FORMAL LOGIC

Yiannis N. Moschovakis
Department of Mathematics,University of California, Los Angeles, CA, USA and University of Athens, Greece

Keywords: algorithm, analytical hierarchy, arithmetical hierarchy, Axiom of Choice, axiomatizable theory, cardinal arithmetic, Church-Turing Thesis, complete theory, computable function, conjunctive normal form, consistent theory, countable (enumerable) set, disjunctive normal form, decidable relation, Deduction Theorem, diagonal argument, elementary equivalence, equinumerous, First Order Logic, formal sentence, formal syntax, formal theorem, Formalism, General Comprehension Principle, Generalized Continuum Hypothesis, Gödel Completeness Theorem, Gödel Incompleteness Theorem, Gödel Second Incompleteness Theorem, grounded set, inconsistent theory, Intuitionism, Kuratowski operation, logical axiom, logical consequence, logical truth, Logicism, mathematical logic, metatheorem, Model Theory, Modus Ponens, non-standard model, numeralwise expressibility, ordinal arithmetic, ordinal recursion, partial function, Peano Arithmetic, prenex normal form, primitive recursive function, Principle of Extensionality, programming language, proof by induction, Proof Theory, proofs-as-types, propositional connective, quantifier, Realism, recursive definition, recursive function, recursive program, Russell paradox, satisfaction relation, Scott-continuous, Second Order Language, semantics, set theory, symbolic computation, Tarski Truth Conditions, tautology, transfinite induction, transfinite recursion, Turing machine, Typed λ-Calculus, uncountable set, undecidable relation, unsolvable problem, wellorder, Zermelo-Fraenkel Set Theory.

Contents

1. Cantor’s Set Theory
1.1. Equinumerosity; Countable and Uncountable Sets
1.2. Cardinal Arithmetic
1.3. Transfinite Induction and Recursion
1.4. The Situation in 1900
2. The Birth of First Order Logic
2.1. Frege’s Logicism Program
2.2. The First Order Language of Sets
2.3. Logical Deduction
2.4. Frege’s Set Theory
2.5. Frege’s Definition of Number
3. The Paradoxes
3.1. Intuitionism
3.2. Type Theory
3.3. Hilbert’s Program
4. Axiomatic Set Theory
4.1. The Zermelo-Fraenkel Axioms
4.2. The Axiom of Choice
4.3. Von Neumann Ordinals and Cardinals
4.4. The Cumulative Hierarchy of Sets

4.5. The Emancipation of Logic from Set Theory
5. Mathematical Logic
5.1. Propositional Logic, PL
5.2. The Syntax of First Order Logic, FOL
5.3. FOL -Semantics
5.4. First Order (Elementary) Definability
5.5. Model Theory and Non-Standard Models
5.6. FOL -Deduction
5.7. Soundness and Completeness of FOL
5.8. Second Order Logic, FOL^2
5.9. The Typed λ-Calculus, L_λ
6. Gödel's First Incompleteness Theorem
6.1. The Incompleteness of Peano Arithmetic
6.2. Outline of Gödel's Proof
7. Computability and Unsolvability
7.1. Turing Machines
7.2. The Church-Turing Thesis
7.3. Unsolvable Problems and Undecidable Theories
7.4. Gödel's Second Incompleteness Theorem
8. Recursion and Computation
8.1. Recursive Programs
8.2. Programming Languages
Glossary
Biographical Sketch

Summary

Narrowly construed, Formal Logic is the study of *definition* and *deduction* (proof) in mathematical models of fragments of language. Its development has been intimately connected with that of *Set Theory*, the *Philosophy of Language* and more recently *Theoretical Computer Science* and *Linguistics*, so much so that parts of these fields are normally included in a broad conception of logic. It also has substantial applications to other areas of pure mathematics, especially through the work of Kurt Gödel.

Modern logic started in the last quarter of the 19th century, with Georg Cantor's discovery of set theory and Gottlob Frege's attempt to create a "logical" foundation of mathematics. In the first half of this article, roughly Sections 1–3 and part of Section 4, we will give a brief, elementary account of the contributions of Cantor and Frege and the subsequent development of logic up to about 1908; not a proper history, as we will use modern terminology and oversimplify extensively, but an attempt to trace the roots of our subject. In the rest of the paper we describe in outline the developments in logic in the 20th century, especially the famous Completeness and Incompleteness Theorems of Kurt Gödel and some of what they led to. The pace is somewhat faster in this part, there are no proofs, and the more recent developments are just mentioned, with pointers to the six more advanced articles in this topic.

