Natural numbers as exponents of operations

A formalization of Wittgenstein’s definition

Abstract

Wittgenstein’s definition of natural numbers as “exponents of operations” is investigated and formalized. The definition gives different systems of “numbers”, which are essentially isomorphic, however. Mathematically this is rather simple and uninteresting; an empty formalization, so to say. But philosophically this might throw some light on the concept of number. Although the basic form of a numeral in human languages is usually a cardinal number, numerals like twice or Latin bis that answer the question “how many times?” might reflect better the essence of a number as an idea in human mind.

Preface

I wrote the original version of this text in 1977 and distributed it to various people but never really got any feedback. Either my idea was too trivial, or presented poorly, or it just wasn’t interesting. In 2001, I decided to publish a slightly edited version on the World Wide Web, hoping that someone finds it, and finds it interesting. (The most important change was to substitute the term functional for transformation.) I have learned to see the Web as a great medium. People with odd ideas and hobbies can share them.

As I mentioned in the abstract, this is a philosophical essay rather than a mathematical paper. But it uses mathematical methods, so you probably need to know at least elementary university level mathematics to understand this.

This is also an exercise in applying some of my ideas about authoring mathematical texts on the Web, Math in HTML (and CSS). I have used some techniques that produce better quality of the presentation of mathematical notations, at some reduction in accessibility. In order to read this so that the symbols are correctly displayed, you will need something like Internet Explorer or Netscape 4 or newer and a sufficiently rich font selected (Lucida Sans Unicode should do), and even then you might have problem. Sorry about this. There are some notes on browser settings in my document Using national and special characters in HTML.

Introduction

What is essential in the system of natural numbers?

In modern logic, natural numbers are most often defined as finite ordinals or cardinals. A natural number is the cardinality (intuitively, number of elements) of a finite set. The basic operations like forming the successor of a number are introduced after the definition of numbers. There is also the axiomatic approach, which defines some properties for numbers (Peano axioms) instead of any definition for what the numbers really are. The cardinality approach can be proved to satisfy the axioms.

But is either of the two approaches intuitively satisfactory? Consider how we establish whether two collections of physical object (two sets in the everyday sense) have the same number of elements. If the collections are small, we just look at them. It is of course most likely that our mental processes in this connection involve an unconscious matching procedure which corresponds to the mathematical notion of bijective mapping. But no mathematics in the normal sense is needed. In fact, there are several animals that can learn to recognize that two collections "have the same cardinality". For large collections, we need some conscious and explicit procedure.

The cardinality approach offers one: try to find a bijection between the sets, perhaps physically by drawing lines from an element in one set to an element in the other set. But is this really the approach we apply in practice?

"Large" is here a fuzzy attribute that means 'having more than about seven elements'. It is noteworthy that seven is a holy number in several religions and cults. The most natural way to decompose it—in the sense of visualizing a set of seven elements—is to consider 7 as the sum of 3 and 4, which also have the status of holy numbers. However, the septimal (base 7) number system has hardly ever been used, except in the sense that a week consists of seven days. On the other hand, if seven is still an "intuitively visualizable" number, then we might expect an octal (base 8) system. But such systems, though very common in the modern world, seem to be a very new invention, historically speaking.

If I have a basket of apples and a basket of oranges and I want to see whether the number of apples equals the number of oranges, I can simply pick out one apple and one orange at a time. Picking out a pair can be interpreted as forming one element of a relation between the two sets.

To pick out elements from a collection inevitably means picking them in some order, even if it is randomly chosen. It is tantamount to counting them, except that we don't say the numerals. Couldn't we thus say that the system of natural numbers is not just an abstraction from real instances of comparing the cardinalities of sets but also an idea, a mental construction, and a corresponding procedure that is present in each comparison, explicitly or implicitly? The comparison is carried out by a human being as process that involves some variant of the idea of natural numbers, at least in the general sense of iterating over a finite set.

