Kurt Gödel

Metamathematical results on formally undecidable propositions:

Completeness vs. Incompleteness

Motto: The delight in seeing and comprehending is the most beautiful gift of nature.
(A.Einstein)

1. Life and work[1]

Kurt Gödel was a solitary genius, whose work influenced all the subsequent developments in mathematics and logic. The striking fundamental results in the decade 1929-1939 that made Gödel famous are the completeness of the first-order predicate logic proof calculus, the incompleteness of axiomatic theories containing arithmetic, and the consistency of the axiom of choice and the continuum hypothesis with the other axioms of set theory. During the same decade Gödel made other contributions to logic, including work on intuitionism and computability, and later, under the influence of his friendship with Einstein, made a fundamental contribution to the theory of space-time. In this article I am going to summarize the most outstanding results on incompleteness and undecidability that changed the fundamental views of modern mathematics.

Kurt Friedrich Gödel was born 28 April 1906, the second son of Rudolf and Marianne (Handschuh) Gödel, in Brno, Pekařská 5,[2] in Moravia, at that time the Austrio-Hungarian province, now a part of the Czech Republic. This region had a mixed population that was predominantly Czech with a substantial German-speaking minority, to which Gödel's parents belonged. Following the religion of his mother rather than his "old-catholic" father, the Gödels had Kurt baptized in the Lutheran church. In 1912, at the age of six, he was enrolled in the Evangelische Volksschule, a Lutheran school in Brno. Gödel's ethnic patrimony is neither Czech nor Jewish, as is sometimes believed. His father Rudolf had come from Vienna to work in Brno's textile industry, while his mother's family came from the Rhineland region for the work in textiles. At the age of six or seven Kurt contracted rheumatic fever and, despite eventual full recovery, he came to believe that he had suffered permanent heart damage as a result. Here are the early signs of Gödel's later preoccupation with his health.

From 1916 to 1924, Kurt carried on his school studies at the Deutsches Staats-Realgymnasium (Grammar School), where he excelled particularly in mathematics, languages and religion. Dr. Cyril Kubánek, professor of catholic religion, was Gödel's professor of philosophical propedeutics. Gödel probably obtained his interest in the philosophy of Immanuel Kant at this stage, which appears to have prevented him from later fully accepting the neo-positivist ideas of the Vienna circle. All the school register records testify to a great intellectual talent[3].

The World War I took place during Gödel's school years; it appears to have had little direct effect on him and his family. The collapse of the Austro-Hungarian empire at war's end together with absorption of Moravia, Bohemia and Slovakia into the new Czechoslovak

[1] Sources: Feferman (1986), Dawson (1984), Malina-Novotný (1996), Archives of Brno

[2] At that time also Brünn, Bäckergasse 5.

[3] A rarity can be found, however, in these reports: the evaluation excellent ("sehr gut") is stereotypically repeated in the end of every year; the only lower evaluation was on the first semester report – in mathematics.

Republic in 1918 also little affected the Gödels. After the war, the family continued life much as before, comfortably settled in the villa on Pellicova 8a in Brno.

Following his graduation from the Real-gymnasium in Brno in 1924, Gödel went to Vienna to begin his studies at the University. He was influenced by the number-theorist Philipp Furtwängler, but Professor Hans Hahn became Gödel's principal teacher, a mathematician of the new generation, interested in modern analysis and set-theoretic topology, as well as logic, the foundations of mathematics and the philosophy of science. It was Hahn who introduced Gödel to the group of philosophers around Moritz Schlick, which was later known as the "Vienna Circle" and became identified with the philosophical doctrine of *logical positivism* or *logical empiricism*[4]. Gödel attended meetings of the Circle quite regularly in the period 1924-1928. But in the following years he gradually moved away from it, though he maintained contact with some of its members, particularly with Rudolf Carnap.

