Forthcoming in: *History and Philosophy of Logic*

NEO-LOGICISM AND ITS LOGIC

PANU RAATIKAINEN

Philosophy, School of Social Sciences and Humanities
FI-33014 University of Tampere, Finland

The rather unrestrained use of second-order logic in the neo-logicist program is critically examined. It is argued in some detail that it brings with it genuine set-theoretical existence assumptions, and that the mathematical power that Hume's Principle seems to provide, in the derivation of Frege's Theorem, comes largely from the 'logic' assumed rather than from Hume's Principle. It is shown that Hume's Principle is in reality not stronger than the very weak Robinson Arithmetic Q. Consequently, only few rudimentary facts of arithmetic are logically derivable from Hume's Principle. And that hardly counts as a vindication of logicism.

1. Introduction

The view called 'neo-logicism' has received quite a lot of attention recently. There are in fact several different positions which may be called 'neo-logicism',[1] but my interest here is in the neo-logicist program of 'the Scottish school' (also known as the 'neo-Fregean program' or 'Abstractionism') – as defended by Bob Hale and Crispin Wright, for example. The gist of this program is 'Frege's Theorem': the observation that, roughly, the second order Peano axioms can be derived, in second-order logic, from what is known as 'Hume's Principle'. Neo-logicists submit that, even if perhaps not strictly speaking a logical truth, Hume's Principle constitutes a definition of 'cardinal number', and has therefore the same epistemological status as logical truths – it is often said that it is 'quasi-logical' – and that, consequently, the general spirit of logicism can be vindicated.

Hume's Principle is the following:

> (HP) The number of *Fs* is equal to the number of *Gs* if and only if there is a one-to-one correspondence (a bijection) between the *Fs* and the *Gs*.

[1] For example, another well-know neo-logicist program is 'the Stanford-Edmonton school' whose project has been put forward and defended by Bernard Linsky and Ed Zalta (see e.g. *Linsky and Zalta 1995, 2006*).

Hale and Wright summarize the program thus:

> … there are two main claims … which must be seen to hold good if the neo-Fregean leading thesis is to be sustained. The logical claim is that the result of adjoining Hume's Principle to second-order logic is a consistent system which suffices as a *foundation* of arithmetic, in the sense that *all the fundamental laws* of arithmetic are derivable within it as theorems. The philosophical claim is that if that is so, that constitutes a vindication of logicism, on a reasonable understanding of that thesis…. the formal part of the logical claim may be taken to have been established. (*Hale & Wright 2001*, 4–5; my emphasis)

The larger part of the philosophical discussion around neo-logicism has focused on the status of Hume's Principle and other abstraction principles. Much less attention has been paid to the rather unrestrained use of second-order logic (SOL, in brief) in the neo-logicist program. Yet many of those who work actively in the foundations of mathematics and mathematical logic take a dim view of second-order logic. Certainly this tension deserves closer scrutiny. As Wright has recently put it:

> … if the logic used in the abstractionist programme is indeed, as Quine thought, nothing but set theory in disguise, then execution of the various local abstractionist projects, however technically successful, will be of greatly diminished philosophical interest (*Wright 2007*, 152).

There have been few attempts to address these worries (e.g. *Hale & Wright 2001*, *McBride 2003*, *Wright 2007*; cf. *Rossberg 2004, 2006*). The usual response, though, seems to be to suggest that such suspicions are based on some controversial Quinean doctrines, either on Quine's specific criterion of ontological commitment, or on general scepticism concerning properties. Such more explicit discussions of the status of second-order logic in the context of the neo-logicist program are certainly useful, but I shall argue that they greatly underestimate the problems involved here, and that there are deeper and much more basic reasons to resist the uncritical use of second-order logic as the background logic, at least in the context of the foundations of mathematics.[2]

The paper is structured as follows. In Section 2, different senses, and various axiomatic systems, of second-order logic, as well as their history and development, are reviewed; in Section 3, a series of axiomatic theories of arithmetic and set theory, from weak to strong, is presented; they function as the measure of set-theoretical power employed in the remainder of this paper. After all this necessary stage-setting, Section 4 provides our central critical argument; in section 5, we draw some conclusions.

2. Second-Order Logic

[2] I should perhaps add that I am not advocating any general scepticism concerning SOL. In many contexts, its use may be harmless enough, just as it is often unproblematic to lean on mathematics in studying various other fields. I only want to press that one must be particularly careful when one is working in the foundations of mathematics – a view on which Wright, and also Rossberg, apparently agree (see below).

2.1. Different variants of SOL

To begin with, one must clearly distinguish the *full-blown second-order logic*, defined with the help of the model-theoretic relation of logical consequence, and the various *axiomatic deductive systems of second-order logic*. The former is unquestionably extremely strong and is entangled with strong set theoretical issues; it is anything but accessible to the human mind.[3] It may have been a bit unclear which one of these is intended in the early discussions of neo-logicism, but it has now become quite clear that is must be the axiomatic deductive system that is relevant. But arguably that choice does not make all the problems disappear.

As to the axiomatic theories of SOL, the classical and most usual set of rules seems to be the following; indeed, many expositions of SOL[4] present these rules only, or rules equivalent to them (We'll skip ($\exists^2$E), as its formulation is complicated, and it does not play an important role in the following):

$$\frac{A(X_i)}{\forall X_i\, A(X_i)}\ (\forall^2\mathrm{I}) \qquad \frac{\forall X_i\, A(X_i)}{A^*}\ (\forall^2\mathrm{E}) \qquad \frac{A^*}{\exists X_i\, A(X_i)}\ (\exists^2\mathrm{I}) \qquad \ldots\ (\exists^2\mathrm{E})$$

where A^* is obtained from $A(X^n)$ by replacing all occurrences of $X(t_1,\ldots, t_n)$ by $B(t_1,\ldots, t_n)$ for a fixed formula $B(x_1,\ldots, x_n)$.

Note that ($\forall^2$E) entails the full (impredicative) comprehension scheme (as does ($\exists^2$I):

$$\exists X^n \forall x_1 \ldots x_n\, [A(x_1 \ldots x_n) \leftrightarrow X^n(x_1 \ldots x_n)],$$

where A is any formula not containing X^n free.

