AGORA —Papeles de Filosofía— (1995), 14/1: 149-156 *ISSN 0211-6642*

SEMANTIC ARITHMETIC: A PREFACE

John Corcoran

Abstract

Number theory, or pure arithmetic, concerns the natural numbers themselves, not the notation used, and in particular not the numerals. String theory, or pure syntax, concerns the numerals as strings of «uninterpreted» characters without regard to the numbers they may be used to denote. Number theory is purely arithmetic; string theory is purely syntactical... in so far as the universe of discourse alone is considered. Semantic arithmetic is a broad subject which begins when numerals are mentioned (not just used) and mentioned as names of numbers (not just as syntactic objects). Semantic arithmetic leads to many fascinating and surprising algorithms and decision procedures; it reveals in a vivid way the experiential import of mathematical propositions and the predictive power of mathematical knowledge; it provides an interesting perspective for philosophical, historical, and pedagogical studies of the growth of scientific knowledge and of the role metalinguistic discourse in scientific thought.

1. Pure Arithmetic

Pure arithmetic, or pure number theory, is about the natural numbers, or finite cardinals: zero, one, two, and so on. These numbers are often taken to be properties of finite sets. For example, the number two is taken to be a property that belongs to a given set if and only if that set has exactly two members.

One of the most fundamental principles of arithmetic is the principle of mathematical induction which is expressed using a sentence involving a numeral for zero.

PMI Every property that belongs to zero and to the successor of each number to which it belongs also belongs to every number.

The properties referred to in PMI are the arithmetic, or numerical, properties such as being odd, being even, being prime, being zero or a successor,

Aceptación: VI/1995

being distinct from its own successor, etc. The principle of mathematical induction gives a sufficient condition for a property to be universal, that is, numerically or arithmetically universal, that is, for a property to belong to each and every number without exception. The property of being distinct from its own successor is arithmetically universal. In fact, every universal arithmetic proposition amounts to a proposition to the effect that a certain arithmetic property is universal.

In order for a given property to be arithmetic it is necessary and sufficient for that property to be coherently predicable (that is, truly or falsely predicable) of each and every number. The property of being happy, the property of being intelligent, the property of being female, and the property of being three characters in length, are of course not arithmetic. It would be incoherent to say that a number is happy, intelligent, female, or three characters in length.

Although many arithmetic propositions are expressed using sentences which employ numerals, or number names, there are many arithmetic propositions which are normally expressed using numeral-free sentences. The principle of joint multiples is an especially apt example because this principle is used in proofs of some of the most beautiful elementary results of semantic arithmetic.

PJM Every two numbers whose sum or difference is a multiple of a given number are either both multiples of that given number or both not multiples of that given number.

This is of course logically equivalent to the proposition that if the sum or the difference of two numbers is a multiple of a given number then in order for one of them to be a multiple of the given number it is necessary for the other to also be a multiple of that given number.

It is important to notice that the word «two» occurs in the above expression of PJM as an adjective, not as a proper name, and, *afortiori*, not as a proper name of a number. The whole expression «every two numbers» functions as a unit, like a double universal quantifier, «for every number x, for every number y....» Perhaps other examples of numeral-free sentences are in order. Consider the principle of infinitely many primes.

PIP Every number is exceeded by a prime number.

Consider the arithmetic Pythagorean theorem.

APT No square number is the sum of a smaller square number with itself.

This is equivalent to the proposition that no square (number) is twice a smaller square number. The qualification «smaller» is needed because zero is the sum of zero with zero. The arithmetic Pythagorean theorem amounts, of course, to the theorem that the square root of two is not rational, that is, is not the quotient of a natural number with a natural number... but reference to the square root of two takes us beyond number theory (i.e. natural number theory) to the theory of the real numbers, which is also known as analysis.

The theory of natural numbers can be and has been formulated without numerals. We are so used to using numerals, actually arabic numerals, to discuss the natural numbers that we must remind ourselves that this useful notation is just that... notation. It is regrettable that the word «number» has become ambiguous and that in one sense it is synonymous with «numeral».

Modern formulations of number theory generally depend in one way or another on the work of Peano and Dedekind. One very convenient formulation, due to Gödel, is presented in an accessible way on page xli of the new edition of the Cohen-Nagel text listed as Cohen-Nagel (*1993*) in the bibliography.