The sphere of concerns of the Circle members must have influenced Gödel. He was acquainted with Ernst Mach's empiricist-positivist philosophy of science[5] and Bertrand Russell's logistic program, although he reports first studying the Principia Mathematica several years later. It seems that the most direct influences on Gödel were Carnap's lectures on mathematical logic and the publication of *Grundzűge der theoretischen Logik* by David Hilbert and Wilhelm Ackermann.

Hilbert posed as an open problem the question of whether there is a complete system of axioms and derivation rules for the first-order predicate logic. In other words, whether using the rules of the first-order system, it is possible to derive every logically valid statement. Gödel arrived at a positive solution to this *completeness problem* in the summer of 1929. The work became his doctoral dissertation and was published in a revised version in 1930. The Completeness Theorem is now the most fundamental theorem of model theory and mathematical proof theory.

Gödel's personal life changed in 1927 when he met Adele Nimbursky, a dancer who had been married before and was six years older than Kurt. Owing to the difference in their social situation, the developing relationship led to objections from Kurt's father. Although Kurt's father died not long after, Kurt and Adele were not to be married for another ten years. The death of Kurt's father in 1929 was unexpected; fortunately he left his family in comfortable financial circumstances. Gödel's mother retained the villa in Brno and took an apartment in Vienna with her two sons. Three days after his father's death, Kurt made an application[6] for release from nationality obligations in Czechoslovakia. He was released on condition that he would acquire state nationality in Austria within two years.

The ten years 1929–1939 were the most productive period in Gödel's intense life in mathematical logic, culminating in his greatest discoveries. It was the period of pursuit of avoiding paradoxes and inconsistencies in mathematics that had destroyed Frege's effort to establish a formal proof system for mathematics at the end of the 19th Century. David Hilbert (1862-1943), an outstanding German mathematician, put forward a new proposal for the foundation of classical mathematics which has come to be known as Hilbert's Program.

Pursuing Hilbert's program, Gödel started to work on the consistency problem for arithmetic and realized that the notion of provability can be formalized in arithmetic. He also saw that non-paradoxical arguments analogous to the well-known Liar paradox in ordinary language could be carried out by substituting the notion of provability for that of truth. Surprisingly, these efforts eventually led him to a most unexpected result, his proof of

[4] For a detailed discussion of Gödel's relations to the Vienna circle, see Köhler (1991)
[5] Max Plank's lectures, 1907, the most brilliant exposition of relativity of the period.
[6] Source: Archiv Brno, Certificate of domicile dated Feb 26, 1929.

incompleteness of arithmetic— effectively destroying the conclusions Hilbert had intuitively begun from when he originated his program. Gödel's work is generally taken to show that Hilbert's Program cannot be carried out. The latter has nevertheless continued to be an influential position in the philosophy of mathematics, and, starting with the work of Gerhard Gentzen in the 1930s, work on so-called Relativized Hilbert Programs have been central to the development of proof theory.

Gödel first announced his Incompleteness Theorem in 1930 to Carnap in Café Reichsrat in Vienna, a habitat of the Vienna Circle. The work on incompleteness was published early in 1931, and defended as a *Habilitationschrift* at the University of Vienna in 1932. The title of *Privatdozent* gave Gödel the right to give lectures at the university but without pay. As it happened he delivered lectures in Vienna only intermittently during the following years.

In 1933–1934 his unsalaried position in Vienna was supplemented by income from visiting positions in the United States of America. Gödel's first visit was to the Institute for Advanced Study in Princeton where he gave lectures on incompleteness results. The Institute had been formally established in 1930, with Albert Einstein and Oswald Veblen appointed its first professors. At that time Gödel apparently began to work on problems in axiomatic set theory. In the following years he felt rather depressed and lonely, particularly at Princeton. He had several nervous attacks of mental depression and exhaustion. In 1936 he spent almost the whole year in a sanatorium on account of mental illness. On September 20th, 1938 Kurt Gödel and Adele Nimbursky finally got married and their marriage proved to be a warm and enduring one. Adele was a source of constant support for Kurt in the difficult times ahead.