However, it would be a mistake to assume, as some seem to do, that these rules would be the only natural choice, or have some privileged status. There is a more minimal and neutral system for the second order quantifiers, called the 'Basic Calculus' by Takeuti

[3] It is often noted that SOL is not complete, that is, that the set of its logical truths is not recursively enumerable (r.e.), or Σ_1^0. Though true, this formulation greatly underestimates the difficulty: logical truth in full-blown second-order logic is not co-r.e., or Π_1^0, either; it is not trial-and-error decidable, or decidable-in-the-limit (Δ_2^0); actually it is nowhere in the arithmetical hierarchy (that is, it is not Σ_n^0, for any finite n); it is not definable in the second-order arithmetic (that is, it is not Σ_n^1, for any finite n); in fact, it is not definable in any finite-order arithmetic, i.e., it is not in any finite level of the Kleene hierarchies; it is not Σ_n^m, for any finite m and n).
Montague was apparently fond of saying that the set of second-order validities did not belong to any Kleene hierarchy, 'past, present, or future'. Vaught in turn has reportedly remarked that studying second-order logic is like studying 'the standard model of set theory'. (Enderton, FOM, Sept 1 2000.)

[4] E.g. *van Dalen 2004; Leivant 1994; Troelstra & Schwichtenberg 1996; Prawitz 1965, 1973; Shapiro 1993; Rossberg 2006*.

(*1975*), and by Manzano (*1996*), simply 'C_2^-' (thought it is more typically formulated in the sequent calculus rather than in the natural deduction form). The possibility of such a system must have been clear at least since (*Henkin 1953*) (see below).

We take as rules the obvious analogues of the first-order rules, the Basic Rules:

$$\frac{A\,(X_i)}{\forall X_i\, A(X_i)}\ (\forall^2\mathrm{I}) \qquad \frac{\forall X_i\, A(X_i)}{A(X_i/X_j)}\ (\forall^2\mathrm{E}) \qquad \frac{(X_i/X_j)}{\exists X_i\, A(X_i)}\ (\exists^2\mathrm{I}) \qquad \ldots\ (\exists^2\mathrm{E})$$

with the usual restrictions. Note that these rules require instantiation by a *variable*.

The differences between these two systems are highly relevant in what follows. We can get the "standard system' by adding the Comprehension scheme as an extra-logical assumption to the Basic Calculus. There is also an intermediate system, with only predicative comprehension.

It is by no means obvious that the strongest of these, with unrestricted and impredicative comprehension scheme (or rules equivalent to it), is the only plausible, or even the best or most natural choice in all contexts. Its status as a neutral background logic in the context of the foundations of mathematics is also more problematic – or so I shall argue. However, it is this version that neo-logicism requires.

Now both sets of rules can be motivated – to some extent – by leaning on certain (different) analogies with the basic first order rules for quantifiers. But such analogies are, of course, quite a weak justification, and analogy is obviously a potentially misleading guide. What is needed is a more systematic, theoretical study of the properties of the system, and their behavior in various critical contexts.

2.2. Digression: Development of the Axiomatic SOL

The axiomatic deductive systems were not always formulated in the above way – in the form of introduction and elimination rules for quantifiers. In fact, they were not even anything close to this. In the first half of the 20th century, people were still hung up over *rules of substitution* for predicate variables, and in particular Church, in his well-known book *Introduction to Mathematical Logic* (*1944/1956*), formulated second-order logic (following Hilbert-Ackermann 1928 book *Grundzüge der theoretischen Logik*) using a very complicated rule of substitution; people were always getting it wrong. Another influential text-book of the period, Kleene's *Introduction to Metamathematics* (*1952*), only mentions SOL in passing and refers to the two above sources.

In first-order logic, in addition to some convenient set of rules, e.g. *Modus Ponens* and *Universal Generalization*, one could have either infinite number of axioms, given be an axiom schema, or, alternatively, the number of axioms may be finite, if there is added a rule of substitution. However, because an accurate definition of the substitution operation

is very complex and difficult, a formulation which avoids it has been generally taken to be preferable. However, in the case of SOL, in contrast, it was not at all clear whether and how the substitution operation could be avoided. And for SOL, the rule of substitution was even more complicated, and people had great difficulties in getting is right, and understanding what was going on in it.

Henkin then published an important paper 'Banishing the Rule of Substitution' (*1953*) in which he pointed out that substitution, in SOL, could be replaced by the comprehension scheme. He formulates a weaker system *F** which – although a Hilbert-style axiomatic system – is equivalent to the Basic Calculus. Henkin next demonstrates that by adding the comprehension schema, one gets a system *F***, which is equivalent to the more traditional 'standard systems' of second-order logic; but neither system of Henkin requires any troublesome substitution operations. In a footnote, Henkin refers to Tarski, who had pointed out to him that essentially the same more neutral system is described in Tarski's famous paper on truth-definitions (*Tarski 1933/1935*), where the system is credited to Lesniewski; the instances of comprehension schema were called by Lesniewski and Tarski 'pseudo-definitions'.[5]

In any case, Henkin's analysis also calls attention to the subsystem *F** (or Basic Calculus); he notes that its existence had been previously generally obscured. We can then also naturally distinguish the neutral *logical* axioms and rules of inference, the Basic Calculus, and view the comprehension scheme as an additional non-logical assumption. Henkin also considered intermediate subsystems which one would get by weakening the comprehension schema, e.g. restricting it to the predicative instances. These contributions of Henkin were path-breaking; but it seems to me that – unlike his other contributions to e.g. general vs standard models of SOL – even today they are not as well-known as they deserve to be.

The axiomatic systems of second-order logic can be viewed as systems in many-sorted first-order logic (perhaps, as in the case of the standard system, equipped with some non-logical axioms). Schmidt had already presented many-sorted predicate logic in *1938*, but it seems that only Hao Wang's *1952* paper on many-sorted logic in JSL started to have some impact.