2. Pure Syntax

Pure syntax, or pure string theory, is about strings of characters. For present purposes the most prominent characters are the ten arabic digits: '0', '1', '2', '3', '4', '5', '6', '7', '8' and '9'. Aside from the null string, which is zero characters in length, or which has zero character-length, every string is the result of concatenating one or more characters. If we limit our universe of strings to the arabic numerals then we can state a principle of string induction analogous to the principle of mathematical induction.

PSI Every property that belongs to the null string and to the concatenation of each digit with each string to which the property belongs also belongs to every string.

Concatenation is the most fundamental operation on strings; the concatenation of '456' with '123' is '456123' (and not '123456'... concatenation is not commutative like addition, but it *is*, of course, associative). The properties mentioned in the principle of string induction are syntactical properties such as being a digit, being a non-digit, being a palindrome (as '3223' and '32123'), or being distinct from a concatenation of a digit with itself.

In order for a given string to be palindromic it is necessary and sufficient for that string to «read the same backwards». The null string and the ten digits are palindromic, and the result of concatenating the same digit on both ends of a palindrome is again a palindrome. One sequence of palindromes is: null string, '00', '1001', '210012'. Another sequence of palindromes is: '0', '101', '21012', '3210123'. In order for a string to be periodic it is necessary and sufficient for it to be the result of repeated concatenation of some one string: '01', '0101', '010101', etc.

Syntactic (or string-theoretic) properties are coherently predicable of strings but not coherently predicable of any non-string. It is incoherent to say of a number that it is palindromic or that it is periodic (in the above sense). Likewise it is incoherent to say of a string that it is even, odd, prime, a multiple of nine, etc. The numbers form one category and the numerals (more generally, the strings) form another. Category mistakes result from predicating a property of an object outside of the range of applicability of that property.

The syntactic (or string-theoretic) operation of concatenation is one of the most fundamental operations that apply to strings and yield strings. Another syntactic operation is reversal: the reverse of a string is the result of «writing the string backward.» The null string and the ten digits are all reverses of themselves and if one given string is the reverse of a second then the concatenation of the first with a given digit is the reverse of the concatenation of the given digit with the second. The last sentence suggests a recursive definition of reversal in terms of concatenation, the digits and the null string. Once we have mentioned definition, it should be clear that the property of being palindromic is definable in terms of reversal: in order for a given string to be palindromic it is necessary and sufficient for it to be its own reverse.

Development and application of string theory has a very constructive and geometrical feel to it; it involves what has been aptly but metaphorically called «symbol manipulation.» Rudolf Carnap called pure syntax «the geometry of symbol shapes», a picturesque and suggestive phrase, Carnap *(1937)*.

There are two classical formulations of the theory of strings. One due to Alfred Tarski occurs as part of the famous truth-definition paper, Tarski *(1935)*. One due to Hans Hermes appeared about the same time. Both are discussed in detail in Corcoran, Frank, Maloney *(1974)* where it is shown that the two theories are definitionally equivalent in the sense that even though the two theories use different sets of primitive concepts (and thus different axiom sets) it nevertheless is the case that adding suitable definitions to one makes it possible to deduce the axioms and definitions of the other.

Discussions of string theories, number theories, and the varieties of induction principles that arise in them can be found in my paper «Categoricity», listed as Corcoran (*1980*) in the bibliography.

3. Semantic Arithmetic

Semantic arithmetic involves at a minimum the construction of one theory having two universes of discourse, e.g. the natural numbers and the arabic numerals, where the numerals are taken as names of the numbers. This integrated framework gives rise to a new class of arithmetic properties, e.g. being a «two-digit» number, and a new class of syntactic properties, e.g. being a «prime» numeral, i.e. being a numeral that denotes a prime number.

Perhaps the most obvious of the new syntactic relations is that of being coreferential: two strings are coreferential if and only if the two denote one and the same number. This extrinsically syntactic relation is coextensive with an intrinsically syntactic relation. '00123' is coreferential with '0123' and with '123': two numerals are coreferential if and only if they are both strings of ciphers ('0') or there is a string without initial ciphers from which each can be constructed by concatenation of zero or more initial ciphers, as '123' is a cipherless string from which '0123' and '00123' are constructed.