In March, 1939, after the occupation of Austria by Hitler, Gödel's unpaid position of *Privatdozent* had been abolished and he had to ask for a new paid position called *Dozent neuer Ordnung* (Docent of the New Order). He was also called up for a military physical examination, and much to his surprise found fit for the duty. On 27th of November he wrote a letter to Osvald Veblen in Princeton asking for help. Somehow German exit permits were arranged, and Kurt and Adele managed to leave Vienna in January 1940. They travelled by train through Eastern Europe, then via the Trans-Siberian Railway across Russia and Manchuria to Yokohama where they took a ship to San Francisco. In March 1940 they finally came by train to Princeton. Gödel was never to return to Europe.

So it was in 1940 Gödel was made an Ordinary Member of the Institute for Advanced Study, and he and his wife settled in Princeton. Among his closest friends there were Albert Einstein and Oskar Morgenstern; the latter was another ex-Viennese, an economist who emigrated from Austria in 1938. At the Institute Gödel had no formal duties and was free to pursue his research and studies. In the springtime 1941 he gave a series of lectures, and on April 15th he gave a lecture at the Yale University on "In which sense is intuitionistic logic constructive"? He continued his work in mathematical logic; in particular he made efforts to prove the independence of the axiom of choice and the continuum hypothesis. He partially succeeded on this problem. His masterpiece *Consistency of the axiom of choice and of the generalized continuum-hypothesis with the axioms of set theory* (1940) is a classic of modern mathematics. In this he proved that if an axiomatic system of set theory of the type proposed by Russell and Whitehead in *Principia Mathematica* is consistent, then it will remain so when the axiom of choice and the generalized continuum-hypothesis are added to the system. This did not prove that these axioms were independent of the other axioms of set theory, but when this was finally established by Cohen in 1963 he built on these ideas of Gödel. Another achievement early in this period (published only in 1958) was a new constructive interpretation of arithmetic that proved its consistency, but via methods going beyond finitary means in Hilbert's sense.

From 1943 on, Gödel devoted himself almost entirely to philosophy, first to the philosophy of mathematics and then to general philosophy and metaphysics. Gödel is noted for his support of mathematical realism and Platonism[7]. In this general direction he joins such noted mathematicians and logicians as Cantor, Frege, Zermelo and Church, and the implicit working conceptions of most practicing mathematicians. An expository paper on Cantor's continuum problem in 1947 brought out Gödel's Platonist views quite markedly in the context of set theory. As for general philosophy, Gödel continued his long-pursued study of Kant and Leibniz.

Beginning in 1951, Gödel received many honours. Particularly noteworthy was his sharing of the first Einstein Award (with Julian Schwinger) in 1951. John von Neumann, one of the first to understand Gödel's incompleteness results, compared Gödel's contribution in the field of logic with the work of Aristotle; von Neumann died on January 8th, 1957, Einstein died on April 18th 1955. This was Gödel's best friend and regular companion on their walk home from the Institute. Einstein and Gödel seemed very different in almost every personal way — Einstein full of laughter and common sense and Gödel solemn, serious and solitary — but they shared a fundamental feature: both went directly and rigorously to the fundamental questions at the very heart of things.

From 1959 on, in addition to Gödel's primary interest in logic, philosophy and, to a lesser extent, mathematics and physics, he was interested in phenomenology. Gödel's notes are preserved in his *Nachlass* (inheritance), and many of them are concerned with the phenomenology of Edmund Husserl. These notes are unexpectedly wide-ranging, revealing interests in history and theology. A logical attempt at the proof of God's existence is found here. The 'proof' was written in 1970 and it reminds a sacral text: it has no introduction, no motivation, and no explication of the modal system used; just axioms, definitions, and the proof. It is an ontological proof, based on Anselm principle, but Gödel does not refer to St. Anselm, or to other philosophers and theologians.

In the last fifteen years of his life, Gödel was busy with Institute business[8] and his own philosophical studies; during this time he returned to logic only occasionally, devoting some efforts to revision and translation of his old papers. He translated and revised his 1958 paper[9], which gave a constructive interpretation of arithmetic, but the revised version was never published.