Gentzen very briefly considered second-order quantifiers in his late 1943 paper. He did not present any calculus governing them, but suggested in passing that one can apply the rules (∀E) and (∃I). The explicit formulation and systematic study of a Gentzen-type system for SOL (sequent calculus, not natural deduction) was left to Takeuti (*1953*). He only considered sequent calculus incorporating impredicative comprehension; and such systems became standard in proof theory, especially in the 1950s and 1960s. However, it seems that the focus on these rules was more a result of the interest in proving the consistency of the second-order arithmetic **PA2** via the proof of ('Takeuti's conjecture') cut-elimination

[5] Apparently *this* part of Tarski's seminal work had failed to have any notable impact on the logical community.

for this system (the latter entails the former) than any interest in arguing that these rules are strictly speaking logically true – the interest in them was much more pragmatic.

In his classic 1965 book *Natural Deduction*, Prawitz presented natural deduction rules for SOL. These were, again, strong rules. Also Prawitz was at the time involved in the study of Takeuti's conjecture, which was settled non-constructively by *Tait 1966, Prawitz 1967, and Takashi 1967*; and constructively by *Girard 1971*. By the publication of his book *Proof Theory* (*1975*, lectures in 1968), Takeuti had in any case clearly distinguished three Gentzen-type sequent systems: Basic Calculus, and systems with impredicative or predicative comprehension.

3. Some Axiomatic Theories

In what follows, we need to compare several systems of first- and second-order arithmetic and set theory. We next survey some of the key systems.

3.1. First Order Theories of Arithmetic

The weakest standard system of arithmetic we'll consider is the so-called Robinson arithmetic (due to Raphael M. Robinson; see *Tarski, Mostowski and Robinson 1953*), standardly denoted as **Q**. As axioms, it has the following seven assumptions:

$$\neg(0 = S(x))$$
$$S(x) = S(y) \rightarrow x = y$$
$$\neg(x = 0) \rightarrow (\exists y)(x = \mathrm{S}(y))$$
$$x + 0 = x$$
$$x + S(y) = S(x + y)$$
$$x \times 0 = 0$$
$$x \times S(y) = (x \times y) + x$$

The intended interpretation of 'S' is the successor function, and obviously, of + and ×, the addition and the multiplication functions, respectively. '0' is the only constant and denotes the first element, zero. **Q** is a very weak theory. To quote Burgess, 'virtually no mathematics can be done in the system **Q**' (Burgess 2005, 56). For example, we can't even prove the general claim that addition is commutative, i.e., we cannot derive $(\forall x)\ (\forall y)(x + y = y + x)$.

If we add to the axioms of **Q** the first order induction scheme,

(IND) $\quad [\varphi(0)\ \&\ (\forall x)\ (\varphi(x) \rightarrow \varphi(S(x)))] \ \rightarrow\ (\forall x)\varphi(x),$

where $\varphi(x)$ can be any formula of the language, we get the standard first order theory of arithmetic, i.e. the (first-order) Peano Arithmetic, **PA**. In this theory, in contrast, all ordinary number theory can be easily developed.

Another natural and much-studied arithmetical system, which lies between **Q** and **PA** for its strength, is the Primitive Recursive Arithmetic (**PRA**), which contains (not just the above axioms of **Q** governing successor, addition and multiplication, but) defining axioms for all primitive recursive functions, and the application of the induction scheme is restricted to quantifier-free formulas (i.e. $\varphi(x)$ is not allowed to contain any (unbounded) quantifiers. However, essentially the same system is achieved if we take just the axioms of **Q** and the induction scheme restricted to (roughly) purely existential formulas (Σ_1^0-formulas, as logicians call them) (this was first showed by *Parsons 1970*).

We thus have the following containments:

		Q + Σ_1^0-induction		
Q	$\subset$	PRA	$\subset$	PA

3.2. Subsystems of the Second Order Arithmetic and Reverse Mathematics

Unquestionably one of the greatest success stories in the foundations of mathematics in the last few decades is the study of various subsystems of second order arithmetic and the research program of so-called Reverse Mathematics initiated by Harvey Friedman (*1975*) and developed, in addition of Friedman himself, especially by Stephen Simpson (see esp. *Simpson 1999*). The purpose of this research program is to obtain precise answers to the following question:

> Which set existence axioms are needed to prove the theorems of ordinary mathematics?

I want to suggest that it is highly illuminating to compare the results of this program to the neo-logicist program.

The starting point of the program is the old observation; namely, already Hilbert and Bernays (*1934-39*) had pointed out that all ordinary mathematics can be developed in the full second-order arithmetic **PA2**. Therefore, by the way, it would be indeed a fantastic achievement of it could be demonstrated that it has, in some sense, the same epistemological virtues that logical truths have.

In his 1974 address to the International Congress of Mathematics, Harvey Friedman undertook to study the following question: 'What are the proper axioms to use in carrying out the proof of particular theorems, or bodies of theorems, in mathematics? What are the formal systems which isolate the essential principles needed to prove them?' He went on to explain that certain subsystems of **PA2** are of fundamental importance in this context (see *Friedman 1975*).

The main theme of Friedman is that, when a theorem of ordinary mathematics is provable in one of the isolated axiom systems, then surprisingly often, the theorem is in fact provably equivalent to the principal axiom of the system needed to prove it. That is,

opposite to the usual mathematical practice, in which theorems are deduced from the axioms, Friedman has also deduced certain central axioms from various theorems of ordinary mathematics. Such a deduction of axioms from theorems is known as *Reverse Mathematics*. This provides exact answers to the questions such as:

> How much set existence is needed to prove a given theorem of ordinary mathematics?
>
> Which set existence postulates are necessary and sufficient?

The *language* of **PA2**, denoted L2, is a two-sorted first order language with number variables $i, j, n, m, \ldots$ and set variables $X, Y, Z, \ldots$ The number variables are intended to range over the set of natural numbers $\mathbf{N} = \omega = \{0, 1, 2, 3, .\}$, The set variables are intended to range over subsets of **N**. Numerical terms are built up as usual from number variables, constant symbol 0, and an unary operation S and binary operations + and $\times$. Atomic formulas are $t = t'$, $t < t'$ and $t \in X$, where t and t' are numerical terms. Formulas are built up from atomic formulas by means of propositional connectives, number quantifiers $\forall n$ and $\exists n$, and set quantifiers $\forall X$ and $\exists X$. A formula is called *arithmetical* if it contains no set quantifiers – however, an arithmetical formula may contain free set variables.