1. Cantor's Set Theory

According to Cantor, *by a set we are to understand any collection into a whole of definite and separate objects of our intuition or thought.*

Thus the basic relation of set theory is *membership*, which holds between a set and its *members* (or *elements*), and which we denote by "$\in$", the first letter of the Greek word for "is":

$$x \in A \Leftrightarrow \text{ the object } x \text{ is a member of the set } A.$$

If, for example,

$$\mathbb{N} = \text{the natural numbers} = \{0, 1, 2, \ldots\}$$
$$\mathbb{Z} = \text{the rational integers} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\},$$
$$\mathbb{Q} = \text{the (proper) fractions } \frac{x}{y} \text{ with } x, y \in \mathbb{Z}, y > 0,$$
$$\mathbb{R} = \text{the real numbers},$$

then $-3 \notin \mathbb{N}$, $-3 \in \mathbb{Z}$, $\sqrt{2} \notin \mathbb{Q}$, $\sqrt{2} \in \mathbb{R}$, etc. Somewhat peculiar is the *empty set* $\emptyset$ which has no members, so that $x \notin \emptyset$ for every object x.

A set is completely determined by its members, i.e., for all sets A, B,

$$A = B \Leftrightarrow (\text{for all } x)[x \in A \text{ if and only if } x \in B].$$

This is the *Principle of Extensionality*, and it implies, in particular, that there is only one empty set.

For any *definite* (unambiguous) property of objects $P(x)$, we define its *extension*

$$\{x \mid P(x)\} = \text{the set of all objects } x \text{ such that } P(x),$$

so that, for example, $\{x \mid x \text{ is an odd, natural number}\}$ is the set of odd numbers, and $\{x \mid x \neq x\} = \emptyset$. Notice that

$$\text{if } A = \{x \mid P(x)\}, \text{ then } x \in A \Leftrightarrow P(x).$$

We can use this basic *Comprehension Operation* to define the usual "Boolean" operations on sets, e.g.,

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

is the *union* of A and B, the set whose members are all the members of A and B put together "into a whole"; and if $A_0, A_1, \ldots$ is a sequence of sets indexed by the natural numbers, then the union of the sequence is specified by

$$\bigcup_{n\in\mathbb{N}} A_n = \{x \mid \text{for some } n \in \mathbb{N}, x \in A_n\}.$$

The binary and infinitary *intersection* operations $A \cap B$, $\bigcap_{n\in\mathbb{N}} A_n$ are defined similarly, and so is the *difference* $A \setminus B$, comprising all the objects which are in A but not in B, and the *Cartesian product*

$$A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\}, \qquad (1)$$

where (x, y) is the *ordered pair* of x and y.

A *subset* (part) of a set A is any set whose members all belong to A, in symbols,

$$A \subseteq B \Leftrightarrow (\text{for all } x)[\text{if } x \in A, \text{ then } x \in B];$$

and using the comprehension operation again, we can collect all the subsets of a set A into its *powerset*,

$$\mathcal{P}(A) = \{X \mid X \subseteq A\}.$$

1.1. Equinumerosity; Countable and Uncountable Sets

Next to membership and identity, the most fundamental relation between sets is *equinumerosity*, which holds between A and B when a one-to-one correspondence exists between the members of A and the members of B; in symbols, and using the usual, mathematical notation for *functions* with *domain* A and *range* B,

$$A \sim_c B \Leftrightarrow (\text{there exists a bijection})\ f : A \rightarrow B.$$

If A and B are finite, then, $A \sim_c B$ exactly when A and B have the same number of elements. This is what allows us to deduce that "there are just as many left shoes as there are right shoes in a shoe store" without actually counting either set, and it follows from the *Pigeonhole Principle*, the basic tool of counting: *no finite set is equinumerous with one of its proper subsets*. On the other hand, the function $f(n) = 2n$ establishes a one-to-one correspondence between the natural numbers and the even numbers, so that

$$\mathbb{N} \sim_c E = \{0, 2, 4, \ldots\} \subsetneq \mathbb{N},$$

and so the Pigeonhole Principle fails for infinite sets. This fact was the source of many puzzles going back to antiquity, and (perhaps) also one reason why this way of comparing arbitrary sets for *cardinality* (size) was never studied before Cantor. It was Cantor's bold stroke to transcend this "paradox" and to use equinumerosity as the basic relation in the construction of a far-reaching theory of cardinality for finite and infinite sets alike. To describe his first results, let us also introduce a notation for *inequality* and *strict inequality* of cardinal size:

$A \lesssim_c B \Leftrightarrow$ (there is a subset $C \subseteq B$)$[A \sim_c C]$,
$A <_c B \Leftrightarrow A \lesssim_c B$ and not $B \lesssim_c A$.

