relation on a domain D and $f : \mathscr{P}(D) \to D$ the corresponding operator. These notions are:

- Invariance of the equivalence *relation* R;
- Invariance of the *operator* f;
- Invariance of the abstraction *principle* Ab_R: $f(X) = f(Y)$ iff $R(X, Y)$.

More formally:

Definition

- R is *simply invariant* iff $R(X, \pi[X])$ holds for any permutation π.
- f is *objectually invariant* if it is invariant as a set-theoretic entity: i.e., if and only if $\pi[f] = f$ for any π.
- Ab_R is *contextually invariant* iff, for any operator f and permutation π, $\pi[f]$ satisfies the principle whenever f does.

Is Abstraction Logically Invariant?

Proposition

No function f satisfying HP is objectually invariant.

In fact, the above can be generalized:

Theorem

Let f be an abstraction operator and suppose that $|D| > 1$ and suppose R is simply invariant. Then f is not objectually invariant.

Remark

Simple invariance is the strongest notion of invariance for R, and a very plausible necessary condition on R, but is not germane to the invariance of *abstraction*. The equinumerosity relation $\approx$ is simply invariant.

Objectual invariance is the notion that speaks to the character of abstraction as a *logical operation*. We see that objectual invariance is quite rare and mostly incompatible with simple invariance.

REMARK

If every f_R satisfying an abstraction principle is objectually invariant, then the principle itself is contextually invariant.

PROPOSITION

If R is simply invariant then the corresponding abstraction principle is contextually invariant.

COROLLARY

HP is contextually invariant (since $\approx$ is simply invariant), but contextual invariance does not amount to much, implied by a notion of invariance of R even weaker than simple invariance.

M. Dummett pointed out that in the true spirit of logicism:

"Cardinality is *already* a logical notion"

and does not need a *reduction* to a more fundamental principle to show that it is.

And in fact, cardinality is *permutation invariant*: the cardinality of a set does not change under permutations of the domain.

Reductionist versions of logicism and neo-logicism would seem to suffer from an equivocation: it is the notion of *cardinality*, not that of *number*, that has some claim to being logically innocent.

We want to pursue this more general construal of logicism by building cardinality right into the language in the form of a *primitive quantifier*.

Detour: the modern view of quantifiers

Following Frege in §21 of *Grundgesetze der Arithmetik*, Montague and Mostowski defined a quantifier Q over a domain D as a collection of *predicates*. For instance:

- $\forall = \{D\}$;
- $\exists!^k = \{X \subseteq D : |X| = k\}$;
- $\mathsf{Some} = \{(A, B) : A \cap B \neq \emptyset\}$;
- $\mathsf{Most} = \{(A, B) : |A \cap B| > |A - B|\}$.

Quantifiers are distinguished by the *number* as well as the *dimensions* of their arguments, i.e., the number of formulas appearing as arguments as well as the number of *their* variables.

The notion of *Permutation invariance* applies to quantifiers: $\mathsf{Q}(A, B)$ holds iff $\mathsf{Q}(\pi[A], \pi[B])$ holds.

Whereas *first-order* quantifiers are collections of relations over D, *second-order* quantifiers are relations over first-order quantifiers. ***The distinction between first- and second-order is semantical, not merely notational.***

The Frege quantifier F

Semantic Definition

The Frege quantifier F has the semantics: $\mathsf{F}(A, B) \iff |A| \leq |B|$.

Similarly for the *polyadic* version of F: for $R, S \subseteq D^n$: $\mathsf{F}(R, S) \iff |R| \leq |S|$.

Syntactic Definition

The *first-order* language $\mathscr{L}_{\mathsf{F}}$ comprises formulas built up from individual, predicate, and function constants by means of connectives and the quantifier $\mathsf{F}x$ satisfying the clause:

> *if $\phi(x)$, $\psi(x)$ are formulas in x, then $\mathsf{F}x(\phi(x), \psi(x))$ is a formula;*

and similarly for the polyadic version.

In either its monadic or polyadic variants, F is a *binary first-order* quantifier (just like Most).

Equinumerosity

The equinumerosity (Härtig's) quantifier $\mathsf{I}x(\phi(x), \psi(x))$ is definable as the conjunction $\mathsf{F}x(\phi(x), \psi(x)) \wedge \mathsf{F}x(\psi(x), \phi(x))$.

The definition is correct by the Schröder-Bernstein theorem.