The following relation between theories plays a fundamental role:

Conservative extension: a theory T_1 is conservative over a theory T_2 for a class of formulas F, if

For all S $[(S \in F \ \& \ T_1 \vdash S) \rightarrow T_2 \vdash S]$.

Often, F is simply the whole language of T_2; and in all the cases that we shall consider below, F is the class of (first order) arithmetical formulas.

Q+ and PA2

Let us again begin with the weak first order theory **Q** (i.e. the Robinson Arithmetic). We then add to it the following *Induction Axiom*:

(**IA**) $\quad (0 \in X \wedge \forall n\,(n \in X \rightarrow S(n) \in X)) \rightarrow \forall n\,(n \in X)$.

Note that this is a single axiom, not an axiom schema. Indeed, it is *essential* that we do not, at this point, allow the induction schema. (It is also critical which underlying logic we accept; I am getting ahead, but it is obligatory in this context to assume only the rules of Basic Calculus; more of this below.)

Let us focus for a moment to the basic theory consisting of **Q** and **IA**. Now it is very important not to confuse this theory with the full second order arithmetic **PA2**. For this

theory is conservative over **Q**: it can only prove the same (very few) arithmetical theorems as **Q** – it is much weaker than even the first-order Peano Arithmetic **PA** – indeed, weaker than even Primitive Recursive Arithmetic **PRA**. For this reason, I shall call this extremely weak second order theory **Q+**. It is mathematically quite useless. As with **Q**, in the words of Burgess, 'virtually no mathematics can be done' in **Q+**. In the literature in Reverse Mathematics, the axioms of **Q+** are simply taken as 'basic axioms', and what is studied is various substantial extensions of them.

At the other extreme is:

The Full Second Order Arithmetic, PA2 (aka Z_2)

Add to **Q+** the unrestricted comprehension scheme:

$$\exists X\, \forall n\, (n \in X \leftrightarrow \varphi(n)),$$

where $\varphi(n)$ can be any formula of the language of **PA2** in which X does not occur freely.

This is a very strong theory: already Hilbert and Bernays (1934-39) had pointed out that all ordinary mathematics can be developed in **PA2**.

For example, Descriptive Set Theory is a field of mathematics related to topology. It was initiated by the French semi-intuitionists (Lebesgue, Baire, Borel), and studies certain 'well-behaved' sets which possess relatively simple definition (in contradistinction to the ideas of arbitrary sets and various higher power-sets, which the semi-intuitionists doomed as meaningless). In particular, Descriptive Set Theory studies so-called Borel and Analytic sets. Now **PA2** can prove, among other things, the existence of sets that are non-Borel and non-Analytic.

In sum, **PA2** is a very powerful theory, and it would be striking indeed if one could show that it is quasi-logical in its nature – that it has in some sense the same epistemological status as logical truths. But I am getting ahead.

Subsystems of PA2

RCA_0. The weakest system that is taken seriously is the one known as **RCA_0**; the name refers to its principal axiom, 'the Recursive Comprehension Axiom', which postulates the existence of at least recursive sets.[6] The subscript 0 denotes restricted induction, i.e. we

[6] For technical reasons, the formulation of the relevant comprehension axiom (scheme) requires some care. As a set is recursive if and only if both the set and its complement are definable by a Σ_1^0-formula (a purely existential first order arithmetical formula), the key axiom scheme can be stated as:

$$\forall n\, (\varphi(n) \leftrightarrow \psi(n)) \rightarrow \exists X\, \forall n\, (n \in X \leftrightarrow \varphi(n)),$$

where $\varphi(n)$ and $\psi(n)$ are Σ_1^0-formulas and Π_1^0-formulas, respectively. Obviously we also have the basic axioms of **Q+**. The minimum ω-model of **RCA_0** is just the class of recursive subsets of ω.

assume only the Induction Axiom and not the full induction scheme. Such a restricted system T_0 is in general weaker than T since it contains induction only for those sets which can be proved to exist within T_0.

The full induction scheme is itself equivalent to a set existence principle, namely the bounded comprehension scheme

$$\forall n\, \exists X\, \forall m\, (m \in X \leftrightarrow (m < n \wedge \varphi(m)),$$

where $\varphi(m)$ is any formula of the language of **PA2**. Therefore, it cannot be uncritically accepted.

In $\mathbf{RCA_0}$, one can prove some basic facts of number theory such as the fundamental theorem of arithmetic, and develop a certain amount of analysis. $\mathbf{RCA_0}$ is conservative over the Primitive Recursive Arithmetic **PRA**, i.e. they prove exactly the same first-order or arithmetical theorems.

$\mathbf{ACA_0}$**.** Another important subsystem of **PA2** that Friedman has isolated is the one obtained by restricting the comprehension scheme to *arithmetical* ('first order') formulas. This theory is denoted by $\mathbf{ACA_0}$**.** (If we focus on comprehension and its restrictions, there aren't really systems between $\mathbf{ACA_0}$ and $\mathbf{RCA_0}$, for allowing either Π_1^0**-CA** or Σ_1^0**-CA** already gives $\mathbf{ACA_0}$**.)**

$\mathbf{ACA_0}$ is enough to prove all the theorems of ordinary analysis and algebra. $\mathbf{ACA_0}$ proves König's Lemma, and $\mathbf{ACA_0}$ is indeed equivalent to it. $\mathbf{ACA_0}$ is conservative extension of the ordinary (first-order) **PA**.

Π_1^1**-CA**$_0$. This strong subsystem has the comprehension for simple, purely universal second-order formulas, i.e. formulas with the form $\forall X\, \varphi$ (where φ is arithmetical) which logicians call Π_1^1-formulas.