Apparently, the operations of addition and multiplication should have a natural interpretation in any theory that tries to lay the foundations of arithmetic. In the cardinality theory, the interpretations are "to form the union of disjoint sets" and "to form the Cartesian product", but they are not very enlightening. If I decide to count the number of elements in a collection and for that purpose divide it into two subsets, then count the elements in each of them, the result is formed by putting the two counts (enumerations) together, not by forming the union of the two subsets. The point is that although the union concept is relevant, it describes one look at the purpose of arithmetic, not arithmetic itself, or counting and operating on the results of counting. And if I have counted the number of elements in collection terni, three at a time, the result is formed by putting this enumeration and the enumeration 1, 2, 3 together—by composing them, in the sense of composition of functions. The Cartesian product idea doesn't really reflect the use of multiplication.

Calculation with natural numbers does not start with 0 or 1, or even with 2, but around 7. Of course, this note by itself is not intended to answer the question whether 0 or 1 (or some other number!) is the smallest natural number. When we switch to counting, we need to start from zero. But we wouldn't really need arithmetic if we only worked with small sets. It is the practical meaning of large sets (in the physical world) that makes the counting system

0, 1, 2,...

and all of arithmetic proper necessary. We can see that 2+2=4 (think about this as II+II=IIII) but we need to calculate what 197+456 makes.

Wittgenstein's definition

Ludwig Wittgenstein defined, in his famous Tractatus logico-philosophicus, in clause 6.0, (natural) numbers as "exponents of operations". This was related to his treatment of truth functions and had an anti-metaphysical purpose. Wittgenstein's thesis can, however, be studied as an idea concerning the logical foundations of mathematics, in the sense of symbolic rather than philosophical logic.

Wittgenstein considered the theory of classes (sets) as redundant in mathematics. It seems that he deduced this from considerations concerning natural numbers and that his objection was in fact directed towards the “static” set-theoretical notion of numbers.

In any case, Wittgenstein did not explicitly develop further his idea of the system of natural numbers, although this system played an important rôle in his metamathematical investigations. However, he was interested in interpreting mathematics as a family of calculi, and in this connection he rhetorically asks, as posthumously published in Remarks on the Foundations of Mathematics:

> Tell me: have I discovered a new kind of calculation if, having once learnt to multiply, I am struck by multiplications with all the factors the same, as a special branch of these calculations, and so I introduce the notation ‘$a^n = \ldots$’?
> Obviously, the mere ‘shortened’, or different, notation—‘16^2’ instead of ‘16×16’—does not amount to that. What is important is that we now merely count the factors.

In the original German text, there is the word “andere”, ‘another’, in place of “another”.

Note that in the above-quoted treatment, natural numbers were used not only for illustration as data items but also in the construct ‘a^n’ as exponents.

In my presentation, I will use set-theoretical notions. I assume that we have a set of objects, but the definitions will be essentially based on operations (mappings, functions) in the set of objects and on operations on such operations, or functionals. In a sense, set theory is present, and I do not pretend that Wittgenstein would be satisfied.

“An exponent of an operation” expresses how many times the operation is to be executed, in the sense of being applied repeatedly. This can be interpreted so that

- f^0 = identical operation (i, defined by $(\forall x)(i(x)=x)$
- f^1 = f itself
- $f^2 = f \bullet f$ (f composed with itself)
- $f^3 = f \bullet f \bullet f$
- ...

Whatever we do with natural numbers ultimately relates to properties of exponentiation of operations. Thus '15' refers to 'f^{15}', where f is a generic symbol for operations.

Theory of X-numbers

Let X be a nonempty set. Its elements will be called objects. We shall consider the set F consisting of all mappings of X into itself, i.e. from functions from X to X, called operations in X; naturally F depends on X, so we denote it by F_X when needed.

We also consider mappings of F into itself and we shall call them functionals over X and denote their set by T, or T_X for clarity. Thus, pragmatically speaking, an element of T is a function that takes a function (in X) as its argument and returns a function (in X) as its value. For example, there's the trivial identity functional that maps each function into itself.

We define the composition of operations in the usual way: if $f,g \in F$, then the composite of f and g is the operation $f \bullet g$ defined by

$$(f \bullet g)(x) = f(g(x)) \text{ for all objects } x \in X$$

Definition 1. The functional $0 \in T$ such that

$$0(f) = i \text{ for every operation } f \in F,$$

where i denotes the identity operation (i.e., the identical mapping of X into itself), will be called the X-number zero.

Comment. Here the attribute "X-number" only refers to the fact that distinct sets of objects have distinct zeros. The functional 0 is simply the constant functional that "annuls" every operation by mapping it to the identity operator.