On April 21–23, 1966, a 60th birthday symposium was organised at Ohio State University; but the invitation to attend was declined by Gödel. On July 23rd Marianne (Handschuh) Gödel (mother) died in Vienna, and in August Gödel refused an honorary membership in Austrian Academy of Sciences. In fact, Gödel's health was poor from the late 1960s on. His wife Adele was not able to help him as before, being herself partially incapacitated, and for a time moved to a nursing home.

Gödel's depressions returned accompanied by paranoia; he developed fears about being poisoned and would not eat. Kurt Gödel died in Princeton Hospital on January 14th, 1978 of "malnutrition and inanition caused by personality disturbance". Adele survived him by three years. Kurt and Adele had no children, leaving Kurt's brother Rudolf as the sole surviving member of the Gödel family.

In 1987 an international Kurt Gödel Society was established in Vienna, the first president of which was Gödel's student and friend Hao Wang. In 1992 the Society of Kurt Gödel was

---

[7] See, e.g., Köhler (2002a)
[8] He was made a Permanent Member of the Institute in 1946, and promoted to Professor in 1953
[9] The last published paper apart from revisions of earlier works: 'Über eine bisehr noch nicht benützte Erweiterung des finiten Standardpunktes", *Dialectica* 1958.

founded in Brno, and the Society is an organiser of the International conference *Logical Foundations of Mathematics, Computer Science and Physics – Kurt Gödel's Legacy* held regularly every four years. The first conference Gödel'96 was held in Brno on August 25–29, 1996 on the occasion of Gödel's 90^{th} birthday.

2. Completeness of the 1^{st}-order predicate logic proof calculus

Now we are going to deal in more details with Gödel's undoubtedly greatest results, namely those on completeness and incompleteness. There is a question, however: How to communicate something from those ingenious thoughts to a reader enthusiastic about rigorous science but not being a specialist in mathematical logic? There are at least three ways of doing so. First, to give a systematic historical exposition of the development of ideas that led to the major Gödel's achievements. Second, to outline a philosophical interpretation of the results; and third, to enunciate basic ideas of Gödel's work in a comprehensive way, in terms of current logical systems. Instead of doing the first, I briefly summarized Gödel's life. Now I will confine myself to a mathematical exposition accompanied by brief philosophical and historical comments, because I am convinced that without a good understanding of the mathematical fundamentals any historical and philosophical considerations would not be well-founded. I will not reproduce original Gödel's formulations and proofs. Instead, I will give an exposition from the point of view of current mathematical logic. I would just like to stress that Gödel's results are mathematical *facts.* Despite a strong resemblance to Liar paradox (and an obvious inspiration by it) they are no paradoxes, no hypotheses.

2.1. First order predicate logic.

In mathematical logic we work with closed well-formed formulas (called *sentences*) which in a less or more precise way render the logical structure (meaning) of our statements. We define, what it means that a formula φ is *provable* (from some premises) and that a formula φ is *true* (under some interpretation). The notions of provability and truth are two basic notions of mathematical logic; in which way are they related? *Are provable formulas exactly those that are true (under some or all interpretations)?*

To make sense of this fundamental question, we have to explicate the notions of *formula, provability* and *truth*. Gödel's results on completeness and incompleteness are two answers to our question — one positive (and coming up to expectations of that time) and one negative (at that time an unexpected surprise). I will also try to elucidate a terminological confusion that is frequently caused by a nodding acquaintance with Gödel's work, when two distinct notions of *completeness* are commingled: completeness of a *proof calculus* and completeness of an *axiomatic theory* (formulated within a calculus). I will also distinguish two notions of *decidability* (Entscheidbarkeit): decidability of a formula in a given theory and decidability of the whole theory.