In this subsystem, even a number of central theorems of Descriptive Set Theory can be proved, e.g. Kondo's theorem on uniformization for co-analytic sets, and the Cantor-Bendixson theorem (both these theorems turn out to be equivalent with Π_1^1**-CA**).

Π_2^1**-CA**$_0$**.** Even stronger is the system with the comprehension scheme for the second-order formulas of the form $\forall X \exists Y\, \varphi$ (where φ is arithmetical). Recently, Mummert and Simpson have shown that a certain metrization theorem in Topology necessarily require this much (*Mummert & Simpson 2005*). This is apparently the first convincing instance of a core mathematical theorem which requires exactly Π_2^1**-CA**$_0$. Still, this is only a fragment of the full **PA2**. Hence, I repeat, it would be a great discovery if **PA2** could be, in some sense, founded on pure logic.

* * *

In sum we have the following containments of second-order (SO) and first-order (FO) (arithmetical theories; $\Downarrow$ denotes the relation of being a conservative extension:

$$\begin{array}{llllllll}
\text{SO theories:} & \text{Q+} & \subset & \text{RCA}_0 & \subset & \text{ACA}_0 & \subset & \Pi_1^1\text{-CA}_0 \subset \Pi_2^1\text{-CA}_0 \ldots \subset \text{PA2} \\
 & \Downarrow & & \Downarrow & & \Downarrow & & \\
\text{FO theories:} & \text{Q} & \subset & \text{PRA} & \subset & \text{PA} & &
\end{array}$$

There are a couple of important intermediate systems (WKL_0, ATR_0) also studied in the program of reverse mathematics, but we need not consider them now – what matters is the big picture; and, in particular, the great distance between **Q+** and **PA2**.

3.3. Interpretability

Another very important relation between theories, which we soon need, is (relative) interpretability. It allows the comparison of theories when one is not a direct extension of the other or when the two have totally different languages. Very roughly, T_1 is interpretable in T_2 if the language of T_1 can be translated into the language of T_2 in such a way that T_2 proves the translation of every theorem of T_1. We then write $T_1 \leq T_2$.

Slightly more exactly, one has an (relative) interpretation of T_1 into T_2 if with each basic relation, function and constant symbol of the language $L(T_1)$ is associated as its interpretation a definition of it in the language $L(T_2)$ and with each sort of variable s in $L(T_1)$ is associated a defined range of variation given by a formula $D_{T1}(x)$ of $L(T_2)$. Then with each formula φ of $L(T_1)$ is associated as its interpretation in $L(T_2)$ a formula $f(\varphi)$ obtained by substituting the respective definitions for the basic symbols and relativizing quantified variables of sort s to $D_s(x)$. Such an effective map f of the formulas of $L(T_1)$ into the formulas of $L(T_2)$ is defined to constitute a relative interpretation of T_1 into T_2, if we have

$$T_1 \vdash S \Rightarrow T_2 \vdash f(S).$$

T_1 is *interpretable* in T_2, $T_1 \leq T_2$, if there is an interpretation of T_1 in T_2. Two theories are called *mutually interpretable* if each is interpretable in the other.

This notion of interpretability was first given an explicit definition by Tarski in (Tarski, Mostowski and Robinson 1953). It had been, however, already used in practice by logicians for some time. The relation of interpretability serves, among other things, as a measure of strength of formal systems. That is, it is reasonable to say that T is essentially stronger than T_1 under the relation $\leq$ of relative interpretability if $T_1 \leq T_2$, but not $T_2 \leq T_1$.

Some important properties of interpretability are the following:

- If $T_1 \leq T_2$ and T_2 is consistent, so is T_1.

- For all T, $T \leq T$

- Composed interpretation: If $T_1 \leq T_2$ and $T_2 \leq T_3$, then $T_1 \leq T_3$.

Well known examples of relative interpretability include:

PA $\leq$ **ZFC**, the interpretation of Peano Arithmetic in Zermelo-Fraenkel set theory;

ZF+AC+GCH $\leq$ **ZF**, via Gödel's model of **ZFC+AC+GCH** in the constructible sets.

Neo-logicism and interpretability. In fact, the notion of interpretability also plays an essential role in the neo-logicist program. Namely, what Frege's Theorem, more exactly, amounts to, is:

PA2 can be *interpreted* in the theory consisting of Hume's Principle and the standard SOL (impredicative), that is, in Frege Arithmetic.

It is this theorem that the neo-logicists refer to when they announce that **PA2** can be 'derived' or 'developed' in Frege Arithmetic.

Faithful interpretability. Now mere interpretability (and also mutual interpretability) is yet a rather weak and flexible relation between theories, and it is possible to have a pair of theories with clearly different strengths that are mutually interpretable. The following notion, due to Feferman, Kreisel and Orey (*1960*), is stricter in this respect:

We say that T_2 *faithfully interprets* T_1, if and only if, for some translation f, and, for all T_1-sentences S,

$$T_1 \vdash S \Leftrightarrow T_2 \vdash f(S).$$

Note the change from $\Rightarrow$ to $\Leftrightarrow$. In other words, faithful interpretability requires, in addition to ordinary interpretability, that translations of non-theorems be non-theorems. There are many cases in which a theory is interpretable but not faithfully interpretable in another theory.

In the present paper, though, we focus solely on theories which are (presumably) Σ_1^0-sound, i.e. prove only arithmetical Σ_1^0-statements which are true, or to theories which can at least interpret such an arithmetical theory. We can therefore apply Visser's extension of Lindström's result[7] and achieve the conclusion that:

[7] *See Lindström 1984; Lindström 2003* (Chapter 6, Section 2, Theorem 13); *Visser 2005* (Lemma 5.4).

Corollary (Lindström-Visser): If we have a pair of mutually interpretable theories both of which can interpret a Σ_1^0-sound arithmetical theory, we can always strengthen the interpretations between them to faithful ones.

Consequently, all relations of mutual interpretability in this paper can be improved to faithful ones.