These relations are both *transitive*, e.g.,

if $A \lesssim_c B$ and $B \lesssim_c C$, then $A \lesssim_c C$.

A set A is *countable* if $A \lesssim_c \mathbb{N}$, otherwise it is *uncountable*.

Theorem 1. (Cantor's basic results)

(a) $\mathbb{N} \sim_c \mathbb{Z} \sim_c \mathbb{Q}$.
(b) *For every set* A, $A <_c \mathcal{P}(A)$.
(c) $\mathbb{N} <_c \mathbb{R}$, i.e., *there are uncountably many real numbers*.

Proof. (a) The equinumerosity $\mathbb{N} \sim_c \mathbb{Z}$ is proved much as we showed $\mathbb{N} \sim_c E$: we define a bijection of $\mathbb{Z}$ with $\mathbb{N}$ by setting $f(n) = 2n$ for $n \geq 0$ and $f(-n) = 2n - 1$ for $n < 0$.

To prove $\mathbb{N} \sim_c \mathbb{Q}$, set $A_0 = \{0\}$, $A_1 = \{-1, 1\}$, and for each $n > 1$, let A_n be the set of (proper) fractions $\frac{\pm d}{n}$ with $d \leq n$, i.e.,

$$A_0 = \{0\},\ A_1 = \{-1, 1\},\ A_2 = \left\{-\frac{1}{2}, \frac{1}{2}\right\},\ A_3 = \left\{-\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, \frac{2}{3}\right\}, \ldots$$

These sets are all finite, and if we list their members in a row we get an enumeration of all the fractions with absolute value ≤ 1:

$$0,\quad -1, 1,\quad -\frac{1}{2}, \frac{1}{2},\quad -\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, \frac{2}{3} \cdots$$

We can enumerate similarly all the fractions with absolute value > 1,

$$-2, 2, -3, -\frac{3}{2}, \frac{3}{2}, 3, -4, -\frac{4}{3}, \frac{4}{3}, 4, \cdots$$

and then interweave these two enumerations to get an enumeration of all the fractions, which now establishes a correspondence of $\mathbb{Q}$ with $\mathbb{N}$:

0	-2	2	-1	1	-3	$-\frac{3}{2}$	$\frac{3}{2}$	3	$-\frac{1}{2}$	$\frac{1}{2}$	-4	$-\frac{4}{3}$	$\frac{4}{3}$	4	$\cdots$
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	$\cdots$

(b) To see that $A \lesssim_c \mathcal{P}(A)$, assign to each $x \in A$ its *singleton* $\{x\}$, the subset of A whose only member is x; this is a one-to-one correspondence, and so A is equinumerous with the subset of $\mathcal{P}(A)$ consisting of all the singletons, so $A \precsim_c \mathcal{P}(A)$. To see that $\mathcal{P}(A) \precsim_c A$ cannot hold, suppose towards a contradiction that $f : C \to \mathcal{P}(A)$ is a one-to-one correspondence of some set $C \subseteq A$ with the powerset $\mathcal{P}(A)$, and let

$$D = \{x \in C \mid x \notin f(x)\}.$$

Now D is a subset of A, and so $D = f(d)$ for some $d \in C$; but then, by the definition of D and the choice of d, $d \in f(d) \Leftrightarrow d \in D \Leftrightarrow d \notin f(d)$, which is absurd.

(c) Let us first assign to each $A \subseteq \mathbb{N}$ the real number $f(A)$ with decimal expansion $.a_0 a_1 a_2 \cdots$, where $a_i = 0$ if $i \in A$ and $a_i = 2$ if $i \notin A$, so that, for example,

$$f(\varnothing) = .2222\cdots = \frac{2}{9}, \quad f(\mathbb{N}) = .0000\cdots = 0.$$

Simple properties of decimal expansions imply that f is a one-to-one function, and hence a correspondence of $\mathcal{P}(\mathbb{N})$ with some set of real numbers—those whose decimal expansions involve only 0's and 2's; thus $\mathcal{P}(\mathbb{N}) \precsim_c \mathbb{R}$, which with (b) prohibits $\mathbb{R} \precsim_c \mathbb{N}$.

Soon after these results and by the same "counting" methods, it was shown that the relation $\precsim_c$ *partially orders* the "equinumerosity classes" of sets:

Theorem 2. (Schröder-Bernstein). *For any two sets, if* $A \precsim_c B$ *and* $B \precsim_c A$, *then* $A \sim_c B$.