The distinction between extrinsic and intrinsic properties is a kind of hallmark of semantic arithmetic. An intrinsic arithmetic property is one that belongs (or does not belong) to a number in virtue of the nature of the number itself and/or in virtue of the place the number takes in the system of natural numbers. An intrinsic syntactic property is one that belongs (or does not belong) to a string in virtue of the nature of that string per se and/or in virtue of the place that string occupies in the system of strings. The arithmetic properties normally considered in pure arithmetic are intrinsic and the syntactic properties normally considered in pure syntax are intrinsic —but in neither case is there a limitation of properties to intrinsic ones. Extrinsic properties are those that belong (or not) to one sort of thing in virtue, not of that sort of thing itself, but rather in virtue of its relation to another sort. In the context of semantic arithmetic, an extrinsic arithmetic property belongs (or not) to a number in virtue of the numerals that denote it. The property of being denoted by a numeral ending in '0' is an extrinsic arithmetic property coextensive with the intrinsic arithmetic property of being a multiple of ten. Accordingly, an extrinsic syntactic property belongs (or not) to a string in virtue of the number it denotes. The property of denoting a multiple of nine is an extrinsic syntactic property that belongs to '0', '9', '18', '27', etc.

One elementary result in semantic arithmetic that dramatically illustrates the characteristic interplay of numbers and numerals, of referents and names, has been called Bolzano's corollary: the reverse of a numeral denoting a multiple of nine denotes a multiple of nine. Bolzano's corollary amounts to the proposition that the class of digit strings denoting multiples of nine is invariant under the syntactic transformation of reversal.

This easy but surprising result can be used to illustrate the experiential import of mathematical propositions: a person who knows a mathematical proposition can use that knowledge to predict the experiences that people will have. In fact, one might say that knowledge of mathematics can be used as a substitute for experience in so far as a person who has knowledge of a given mathematical proposition knows how certain experiments are going to come out before they are conducted... and if you know how the experiment will come out, what is the point of conducting it?

Ask a student to choose an arbitrary number, multiply it by nine and then take the reversal of the resulting numeral. You can then predict that if the student divides by nine a whole number will result. Or you can then predict that if the student repeatedly subtracts nine the ultimate result will be zero.

The reversal operation is an intrinsic syntactic operation that in no way involves regarding the strings of digits as numerals. To illustrate an extrinsic syntactic operation consider the repeated digit sum operation which applies to an arbitrarily long numeral. Add the digits and write the sum as a numeral. If the result is not a digit repeat the process. The ultimate result is called the repeated digit sum of the numeral. For example, the digit sum of '987654321' is '45' and the digit sum of '45' is '9'; so the repeated digit sum of '987654321' is '9'. Clearly this operation, which maps the class of all numerals onto the ten digits, is an extrinsic syntactic operation.

Results having to do with repeated digit sums were already known to Leonardo of Pisa, who is also called Fibonnaci, in the late middle ages (Ore, *1948*, 225 ff.). One of Fibonnaci's results is that the repeated digit sum of an arbitrary numeral denoting a positive multiple of nine is the digit '9'. Armed with this result you can predict that a student who chooses a non-zero number, multiplies by nine, and repeatedly adds the digits, will ultimately arrive at the digit '9'.

4. Conclusions

In order to illustrate the fact that Bolzano's corollary and Fibonnaci's result both depend on arabic decimal notation for the natural numbers it is sufficient to consider the arabic ternary notation in which each number is denoted by a string using no digits other than '0' '1' and '2'. The right position is the units position as before. The second position (formerly the tens position) is now the threes position. The third position (formerly the hundreds position) is now the nines position. The numeral '100' denotes nine; its reversal '001' denotes one, which is not a multiple of nine; its repeated digit sum is '1' not '9', of course.

According to many historians the emergence of the natural arabic decimal notation was connected with the discovery of zero and the discovery of positional or place-value notation. Some historians, e.g. Ore (*1948*, 16), think that the discovery of zero or at least the use of a «zero-symbol» to indicate a «void position» as in '204' was essential. However, this reasoning begs the question of whether a void position is necessary. Semantic arithmetic helps to free us from our dependence on the familiar notation so that we can answer the question of whether it was necessary to discover zero in order to develop a positional notation for the positive integers.