3.3. Some Set Theoretical Systems

As to different set theories, let us begin with an extremely weak theory. It is known under many different names:

'Baby Set Theory', *aka* Szmielew-Tarski Set Theory, *aka* Adjunctive Set Theory:

AS1. $\exists x\, \forall y\, (x \notin y)$ the empty set: 'There is a set such that no set is a member of it.'

AS2. $\forall x, y\, \exists z\, \forall u\, (u \in z \leftrightarrow (u \in x \vee u = y))$. Adjunction: "If x and y are sets, then there exists a set z, the *adjunction* of x and y, whose members are just y and the members of x.'

Intuitively, adjunction $x \wedge y$ is simply $x \cup \{y\}$: it provides the successor operation for the finite von Neumann ordinals[8] (sometimes one includes the Axiom of Extensionality, sometimes not, but the details do no matter here). The important fact is that the Robinson arithmetic **Q** and Baby Set Theory (with or without Extensionality) can be interpreted in each other (faithfully), and hence have essentially the same strength – both are extremely weak.

As to the strong systems in set theory, the familiar standard choice is of course **ZFC**. However, actually the Power Set Axiom (**Pow**) is hardly ever used in ordinary mathematics. If we drop it, we end up with a still very strong system **ZF** – **Pow**, often denoted simply as $\mathbf{ZF}^-$. This theory can be (faithfully) interpreted in **PA2**, and *vice versa*; hence, they have essentially the same strength. I have earlier pressed the deep difference between **Q+** and **PA2**. Similarly, there is a huge gap between Baby Set Theory and $\mathbf{ZF}^-$.

If, instead, we drop the Axiom of Infinity (**Inf**) from **ZFC**, we have the system **ZF** – **Inf**, which has essentially the strength of the standard first-order Peano Arithmetic **PA**., i.e. each of these theories can be (faithfully) interpreted in the other.

We can summarize the various containments between our theories, where $\Downarrow$ means conservative extension, and $\uparrow\downarrow$ stands for mutual interpretability, in the following diagram:

[8] The standard, though only one of the many possible, interpretation of arithmetic in set theory, due to von Neumann, is to interpret 0, 1, 2, … as $\varnothing$, $\{\varnothing\}$, $\{\varnothing, \{\varnothing\}\}$, …

SO theory:	Q+	$\subset$	RCA_0	$\subset$	ACA_0	$\subset$	Π_1^1-CA_0	$\subset$	PA2		
	$\Downarrow$		$\Downarrow$		$\Downarrow$						
FO theory:	Q	$\subset$	PRA	$\subset$	PA				$\uparrow\downarrow$		
	$\uparrow\downarrow$				$\uparrow\downarrow$						
Set theory:	Baby Set Theory	$\subset$			ZF – Inf		$\subset$		ZF^-	$\subset$	ZFC

4. Critical Considerations

4.1. Hume's Principle and Q

Now we come to my first critical observation. To speak roughly at first: Enter the 'standard' (impredicative) second order rules of inference, carefreely granted by the neo-logicists, and you lose, in a sense, the difference between these extremely weak and very strong theories. That is, if we accept as little as **Q+**, the background logic flings us directly to the poweful **PA2.** Unless we introduce some specific method of keeping track of different methods of proof used, we can't make the distinction between them, and numerous theories in between. Recall that **Q+** corresponds to (is faithfully mutually interpretable with) Baby Set Theory, whereas in **PA2** it is possible to develop as much set theory as $\mathbf{ZF^-}$.

This is indeed a huge leap, and it is somewhat problematic if it is allowed by the mere rules of the background logic. In particular, focusing on corresponding theories of sets, $\mathbf{ZF^-}$ includes, and Baby Set Theory does not, the unrestricted Axiom of Separation, and the Axiom of Infinity, and the Axiom of Replacement. It is doubtful whether the methods of proof which result in such strong set existence assumptions could be taken simply for granted in foundational considerations: It seems suspicious if mere background logic justifies the jump from a theory with the strength of Baby Set Theory to a theory with the strength of $\mathbf{ZF^-}$.

One doesn't need to commit oneself to any specifically Quinean doctrines to get worried here. It begins to look as if the mathematical power that Hume's Principle **HP** seems to provide, in the derivation of Frege's Theorem, may well come from the logic assumed rather than from Hume's Principle.

The rules incorporating impredicative comprehension give **PA2**, given either the weak **Q+,** or **HP**. However, the fact that **HP** does not need to do any more work than **Q+** does not of course as such entail that it not stronger than Q+. The above observations as such leave this open. We can do better, though. (Here I am deeply indebted to Albert Visser, who, in personal correspondence, confirmed my working hypothesis and pointed out how this can be proved.[9])

[9] In fact, given various known results, the proof is rather simple: First, one can lean on the fact that the theory known as $\boldsymbol{S_2^1}$ is interpretable in **Q**. Second, the theory of sequences with adjunction over the original numbers can be interpreted in $\boldsymbol{S_2^1}$**.** The sequences can then be used to simulate classes. We let the original numbers stand for the objects. The mapping from a class to its cardinality is clearly present in $\boldsymbol{S_2^1}$**.** We take this as the Hume mapping.
The opposite direction is, as such, more interesting, and is studied in depth in (*Visser 2010*).

Namely, Hume's Principle HP, if equipped only with uncontroversial logical principles (the Basic Calculus or something equivalent), can be faithfully interpreted in **Q**, and has at most the strength of **Q** (and nothing like **PA2**, not by a far cry, contrary to what the neo-logicist literature often seems to suggest). There is a definite sense in which only few rudimentary facts of arithmetic are logically derivable from **HP**. And that hardly counts as a vindication of logicism.

It is well known that **Q** can in turn be interpreted in **HP**. However, even this requires some auxiliary assumptions, the (admittedly weak) adjunctive relation theory (not to be confused with the Adjunctive Set Theory) (see *Burgess 2005*). Interestingly, on the one hand, if we start with **Q+** and add the predicative rules which entail predicative comprehension, we get $\mathbf{ACA_0}$; on the other hand, if we begin with **HP** and add these rules, we still get only **Q**. So, **HP** is, in a sense, even weaker than **Q+** (see also *Visser 2010*).