Definition 2. If τ is a functional, its successor is the functional τ^+ defined by

$$\tau^+(f)=f\bullet\tau(f) \text{ for all } f\in F$$

Comment. Denoting the identical functional by ι, (thus having $\iota(f)=f$ for all $f\in F$), we notice that

$$\tau^+(f)=\iota(f)\bullet\tau(f)$$

Definition 3. A set $P\subset T$ is closed under succession, if it contains the successors of its elements, i.e.

$$(\forall\pi\in P)(\pi^+\in P)$$

Definition 4. The smallest subset of T_X that contains 0 and is closed under succession is called the set of X-numbers and denoted by $\bullet_X$.

More explicitly, $\bullet_X$ is the intersection of all subsets of T having the two above-mentioned properties. It is clear that the intersection also has those properties. Moreover, there are such subsets, at least T itself. In general, the properties of $\bullet_X$ depend on X, and only by setting restrictions on X can we interpret X-numbers as natural numbers.

It is easy to see that a functional is an X-number if and only if it is either 0 or the successor of some X-number.

Before a detailed investigation of the structure of $\bullet_X$, let us prove a simple assertion that tells us whether there is something to be investigated.

Proposition 1. If X has at least two elements, i.e. if

$$(\exists x,y \in X)(x \neq y),$$

then each X-number is distinct from its successor.

Proof. Let x and y be two distinct elements of X. Define

$$f:X \rightarrow X,\ f(x)=y \text{ and } (\forall z \neq x)(f(z)=x).$$

Let us consider an arbitrary X-number σ and denote the object $(\sigma(f))(x)$ by q. Then

$$(\sigma^+(f))(x) = (f \bullet \sigma(f))(x) = f((\sigma(f))(x)) = f(q),$$

and $f(q) \neq q$ by the definition of f. Thus we have shown that $\sigma^+(f) \neq \sigma(f)$, which implies that $\sigma^+ \neq \sigma$. Q.E.D.

Next we consider the Peano axioms, which we state as (true or false) propositions concerning our concepts "X-number" and "successor":

PA1. The X-number zero is not a successor of any X-number:

$$(\forall \sigma \in \bullet_X)(\sigma^+ \neq 0).$$

PA2. Distinct X-numbers have distinct successors:

$$(\forall \sigma,\tau \in \bullet_X)(\sigma \neq \tau \Rightarrow \sigma^+ \neq \tau^+).$$

PA3. If a set S of X-numbers contains the X-number zero and successors of all X-numbers belonging to S, then S is the set of all X-numbers:

$$(\forall S \subset \bullet_X)(0 \in \bullet_X \ \&\ (\forall \sigma \in \bullet_X)(\sigma \in S \Rightarrow \sigma^+ \in S) \Rightarrow S = \bullet_X).$$

PA3 or the principle of mathematical induction is an immediate consequence (or an essential part) of the very definition of X-numbers.

It is easy to see that PA1 is false if X is a singleton, since in this case F={i}, and consequently the only element of T (and a fortiori the only X-number) is the transformation ι which maps every element of F (namely i) to i; thus $\bullet_X = \{\iota\} = \{0\}$, which implies that 0 must be its own successor.

Proposition 2. If X has more than one element, i.e. if $(\exists x,y \in X)(x \neq y)$, then PA1 is true.

Proof. Let x and y be distinct objects. Define the constant operation that maps every object to x:

$$f: X \to X,\ f(u) = x \text{ for all } u \in X.$$

It is easy to show by induction that

$$\sigma^+(f)(y) = x \text{ for all } \sigma \in \bullet_X.$$

And this implies that $\sigma^+ \neq 0$ for all $\sigma \in \bullet_X$. Q.E.D.

The question whether PA2 is true is the crucial point in comparing the system of X-numbers with the system of natural numbers defined by the Peano axioms. The following proposition says that PA2, which can be interpreted as an axiom of infinity for $\bullet_X$, is equivalent to the infinity of the object set X. (Thus we can informally say that only by having an infinity of objects at our disposal can we create an infinite sequence of different functionals by succession.)

The trivial case where X is a singleton is again easily studied. In this case there is only one X-number, and consequently PA2 is true.