A model $\mathfrak{M}$ with non-empty domain D provides an interpretation for the non-logical constants of $\mathscr{L}_\mathsf{F}$.

In the monadic case, given a formula $\varphi(x)$ and assignment s, satisfaction $\mathfrak{M} \models \varphi[s]$ is defined recursively in the usual way for atomic formulas and Boolean combinations. For the quantifier we have the clauses:

- $[\![\varphi]\!]^x_s = \{d \in D : \mathfrak{M} \models \varphi[s^d_x]\}$
- $\mathfrak{M} \models \mathsf{F}x(\varphi(x), \psi(x))[s] \iff \exists f : [\![\varphi]\!]^x_s \xrightarrow{1-1} [\![\psi]\!]^x_s$

The definition by simultaneous recursion is standardly extended to the polyadic case.

The standard first order quantifiers are expressible in $\mathscr{L}_\mathsf{F}$:

- $\forall x \varphi(x) = \mathsf{F}x(\neg\varphi(x), x \neq x)$;
- $\exists x \varphi(x) = \neg\mathsf{F}x(\varphi(x), x \neq x)$.

There is an *axiom of infinity* in the pure identity fragment of $\mathscr{L}_{\mathsf{F}}$:

AxInf: $\exists y \mathsf{F} x(x = x, x \neq y)$.

AxInf is true in *all and only* the infinite models, and as a consequence, *compactness fails*.

Abbreviate the statement that $\{x : \phi(x)\}$ is Dedekind finite: $\forall y \neg \mathsf{F} x(\phi(x), \phi(x) \wedge x \neq y)$ by $\mathsf{Fin}\, x(\phi(x))$.

Let φ_ω be the conjunction of the two sentences of $\mathscr{L}_{\mathsf{F}}(<)$:

- $<$ is a strict transitive linear order with a first element; and
- $\forall x\, \mathsf{Fin}\, y(y < x)$.

The sentence φ_ω is true if and only if $<$ has order type $\leq \omega$. Then $\varphi_\omega \wedge \mathsf{AxInf}$ is true precisely if $<$ is a countably infinite linear order. The conjunction of this sentence with axioms giving recursive definitions for addition and multiplication characterizes the standard model $(\mathbb{N}, +, \times)$ up to isomorphism.

The general interpretation of quantifiers

On the standard interpretation, the Frege quantifier is very expressive. Is there an alternative?

Henkin showed that *second-order* quantifiers can be given a general interpretation, on which they range over *some* collection $\mathscr{C}$ of subsets of D, usually satisfying some closure conditions. What seems to have escaped attention is that ***first-order*** **quantifiers can also be so interpreted**.

On the standard interpretation $\exists$ ranges over the collection of *all* non-empty subsets of D and dually $\forall$ denotes $\{D\}$.

On the *general* interpretation, $\exists$ ranges over *some* collection $\mathscr{C}$ of non-empty subsets of D (where $\mathscr{C}$ itself is allowed to be empty), and dually $\forall$ ranges over *some* collection of subsets of D of which D itself is a member.

Interestingly, the question “What is the logic of $\exists$ on the general interpretation?” has a simple if unexpected answer: *positive free logic*.

General semantics for F

A *general model* $\mathfrak{M}$ for $\mathscr{L}_{\mathsf{F}}$ provides a non-empty domain D, interpretations for the non-logical constants, and a collection $\mathscr{F}$ of 1-1 functions $f : A \to B$ with $\text{dom}(f) = A$, and $\text{rng}(f) \subseteq B$, for $A, B \subseteq D^k$. The satisfaction clause for monadic F then is:

$$\mathfrak{M} \models \mathsf{F}x(\phi(x), \psi(x))[s] \iff (\exists f \in \mathscr{F})\, f : [\![\phi]\!]_s^x \xrightarrow{1-1} [\![\psi]\!]_s^x$$

We want $\mathscr{F}$ to satisfy six *closure conditions*:

1. For each $A \subseteq D$, the identity map on A belongs to $\mathscr{F}$, including the empty map on $\emptyset$;
2. if $f_1 : A_1 \to B_1$ and $f_2 : A_2 \to B_2$ are in $\mathscr{F}$, where: $A_1 \cap A_2 = \varnothing$ and $B_1 \cap B_2 = \varnothing$; then $f_1 \cup f_2 \in \mathscr{F}$ as well;
3. if $f : A \to B$ is in $\mathscr{F}$ then also f^{-1} is in $\mathscr{F}$;
4. if $f \in \mathscr{F}$ and $f : A \to B$ and $x \notin A$ and $y \notin B$, then there is a $g \in \mathscr{F}$ such that $g : A \cup \{x\} \to B \cup \{y\}$;
5. if $f : A \to C \in \mathscr{F}$ and $B \subseteq A$ then also $f \restriction B \in \mathscr{F}$;
6. if $f : A \to B$ and $g : B \to C$ are in $\mathscr{F}$ then so is $f \circ g$.