In sum, the picture that emerges is the following: **HP**, augmented with the adjunctive relation theory, gives us a weak subtheory of arithmetic. If we then add more and more strength to our set-theoretical assumptions (comprehension axioms), we get more and more arithmetic and analysis.

4.2. Is our use of interpretability problematic?

We have above used heavily on the relation of *interpretability* between theories. We are not, however, alone in this, for also neo-logicism leans essentially on this very notion. Namely, we noted above that the more exact content of Frege's Theorem is that **PA2** can be *interpreted* in the theory consisting of Hume's Principle **HP** and the standard SOL (impredicative), i.e., Frege Arithmetic. Further, neo-logicism submits that, as Hume's Principle is arguably quasi-logical, and **PA2** is interpretable in it (and logic (SOL)), **PA2** therefore inherits the epistemological status of the former, and is likewise quasi-logical.

Neo-logicism thus tacitly assumes the following principle:

> If a theory T_1 is quasi-logical, and a theory T_2 is interpretable in T_1, then T_2 is also quasi-logical.

Our reasoning here is simply the contrapositive of that:

> If a theory T_2 is *not* quasi-logical, and T_2 is interpretable in T_1, then T_1 is *not* quasi-logical either.

We have here taken T_2 to be a certain relatively strong system of set-theory with substantial set existence postulates ($\mathbf{ZF^-}$), which I take all parties of the debate to agree to be non-quasi-logical, and T_1 to be the standard **SOL** (impredicative) together with **HP**. Hence either Hume's Principle or the impredicative second-order rules for quantifiers must be *non*-quasi-logical; as Neo-logicism insists that the former is quasi-logical, it must be concluded that the latter are not.

In sum, our use of the relation of interpretability cannot be questioned in this context.

5. Logicism that deserves the name?

Here are some representative quotations from the literature:

> [HP] … suffices as a *foundation* of arithmetic, in the sense that *all the fundamental laws* of arithmetic are derivable within it as theorems. (*Hale & Wright 2001*, 4)

> … the fundamental facts about addition can be proved on the basis of this definition … the rest of arithmetic carried out, all on the basis of just one simple, consistent, and trivial-looking axiom, the number principle [HP]… *arithmetic can be derived* from the number principle [HP]. (*Boolos 1998*, 154)

> PA2 is adequate to develop the classical theory of natural and real numbers, and probably adequate for most of mainstream mathematics. … What the book [Wright 1983] indicates is that PA2 can be derived from FA2 ['Frege Arithmetic'], or in other words, that within pure second-order logic, *PA2 can be derived from HP*. (*Burgess 1998*, 139)

Whatever their actual intent, it is very easy to read such brief statements as suggesting that the full second-order arithmetic **PA2** can be derived from **HP** alone, without any other substantial assumptions. But we have now seen that this is not really the case. Only some most rudimentary arithmetical fact can be thus achieved. The rest can only be derived by leaning on something that amounts to substantial set existence assumptions (in this context problematic impredicative comprehension axioms).

Let me end by quoting Marcus Rossberg, from his St Andrews dissertation (*Rossberg 2006*; written under the supervision of Wright); this work is the most detailed *defence* of the use of SOL from the neo-logicist camp:

> The problem with this is not that set theory is false, but that it is a strong mathematical theory, and not logic. Set theory makes enough substantial mathematical claims that virtually all areas of mathematics can be represented in it. …
>
> Since mathematical truths are, presumably, necessarily true, if true at all, why would that matter? It does not always matter. Mathematics is applicable in the sciences, for example, and presumably does not cause any problems there. Indeed, many think it is indispensable. There are areas, on the other hand, where no mathematics should be presupposed. The philosophy of mathematics is one such area, at least as construed by some research projects in this area. … Another research programme that should not presuppose any mathematics is logicism.

> The logic that Neo-Fregeanism uses is second-order logic, as it was for Frege. If second-order logic is indeed set theory, the project loses almost all its interest: a reduction of arithmetic and analysis to set theory is no news. In particular, it would not show that arithmetic inherits its epistemic status from Hume's Principle (which is arguably analytic).

I whole-heartedly agree with Rossberg here, though I am afraid that our conclusions are the opposite; he hopes to save SOL from such accusations; I have aimed to show that they are to the point.

Obviously the observations in this paper do not make all the work around Hume's Principle uninteresting, fruitless or futile; much of the logical research in this area remains interesting. However, they do cast doubt over certain more ambitious philosophical conclusions that some have drawn from Frege's Theorem.

Acknowledgements

This work started from a seemingly innocent question at a dinner table by Peter Clark in St Andrews more than a decade ago; I hope this paper answers the question. My understanding of the development of the axiomatic systems of second-order logic has greatly benefited from certain postings by Martin Davis for the FOM list back in 1999-2000. My debt to Albert Visser in this paper is obvious and enormous; I would like to express my warmest thanks to him for all his advice and help with these issues. I also thank Richard Heck for his helpful advice with the rules of inference. An earlier version of this paper was presented at the Jowett Society in Oxford (October 2010) and at *The Number Concept* -workshop in Nancy (November 2010); I would like to thank all those who participated in the discussions in these occasions. I am very grateful to Juhani Yli-Vakkuri for his detailed comments on a draft of the paper; the structure of the argument also improved due to a discussion thread on Yli-Vakkuri's Facebook wall. I also thank the anonymous referee for HPL, whose comments helped to make the paper much better. Naturally, I am still solely responsible for the content of the paper and the views expressed in it.