One easy consequence of this result is that

$$\mathcal{P}(\mathbb{N}) \sim_c \mathbb{R} \sim_c \mathbb{R}^n,$$

which is rather surprising at first glance, since it is not immediate how to construct a one-to-one correspondence between (say) the line and the plane.

Students find these arguments somewhat strange when they first see them, even today, and so did mathematicians in the 1870's: no "formula" is given for the correspondence $f : \mathbb{N} \to \mathbb{Q}$, although one can easily compute any required value of it, and the *diagonal argument* used in the proof of (b) was totally unconventional. On the other hand, the proofs are elementary and utterly convincing—and the results are unexpected and spectacular. In plain words: there are exactly "as many" fractions as there are natural numbers, but there are "strictly more" real numbers; and there is a sequence of ever-increasing "infinite cardinal sizes",

$$\mathbb{N} <_c \mathcal{P}(\mathbb{N}) <_c \mathcal{P}\big(\mathcal{P}(\mathbb{N})\big) <_c \cdots.$$

Cantor applied the new theory immediately to give a new proof of Liouville’s Theorem, that *transcedental numbers exist*. A real number is *algebraic* if it is a root of an algebraic equation $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$ with coefficients $a_0, \ldots, a_n \in \mathbb{Z}$ and $a_n \neq 0$, otherwise it is *transcedental*. Liouville’s had shown the existence of transcedental numbers by a direct construction, which (naturally) depended on several not-quite-elementary ideas from number theory. Cantor proved something stronger, and more simply: that *the set of algebraic numbers is countable*, and so it cannot exhaust $\mathbb{R}$, which is uncountable. The only (simple) algebraic fact he needed was a that *an algebraic equation has finitely many roots*, and with it he *counted* the algebraic numbers very much like he counted the fractions in the proof above. This was a “killer application” which made set theory quickly known (if somewhat notorious) to the mathematical community.

1.2. Cardinal Arithmetic

Cantor assumed—with some informal explanation—that every set A can be assigned a *cardinal number* $|A|$, so that, as with finite sets,

$$|A| = |B| \Leftrightarrow A \sim_c B, \tag{2}$$

and $0 = |\varnothing|$, $1 = |\{x\}|$, for any x, $2 = |\{x, y\}|$ if $x \neq y$, etc. The least infinite cardinal

$$\aleph = |\mathbb{N}|$$

is “the number of natural numbers”, and

$c = |\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$ is the cardinal size of *the continuum*. To define the arithmetical operations and the order on (possibly infinite) cardinal numbers, choose disjoint sets K, L so that $\kappa = |K|$, $\lambda = |L|$, and set

$$\begin{aligned}
&\kappa + \lambda = |K \cup L|,\\
&\kappa \cdot \lambda = K \times L,\\
&\kappa^{\lambda} = |(L \to K)|,\\
&\kappa \leq \lambda \Leftrightarrow K \lesssim_c L,
\end{aligned}$$

where the *function space* $(L \to K)$ (sometimes denoted by K^L or ${}^L K$) is the set of all functions $f : L \to K$. In particular, if $K = \{x, y\}$ is a doubleton, then $(L \to K) \sim_c \mathcal{P}(L)$, and so $2^{\lambda} = |\mathcal{P}(L)|$.

If κ and λ are finite, we get the usual sum, product, exponential and order on the natural numbers. Moreover, many of the usual laws of finite arithmetic hold: every non-empty set of cardinal numbers has a least member, addition and multiplication are associative and commutative, multiplication distributes over addition, and

$\kappa + 0 = 0, \quad \kappa^0 = 1, \quad \kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu, \quad \left(\kappa^\lambda\right)^\mu = \kappa^{\lambda\cdot\mu}.$

Addition and multiplication are trivial on infinite cardinal numbers:

if $\kappa, \lambda \neq 0$ and one of them is infinte, then $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda)$ (3)

These are the *absorption laws*. On the other hand, exponentiation is the source of some of the deepest problems in set theory: about the only simple inequality involving it is $\kappa < 2^\kappa$, from Theorem 1.

-
-
-

Biographical Sketch

Yiannis Moschovakis is a Professor of Mathematics at the University of California, Los Angeles, and also at the University of Athens, Greece, where he directs the Graduate Program in Logic, Algorithms and Computation. He has worked in Recursion Theory, Descriptive Set Theory, and more recently in the applications of logic and recursion to Theoretical Computer Science and the Philosophy of Language. A past speaker at the International Congress of Mathematicians, he has also served as President of the Association for Symbolic Logic.