The existence of empty maps ensures that $\forall$ and $\exists$, as defined using F, receive their proper interpretation. Similarly, closure under composition gives transitivity.

We further expand the language by introducing the abstraction operator Num. Strictly speaking, Num is a *variable-binding* operator:

> if x is a variable and φ a formula, $\mathsf{Num}_x\,\varphi(x)$ is a term.

A *general model* is a structure $\mathfrak{M}$ providing a non-empty domain D and interpretations for the non-logical constants, a collection $\mathscr{F}$ of 1-1 functions, as well as a function $\eta : \mathscr{P}(D) \to D$ providing an interpretation for the abstraction operator. (No class of functions need be specified for a *standard model*.)

Satisfaction and reference are now defined by *simultaneous recursion*:

- $\mathfrak{M} \models \mathsf{F}\overline{x}(\varphi(x), \psi(x))[s] \iff (\exists f \in \mathscr{F}) f : [\![\varphi]\!]_s^{\overline{x}} \xrightarrow{1-1} [\![\psi]\!]_s^{\overline{x}}$.
- $[\![\mathsf{Num}\; x.\varphi(x)]\!]_s = \eta([\![\varphi]\!]_s^x)$.

Since $\mathscr{L}_{\mathsf{F}}$ interprets the standard quantifiers $\exists$ and $\forall$, we could just use the Peano-Dedekind axioms (with no mention of Num). But we would rather like to focus on the *cardinal* properties of numbers, rather than their *structural* ones.

Accordingly, we give a *first-order theory* formalizing arithmetic *in the Frege-Russell tradition*, with numbers "representing" equivalence classes of equinumerous concepts. The theory will have some claim to being a *non-reductionist* implementation of logicism.

The language $\mathscr{L}_{\mathsf{F}}$ comprises extra-logical constants: a binary relation $\leq$, and a 1-place predicate $\mathbb{N}$.

Special extra-logical axioms will specify the interaction of Num and the Frege quantifier.

- Hume's Principle:

Ax1 $\mathsf{Num}(\varphi) = \mathsf{Num}(\psi) \leftrightarrow \mathsf{I}z(\varphi(z), \psi(z))$

- Definition of the "less-than" relations:

Ax2 $\mathsf{Num}(\varphi) \leq \mathsf{Num}(\psi) \leftrightarrow \mathsf{F}z(\varphi(z), \psi(z))$;

Ax3 $\mathsf{Num}(\varphi) < \mathsf{Num}(\psi) \leftrightarrow [\mathsf{F}z(\varphi(z), \psi(z)) \wedge \neg \mathsf{F}z(\psi(z), \varphi(z))]$;

- Definition of "x is a natural number":

Ax4 $\forall x(\mathbb{N}(x) \leftrightarrow \mathsf{Fin}\, y(\mathbb{N}(y) \wedge y < x) \wedge x = \mathsf{Num}(\mathbb{N}(y) \wedge y < x))$

"x is a natural numbers if and only if x is the number of the concept 'natural number less than x' and moreover such a concept is finite." ***This a <u>contextual</u> definition of $\mathbb{N}$.***

- The successor relation (Succ implicitly binds a variable):

Ax5 $\mathsf{Succ}(\varphi, \psi) \leftrightarrow \exists x(\psi(x) \wedge \mathsf{I}y(\varphi(y), \psi(y) \wedge y \neq x)$;

Comprehension (and induction?)

- A form of *comprehension* axiom:

Ax5 $\forall x[\varphi(x) \to \exists! y(\psi(y) \wedge \theta(x,y))] \to \mathsf{F}x(\varphi(x), \psi(x))$.

The axiom expresses the closure of the set $\mathscr{F}$ of injections under definability, and therefore subsumes the existence of the empty and identity maps.

The axiom allows us to prove that $\mathbb{N}$ is not Dedekind-finite: it's an "infinitary" axiom. It is also important for the principle of induction.