Let us explicitly state the implications of PA1–PA2 on the mapping properties of the succession mapping

$$\Sigma: \bullet_X \to \bullet_X,\ \Sigma(\sigma) = \sigma^+ \text{ for all } \sigma \in \bullet_X.$$

(Note that here we consider a functional on functionals.)

PA3 says that Σ is injective. On the other hand, PA3 (or, more directly, definition 4) implies that $\Sigma(\bullet_X) \cup \{0\} = \bullet_X$, and if PA1 is true, the range of Σ is $\{\sigma \in \bullet_X \mid \sigma \neq 0\}$, which will be denoted by $\bullet_X^*$ in the sequel.

Proposition 3. Assuming that X has more than one element, PA2 is true if and only if X is infinite.

Proof. First, let us assume that PA2 is true. Reformulating the remark above, this means that the mapping

$$\Sigma^*: \bullet_X \to \bullet_X^*, \Sigma^*(\sigma) = \Sigma(\sigma) = \sigma^+ \text{ for all } \sigma \in \bullet_X,$$

is bijective. Since its range is a proper subset of $\bullet_X$, this implies that $\bullet_X$ is infinite. Thus X must be infinite, too, since if it were finite, so would be F and a fortiori $\bullet_X$.

Next we assume that X is infinite. Let σ and τ be distinct X-numbers. This means that there is an operation f such that

$$\sigma(f) \neq \tau(f),$$

which in turn means that there is an object x such that

$$\sigma(f)(x) \neq \tau(f)(x).$$

Consider the following subset of X:

$$S = \{\nu(f)(x) \mid \nu \in \bullet_X\}.$$

Now $S \neq X$ or $S = X$. In the first case there is an object z that does not belong to S. Define the operation g by setting $g(z)=x$ and $g(y)=f(x)$ for all $y \neq z$. Then it is easy to show by induction that

$$\sigma^+(g)(z) = \sigma(f)(x) \text{ for all X-numbers } \sigma.$$

The induction proof goes as follows: We immediately obtain $0^+(g)(z) = g(z) = x$ and $0(f)(x) = i(x) = x$. Now supposing that $\sigma^+(g)(z) = \sigma(f)(x)$ we calculate

$$\sigma^{++}(g)(z) = g(\sigma^+(g))(z) = g(\sigma(f))(x),$$

which is equal to $f(\sigma(f))(x) = \sigma^+(f)(x)$, since $\sigma(f)(x)$ belongs to S and is thus distinct from z.

This implies $\sigma^+(g)(z) \neq \tau^+(g)(z)$, which implies $\sigma^+ \neq \tau^+$.

It remains to consider the other alternative $S = X$. Since X is infinite, there is a bijection b from X to its proper subset, say Y. Now define the operation h by setting $h(w) = b(x)$ if $w \notin Y$ and $h(w) = (b \bullet f \bullet b^{-1})(x)$ if $w \bullet Y$ (where b^{-1} is the inverse of b). Let u be an object not belonging to Y. One can show by induction that

$$\nu^+(h)(u) = b(\nu(f)(x)) \text{ for all } X\text{-numbers } \nu.$$

Induction proof: $0^+(h)(u) = h(u) = b(x)$ and $b(0(f)(x)) = b(i(x)) = b(x)$. If $\nu^+(h)(u) = b(\nu(f)(x))$, then

$$\nu^{++}(h)(u) = h(\nu^+(h)(u)) = h(b(\nu(f)(x))),$$

and since in the last expression the argument of h belongs to $Y=b(X)$, this expression is equal to

$$(b \bullet f \bullet b^{-1})(b(\nu(f)(x))) = (b \bullet f \bullet b^{-1} \bullet b \bullet \nu(f))(x) = (b \bullet f \bullet \nu(f))(x) = b(\nu^+(f)(x)).$$

Because b is injective, we deduce that

$$\sigma^+(h)(u) \neq \tau^+(h)(u),$$

whence $\sigma^+ \neq \tau^+$. Q.E.D.

We defined (in definition 2) the successor of a functional by

$$\sigma^+(f) = f \bullet \sigma(f).$$

It can be asked why we wrote $f \bullet \sigma(f)$ instead of $\sigma(f) \bullet f$. The following proposition says that the two definitions would be equivalent.