References

Boolos, G. 1998. *Logic, Logic, and Logic*, Cambridge, MA: Harvard University Press.
Burgess, J.P. 1998. 'Introduction' (to Part II), in Boolos 1998, 135–142.
Burgess, J.P. 2005. *Fixing Frege*, Princeton: Princeton University Press.
Church, A. 1944/1956. *Introduction to Mathematical Logic*, 2nd revised edition (1956). Princeton: Princeton University Press.
Feferman, S., Kreisel, G. and Orey, S. 1960. '1-consistency and faithful interpretations', *Archiv für Mathematische Logik und Grundlagenforschung* **6**, 52–63.
Friedman, H. 1975. 'Some systems of second order arithmetic and their use', *Proceedings of the International Congress of Mathematicians* (Vancouver, B. C., 1974), Vol. 1, Montreal: Canad. Math. Congress, 235–242.
Gentzen, G. 1943. 'Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie'. *Mathematische Annalen* **119**, 140–161.
Girard, J.-Y. 1972. 'Une extension de l'intérpretation de Gödel à l'analyse et son application a l'élimination des coupures dans l'analyse et la théorie des types', in J.E. Fenstad (ed.) *Proceedings of the 2nd Scandinavian Logic Symposium*, Studies in Logic and the Foundations of Mathematics, vol. 63, North-Holland Publishing Company, Amsterdam, 1971, 63–92.
Hale, B. and Wright, C. (2001). *The Reason's Proper Study*. Oxford: Clarendon Press.
Henkin, L. 1953. 'Banishing the rule of substitution for functional variables', *Journal of Symbolic Logic* **18**, 201–208.
Hilbert, D. and Ackermann, W. 1928. *Grundzüge der theoretischen Logik*, Berlin: Springer.
Hilbert, D. and Bernays, P. 1934-39. *Grundlagen der Mathematik*, vol. 1 (1934), vol. 2 (1939), Berlin: Springer.

Kleene, S. 1952. *Introduction to Metamathematics*, Amsterdam: North-Holland.
Leivant, D. 1994. 'Higher-order logic', in D. Gabbay *et al*. (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 2, Oxford: Oxford University Press, 230–321.
Lindström, P. 1984. 'On faithful interpretability', in *Computation and Proof Theory*, Springer Lecture Notes in Mathematics **1104**, 279–288.
Lindström, P. 2003. *Aspects of Incompleteness* (Second edition), Lecture Notes in Logic, vol. 10. Association for Symbolic Logic, Urbana, IL: A K Peters, Ltd., Natick, MA.
Linsky, B. and Zalta E.N. 1995. 'Naturalized Platonism vs. Platonized Naturalism', *Journal of Philosophy* **XCII**, 525–555.
Linsky, B. and Zalta E.N. 2006. 'What is neologicism?', *Bulletin of Symbolic Logic* **12**, 60–99.
Manzano, M. 1996. *Extensions of First-order Logic*, Cambridge: Cambridge University Press.
MacBride, F. 2003. 'Speaking with shadows: A study of neo-logicism', *British Journal for the Philosophy of Science* **54**, 103–163.
Mummert, C. and Stephen S.G. 2005. 'Reverse mathematics and Π^1_2 comprehension', *Bulletin of Symbolic Logic* **11**, 526–533.
Parsons, C. 1970. 'On number choice schema and its relation to induction', in A. Kino *et al.* (eds.), *Intuitionism and Proof Theory*, Amsterdam: North-Holland, 459–473.
Prawitz, D. 1965. *Natural Deduction: A Proof-Theoretical Study*, Stockholm: Almqvist and Wiksell.
Prawitz, D., 1967. 'Completeness and Hauptsatz for second order logic', *Theoria*, **33**, 246–258.
Prawitz, D. 1971. 'Ideas and results in proof theory', in J. Fenstad, ed, *Proceedings of the Second Scandinavian Logic Symposium*, North-Holland, 235–308.
Rossberg, M. 2004. "First-order logic, second-order logic, and completeness', in Vincent Hendricks *et al.* (eds.), *First-Order Logic Revisited*, Berlin: Logos, 2004, 303–321.
Rossberg, M. 2006. *Second-Order Logic: Ontological and Epistemological Problems*, Doctoral Dissertation, the University of St Andrews.
Schmidt, A. 1938. 'Uber deduktive Theorien mit mehreren Sorten von Grunddingen', *Mathematische Annalen* **115**, 485–506.
Shapiro, S. 1991. *Foundations without Foundationalism: A Case for Second-Order Logic*, Oxford: Oxford University Press.
Simpson, S.G. 1999. *Subsystems of Second Order Arithmetic*, Heidelberg: Springer-Verlag.
Tait, W. 1966. 'A nonconstructive proof of Gentzen's Hauptsatz for second order predicate logic', *Bulletin of the American Mathematical Society* **72**, 980–983.
Takahashi, M. 1967. 'A proof of cut-elimination in simple type theory', *Journal of the Mathematical Society of Japan* **19**, 399–410.
Takeuti, G. 1953. 'On a generalized logic calculus', *Japanese Journal of Mathematics*, **23**, 39–96.
Takeuti, G. 1975. *Proof Theory*, North-Holland, Amsterdam.
Tarski, A. 1933/1935. 'Pojęcie prawdy w językach nauk dedukcyjnych' ('The concept of truth in the languages of the deductive sciences', in Polish), *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizycznych* **34** (1933), Warsaw. Expanded German translation: A. Tarski, 'Der Wahrheitsbegriff in den formalizierten Sprachen', *Studia Philosophica*, **1**, 261–405 (1935). English translation in A. Tarski, *Logic, Semantics, Metamathematics*, 2nd revised edition, Indianapolis: Hackett (1983), 152–278.
Tarski, A., Mostowski, A., and Robinson, R.M. 1953. *Undecidable Theories*, Amsterdam: North-Holland.
Troelstra, A. and Schwichtenberg, H. 1996. *Basic Proof Theory*, Cambridge: Cambridge University Press.
Van Dalen, D. 2004. *Logic and Structure*, 4th edition, Berlin: Springer.
Visser, A. 2005. 'Faith & falsity', *Annals of Pure and Applied Logic* **131**, 103–131.
Visser, A. 2011. 'Hume's principle, beginnings', *Review of Symbolic Logic* **4**, 114–29.
Wang, H. 1952. 'Logic of many-sorted theories', *Journal of Symbolic Logic* **17**, 105−116.
Wright, C. 2007. 'On quantifying into predicate position: steps towards a new(tralist) position', in M. Leng *et al.* (eds.) *Mathematical Knowledge*, Oxford: Oxford University Press, 150–174.