- We *could* explicitly add a Principle of Induction, in the form "Every finite, non empty set of numbers has a maximum":

$$[\exists x(\mathbb{N}(x) \wedge \varphi(x)) \wedge \mathsf{Fin}\, x(\mathbb{N}(x) \wedge \varphi(x))] \to \exists y[(\mathbb{N}(y) \wedge \varphi(y)) \wedge \forall x(\mathbb{N}(x) \wedge \varphi(x) \to x \leq y)],$$

and then prove $\forall x((\forall y < x)\varphi(y) \to \varphi(x)) \to \forall x \varphi(x)$.

But it turns out that we can provide a direct proof of induction from the remaining axioms.

Arithmetical axioms for + and ×

The theory of successor and $\leq$ can be formulated in the monadic fragment of $\mathscr{L}_{\mathsf{F}}$ (where F only binds a single variable at a time). Representation of addition and multiplication requires the *dyadic* version (so we can count pairs).

Multiplication is more naturally represented than addition: let $\mathsf{Prod}(\varphi, \psi, \theta)$ abbreviate "the number of θ equals the number of φ multiplied by the number of ψ," as follows:

Ax6 $$\mathsf{I}xy(\varphi(x) \wedge \psi(y), \theta(x) \wedge x = y)$$

Addition: let $\mathsf{Sum}(\varphi, \psi, \theta)$ abbreviate "the number of θ equals the number of φ plus the number of ψ," as follows:

Ax7 $$\mathsf{I}xy((x = 0 \wedge \varphi(y)) \vee (x = 1 \wedge \psi(y)), \theta(x) \wedge x = y);$$

We interpret Peano Arithmetic in the theory with F and Num. We proceed semantically, and show that (analogues of) the PA axioms hold in every *general* model satisfying the axioms (and *a fortiori* in every standard model)

- $\mathbb{N}(0)$: where 0 abbreviates $\mathsf{Num}_x(x \neq x)$; i.e., 0 is a number.
- For numbers p and q let Num abbreviate $\mathsf{Succ}(\mathbb{N}(x) \wedge x < p, \mathbb{N}(x) \wedge x < q)$; then every number has a unique successor.
- Every number other than 0 is a successor (it's important that this provable without induction).
- Succ is an injective function.

Moreover, by the infinitary axiom, $\neg\,\mathsf{Fin}_x(\mathbb{N}(x))$ holds as well, with Succ witnessing the injection. The infinitary axiom is crucial also in the proof of induction.

The Principle of Induction

Theorem

The Principle of Induction in the form

$$\left[\varphi(0) \wedge \forall n(\varphi(n) \rightarrow \varphi(n+1))\right] \rightarrow \forall n \varphi(n),$$

can be derived in $\mathscr{L}_{\mathsf{F}}$ *from* **Ax1-Ax5**, *relativizing all quantifiers* $\forall n$ *to* $\mathbb{N}$.

Proof sketch

Assume $\varphi(0)$ and $\forall n(\varphi(n) \rightarrow \varphi(n+1))$ but $\neg\varphi(m)$ for some m. Since $\mathbb{N}(m)$, the set of all $p < m$ is Dedekind finite. Now define:

$$f(p) = \begin{cases} p & \text{if } \varphi(p), \\ p-1 & \text{if } \neg\varphi(p). \end{cases}$$

One checks that: the function is well defined, because $\varphi(0)$; $\operatorname{rng} f$ is a proper subset of $\{p : p < m\}$, because $f(m) < m$; and f is injective. The inifinitary axiom now gives $\mathsf{F}p(p < m, p < m - 1)$, *contra* Dedekind-finiteness. (the elementary arithmetical facts needed for the proof can be established without induction.)

Conclusions

We have thus given an account that is:

1. driven by the cardinal properties of the natural numbers, with the structural properties "supervenient" upon the former (note that *ordinal* properties seem harder to capture);
2. firmly seated in the *Frege-Russell tradition* characterizing the natural numbers as either identical, or intimately connected, to equinumerosity classes;
3. entirely at the *first order* from a semantical point of view.

We leave open the choice between the standard and the general interpretation of F: on the former the axioms are categorical, and on the latter significant arithmetical facts are still representable.

The "ultimate nature" of numbers is left unexplained by the account — as it should be. Numbers are picked as "representatives" of equivalence classes (of which they are not themselves members), but need not (and *cannot*) themselves be *members* those classes.

This leads to a deflation of general worries about abstract objects: rather than being drawn from a separate realm, they are just ordinary objects recruited for the purpose.