Proposition 4. For all X-numbers σ and for all operations f,

$$\sigma^+(f) = \sigma(f) \bullet f.$$

Proof. We use induction; for illustration, we apply the mode of inference that is directly justified by PA3. Let S be the set

$$S = \{ \sigma \in \bullet_X \mid (\forall f \in F)(\sigma^+(f) = \sigma(f) \bullet f) \}.$$

Now $0 \in S$, since for all $f \in F$,

$$0^+(f) = f \bullet 0(f) = f \bullet i = f = i \bullet f = 0(f) \bullet f.$$

And when σ is any element of S, then we obtain for all $f \in F$

$$\sigma^{++}(f) = f \bullet \sigma^+(f) = f \bullet (\sigma(f) \bullet f) = (f \bullet \sigma(f)) \bullet f = \sigma^+(f)) \bullet f,$$

and thus $\sigma^+ \in S$. By PA3, we deduce that $S = \bullet_X$. Q.E.D.

The use of X-numbers

If we confine ourselves to infinite sets of objects, then all systems of X-numbers have the same structure in the sense that they all satisfy the Peano axioms. It is noteworthy—although intuitively clear—that the “degree of infinity” of X, i.e. the transfinite cardinal number that describes the cardinality of X, cannot have an effect on the structure of X-numbers.

Adopting the classical theory of axiomatic arithmetic, we could directly apply its defined concepts and proved assertions to any system of X-numbers, provided that X is infinite. (Evidently, some results apply to the case of finite X, too.) It is, however, possible to give other definitions (and consequently different proofs) in some cases. For example, addition and multiplication can be defined as compositions at two levels:

$$(\sigma+\tau)(f) = \sigma(f) \bullet \tau(t),$$

$$(\sigma\times\tau)(f) = (\sigma \bullet \tau)(t).$$

These definitions should appeal human intuition better than the conventional ones based on recursion.

For illustration, let us consider the formula (proposition)

$$2\times2 = 4.$$

In our system, this says exactly that

$$(\forall f \in X)((2\times2)(f) = 4(f)$$

or, using a slightly different notation and using our definition of multiplication,

$$(\forall f \in X)((f^2)^2 = f^4).$$

The formula can be regarded as a general proposition which valid no matter what underlying set of objects X is assumed. The proof does not in fact use symbols referring to objects at all.

We of course imply the standard definitions (notations) $1 = 0^+$, $2 = 1^+$, $3 = 2^+$ etc. Our proof of the formula is quite similar to the one in Tractatus. For all operations f we have

$$(2 \times 2)(f) = (2 \bullet 2)(f) = 2(2(f)) = 1^+(2(f)) = 2(f) \bullet 1(2(f)) = 2(f) \bullet 0^+(2(f)) =$$
$$2(f) \bullet 2(f) = 1^+(f) \bullet 2(f) = (f \bullet 1(f)) \bullet 2(f) = f \bullet (f \bullet 2(f)) = f \bullet 2^+(f) = f \bullet 3(f) = 4(f).$$

(Cf. to Wittgenstein's proof in Tractatus 6.241.) We wrote the proof with almost all details, in order to show that it consists of the use of definitions and the associativity of composition of operations. This structure enlightens Wittgenstein's thesis that it is a property of expressions like '1+1+1+1' that they can be understood e.g. in the form '(1+1)+(1+1)'. And this is based on the associativity of •, which makes the use of parentheses free.

The generality of our definition of X-numbers gives a uniform setting for the introduction of powers of operations and similar concepts where natural numbers are traditionally used as exponents in some sense.

Example 1. Let the object set X consist of all real-valued partial functions of one real variable, i.e. mappings from some subset of • to •. (Since the empty set is a subset of any set, a nowhere-defined real function is an element of X, too.) For any $x \in X$, we define its derivative by the usual formula

$$D(x)(t) = \lim \; \frac{x(s) - x(t)}{}$$

$s \to t \qquad s - t$

with the convention that if the limit does not exist as a finite real number, D(x) is regarded as undefined for the corresponding value t of the argument. Now we can consider the operation $D:X \to X$. The general theory of X-numbers automatically defines derivatives of all higher orders. We might just specify a preferred notation for them, such as the conventional D^n for n(D).