Consider the class of strings of the ten arabic digits. Now replace every occurrence of the cipher, or zero-symbol, by the letter 'T'. Thus '0' becomes 'T', '10' becomes '1T', '90' becomes '9T', etc. Now interpret these new strings just like in natural arabic decimal notation except take 'T' to denote ten. There is no longer a name for zero. The natural arabic name for ten is no longer available but we have the new name 'T'. The natural arabic name for one hundred is no longer available but we have the new name '9T', which shows ninety in the tens position and ten in the units position... a total of one hundred. The upshot is that this new positive arabic decimal notation provides a name for each and every positive integer. Moreover, it is more efficient in that no two strings of digits denote the same number; in positive arabic decimal notation every numeral is coreferential only with itself, no two distinct numerals are coreferential. To repeat, in order to have a positional notation for the positive integers it is not necessary to have «void positions» and it is not necessary to have a symbol for zero. This conclusion, which contradicts many published sources, is one that I have not seen in print before.

It is not necessary to confine semantic attention to arithmetic. If we expand the numerical universe of discourse to include the class of real numbers then we can speak of semantic analysis. In this expanded field the property of being algebraic (which belongs to a real number that is a solution to an algebraic equation) is seen to be an extrinsic analytic property, a property coherently predicable of real numbers in virtue of their relationship to certain strings of characters, viz. algebraic equations.

This brings us to the observation that even in semantic arithmetic it is not necessary to confine ones syntactic attention to numerals. If we consider sentences, for example constant equations, then we find that truth (in the Tarskian sense not previously used in this paper) is an extrinsic syntactic property.

In order for a constant equation to be true it is necessary and sufficient for the left term to denote the same number that the right term denotes. Thus «truth», in this derivative sense, is a syntactic property, a property coherently predicable of strings, but which belongs to a string not in virtue of its nature as a string but in virtue of its relation to what it is taken to be about. Even the syntactic properties of being a sentence and being an equation are extrinsic. Some people who are baffled about how the property of being a sentence can be syntactic would not be baffled if they were to distinguish extrinsic from intrinsic syntactic properties.

Carnap was confused about this point. He emphasized that truth, even analyticity, is a syntactic property but he failed to distinguish between intrinsic syntactic properties and extrinsic syntactic properties. It was Tarski's famous truth-definition paper, Tarski *(1935)*, that made it clear that truth is an extrinsic syntactic property. Even today we find people emphasizing the fact that deductions are syntactic objects, and that the property of being a deduction is syntactic, without going on to say that the property of being a deduction is extrinsically syntactic.

In so far as a deduction is considered in itself in regard to intrinsically syntactic properties there is no way to understand how it could fulfill its deductive function of showing that a conclusion follows from premises. In many cases it is a confusing and misleading half-truth to refer to a property as syntactic without indicating that it is not intrinsically syntactic, without indicating that it belongs to its exemplifications not merely in virtue of their string-theoretic nature but rather in virtue of their connections to things outside of the universe of strings, in some cases to human beings.

Acknowledgements: This paper is based on lectures in May 1994 at the University of Santiago de Compostela, Spain, in the Symposium on Semantic Problems in Scientific Languages and at the Instituto Superior Tecnico, Lisbon, Portugal. Special thanks to José M. Sagüillo, Juan Vázquez and Luis Villegas Forero of Santiago and to Joâo Martins and Ernesto Morgado of Lisbon. I am also grateful to my Department Chairs, Peter H. Hare and John T. Kearns, and to my Dean, Ross D. MacKinnon, for making arrangements to have me released from other duties so that I could devote time to research and writing. In the final stages of preparation of this paper I was assisted by Ky Herreid, Sriram Nambiar and José M. Sagüillo.

Bibliography

Carnap, Rudolf (1937): *The Logical Syntax of Language*, London.

Cohen, Morris and Nagel, Ernest (1993): *Introduction to Logic,* Indianapolis.

Corcoran, John *et al.* (1974): «String Theory», *Journal of Symbolic Logic* 39: 625-637.

Corcoran, John (1980): «Categoricity», *History and Philosophy of Logic* 1: 187-207.

Ore, Oystein (1948): *Number Theory and Its History*, New York.

Tarski, Alfred (1935): «The Concept of Truth in Formalized Languages», in Tarski 1983.

Tarski, Alfred (1983): *Logic, Semantics, Metamathematics*, second edition, Indianapolis.

John CORCORAN
State University of New York