Example 2. Let G be a multiplicative semigroup with the neutral element e. (Multiplicative means here just that we call the operation of the group multiplication and omit (imply) the operation.) For all $g \in G$ we can define the powers of g as follows. Consider the operation of left multiplication by g:

$$k : G \to G,\ k(x) = gx \text{ for all } x \in G.$$

For any G-number n, we define the n^{th} power of g as the element n(k)(e). It can easily be seen that the verification of the formula

$$g^{m+n} = g^m g^n,$$

that is,

$$(m+n)(k)(e) = m(k)(e)n(k)(e),$$

can be reduced to the problem of showing that k possesses the following property:

$$(\forall n \in \bullet_G)(\forall x \in G)n(k)(x) = n(k)(e)(x)).$$

Example 3. The so-called fixed-point iteration or simple iteration is an algorithm that can, in one form, be described as follows: given a number x_0 and a function f, generate the sequence

$$x_1=f(x_0),\ x_2=f(x_1),\ x_3=f(x_2),\ \ldots$$

Using the concepts discussed here, the algorithm can be expressed compactly as follows:

$$\text{generate } n(f)(x_0) \text{ for } n=1, 2, \ldots$$

Remarks on the nature of arithmetic

It is quite possible—both theoretically and practically—to dispense with the notion of natural numbers. We might just discuss systems of X-numbers and their general properties (i.e., arithmetic) only. It is basically just a linguistic convention to speak of natural numbers and their set •, instead of (isomorphic) sets of X-numbers ($•_X$), though admittedly it also implies some notational convenience.

Arithmetic can be regarded as a formal system, or even a formal language, having its own criteria for truth. But if we consider the practical use of arithmetic concepts, propositions, and calculations, involving concrete numbers or variables, we can see a wide variety of activities which require that arithmetic be related to something external to it. We frequently use concepts like "the n^{th} derivative of a function", "the n^{th} iterated kernel of an integral equation", "recurrence relations", "applying a formula n times", etc. This means using natural numbers in a manner which is related to something external to arithmetic, yet internal to mathematics.

This does not compel us think that natural numbers "exist" in some absolute (metaphysical) sense or as aprioric concepts in human mind. In fact, why would we? We do use natural numbers, and it is a common belief that a thing that is used necessarily exists. But there is a difference between the use of a hammer and the use of numbers. There is also a fundamental difference between the use of a hammer and the use of symbols as e.g., glyphs on paper, although both are material.

In The Blue and Brown Books, composed of Wittgenstein's presentations, the problem is presented as follows:

> — — the explanation that a number is the same thing as a numeral satisfies the first craving for a definition. And it is very difficult not to ask: 'Well, if it isn't the numeral, what is it?'

Numerals can, indeed, play the rôle of numbers. Consider the alphabet A = {|}, with the vertical line (stroke) denoting itself (as a character). Within the general framework of the theory of formal languages, consider the set A^* of all (finite) sequences of the vertical line. The empty string, conventionally symbolized by Λ, is included. We could define that its elements are natural numbers and develop arithmetic on this basis. This has actually been done by •. •. •••••• in •••••• •• ••••••••••••••• •••• ••• (Moscow, 1973). In the X-numbers view, the corresponding method would be to define the function

$$s:A^*\rightarrow A^*,\ s(x) = x|\ \text{for all}\ x\in A^*,$$

where x| means the concatenation of x and |. Then we could note that any A^*-number α can be uniquely denoted by $\alpha(s)(\Lambda)$, which is α copies of | in succession, i.e. a numeral in unary notation. Note that the unary notation is still used in roman numerals like III, and it is assumably the oldest system of written numerals.

However, A^*-numbers are per se something specific and dependent on A, whereas natural numbers have an ubiquous character. The point of view that I suggest is that symbolic numerals (unary, decimal, or other) are general notations that can be used in connection with any set of objects (when expressing cardinalities or ordering positions the traditional way) or in connection with any set of X-numbers. For the most of it, the "X-" prefix is as irrelevant as the question whether we count apples or oranges.

Outside the realms of operations and their powers, arithmetic is only a game, or just formal manipulation. This might sound extremist, but in effect, as we have partly seen and partly will see shortly in the Conclusion section, the usual ideas of using numbers for counting and ordering can be seen as special applications of the idea of exponents of operations.

We could translate the ontological question "Do natural numbers exist?" to an operational, practical question "Is it possible to study the general properties of exponentiation of operations?” The person who asked the question might object to the translation, but in any case, the latter question can be answered. Note that such translations can be applied to seemingly less metaphysical questions too: “Is there a number n such that 3+n=5?” can be translated to “Implying an arbitrary infinite set of objects X, is it possible to find a functional $n\in \bullet_X$ such that $(f\in F_X)(3(f)\bullet n(f) = 5(f))$?”

If the assertion “There is only one zero” is taken as something else (something more general) than “In any system of X-numbers, the X-number zero is unique”, then it is false. (Taking the zeroth derivative of a function is similar to but not the same as raising a number to the zeroth power. In both cases, it is a matter of not applying an operation—differentiation or multiplication—at all.)

In the beginning I said that one can see that 2+2=4. In the theory of cardinals one may say that it is seen by looking at the attached scheme, since it is intuitively clear that the union of A and B can be bijectively mapped onto C. (The natural intuition does not formulate it this way, of course.) There is, however, something arbitrary in this visualization; it is more graphic than mathematical or logical. But one can see the validity of 2+2=4 by looking at the formula

[A scheme with three areas A, B, C, with A and B containing two items each and C containing four items.]

$$f^{2+2} = f^4, \text{ i.e. } (ff)(ff) = ffff.$$

Suppose that we wish to compute a power, say f^{12}, of a function f (e.g., a polynomial), not in the sense of repeated arithmetic multiplication but in the sense of applying the function 12 times, i.e. a power with respect to composition. Perhaps it would be more efficient to compute first a general formula for $g=f^3$ and then calculate g^4. Obviously the arithmetic truth $3\times 4 = 12$ suggests this method, but does it justify it logically. If we regard the arithmetic formula as a proposition of “pure”, axiomatic mathematics, with “zero” and “successor” just as postulated abstract concepts, the formula does not justify anything. Thus it would be necessary, in principle, to perform a tedious calculation in order to verify that $(f^3)^4 = f^{12}$, i.e.

$$(f \bullet f \bullet f) \bullet (f \bullet f \bullet f) \bullet (f \bullet f \bullet f) \bullet (f \bullet f \bullet f) = ffffffffffff,$$

to be rigorous. But in the theory of X-numbers the procedure is immediately justified by an instance of 3×4 = 12.

Conclusion

The theory of natural numbers is the foundation of all systems of numbers and in this sense in the core of mathematics. Thus, this theory should have a formalism which is satisfactory not only logically but also intuitively and philosophically. The formalization of Wittgenstein’s definition seems to give such a solid basis.

The formalization could be given in the more general framework of category theory. But sets and functions and functionals are an intuitively simple and well-known system.

The definition of natural numbers as “exponents of operations” in much better related to actual use of numbers than their definition as finite cardinal numbers. As regards to the relationship between the two definitions, the cardinal number of a finite set A could be defined within our formalism as follows. Let < be some total ordering of A and let s and g be respectively the smallest and largest element of A, with respect to <. Define

$$f: A \to A,\ f(a) = \begin{cases} \min \{x \in A \mid a < x\} & \text{if } a \neq g \\ s & \text{if } a = g \end{cases}$$

Then define card(A) as the unique A-number σ such that $\sigma(f)=i$, where i is the identical mapping of A. This formalizes the idea that the cardinal number of A is found by picking out one element at a time and counting how many times the "picking-out operation" is applied. We can visualize the operator f as follows:

[A sequence of elements of A, from s to g, with arrows labeled with f pointing from an element to the next one and from g back to s.]

(It is clear that card(A) defined that way is independent of the choice of the ordering <.) Thus, cardinality can be interpreted as a very special case of exponent.

To summarize the advantages of the formalization presented here, it

1. is intimately related to the general notion of iteration,
2. provides a unified setting for introducing repeated operations, so that e.g. the definition of a derivative implicitly defines the concept of higher derivatives,
3. make the structure of arithmetic more natural and easier to develop from the basic concepts, so that e.g. the associativity of multiplication of natural numbers is simply an instance of the general associativity of exponentiation of functions with respect to composition.

Date of creation: 2001-03-17. Last modified 2003-09-15.