Language of the 1^{st}-order predicate logic (FOPL) has nowadays become a mathematical stenography. Using the FOPL language we can characterise properties (denoted by predicate symbols of arity 1) of objects of a universe of discourse, and *n*-ary relations (denoted by predicate symbols of arity *n*) between objects of a universe of discourse. We can also express propositions that some (using existential quantifier ∃) or all (using universal quantifier ∀) objects *x* have a property P or are in a relation Q. The language is, however, formal: its symbols are devoid of meaning, they are empty signs. The reader might well wonder what sense it makes to claim that using empty signs we can express meaningful propositions. The answer is that we are walking a subtle midway path between truly empty signs and truly meaningful ones. To evaluate a formula we have to interpret the formula. For instance, the following formula φ

$$\forall x\,[P(x) \to Q(x, a)]$$

"claims": for all x it holds that if x has a property P then this x is in a relation Q with an a. The question whether φ is true does not make sense until we know what 'P', 'Q' and 'a' mean. To evaluate a truth-value of φ we have to choose the universe of discourse over which the variable x can range. For instance, let the universe be the set of natural numbers N. Second, we have to assign a subset of N to the predicate symbol 'P' and a binary relation over N (i.e., a subset of the Cartesian product $\mathsf{N} \times \mathsf{N}$) to the predicate symbol 'Q'. Let P stand for the set E of even numbers and Q for the relation D, "divisible by". Third, we have to assign an element of N to the constant symbol 'a', let it be the number 2. Under this *interpretation* the formula φ is true (all the even numbers are divisible by 2). We say that the *structure*

$$\mathsf{M} = \langle \mathsf{N}, \mathrm{E}, \mathrm{D}, 2\rangle,$$

where the set E is assigned to the symbol 'P', the relation D to the symbol 'Q' and the number 2 to the symbol 'a', is a *model* of the formula φ. There are other models of φ, for instance the structure

$$\mathsf{M}' = \langle \mathsf{N}, \mathrm{Pos}, >, 0\rangle,$$

where Pos (assigned to the symbol 'P') is the set of positive numbers, > (assigned to the symbol 'Q') is the ordinary linear ordering of numbers and 0 is the number zero (assigned to the constant 'a'). Under this interpretation φ claims that all the positive numbers are greater than zero, which is obviously true. The structure

$$\mathsf{M}'' = \langle \mathsf{N}, \mathrm{E}, \mathrm{D}, 3\rangle$$

is not a model of φ (it is not true that all the even numbers are divisible by 3). It is, however, a model of another formula ψ, namely

$$\exists x\,[P(x)\ \&\ Q(x,a)],$$

for there are such numbers that are even and divisible by 3. Formulas φ and ψ are satisfied by a model independently of a valuation of x; we say that x is bound here by quantifiers (general ∀ or existential ∃, respectively), and the formulas φ, ψ are *closed*. We will call closed formulas *sentences.*

Some formulas may have free variables. For instance the formula $[P(x)\ \&\ Q(x,a)]$ is not closed; it cannot be evaluated even if a structure M'' is assigned to it by the realization of 'P', 'Q' and 'a' (i.e., by assigning the set E to the symbol 'P', the relation D to the symbol 'Q' and the number 3 to the constant symbol 'a'), because its truth-value in M'' depends on a valuation e of x. *Valuation* is a total function that assigns elements of the universe of discourse to variables. If e assigns the number 2 to x, the formula is false, if it assigns the number 6, the formula is true, it is satisfied by this valuation.

Using these elementary examples, we illustrated almost all the basic notions we need: predicates of arity n (here P of arity 1 and Q of arity 2), n-ary functional symbols (here constant a of arity 0), variables (here x), logical connectives (such as & ('and'), ∨ ('or'), → ('if … then'), ¬ ('not')), quantifiers (∀–'all', ∃–'some'), and the notion of formula, its interpretation and a model. If a universe U of discourse is chosen, realization of predicate symbols consists in assigning subsets of the universe to symbols of arity 1, and n-ary relations over the universe U (subsets of Cartesian products U^n) to *n-ary* predicate symbols. Constant symbols are realized as elements of the universe U, and n-ary functional symbols as mappings from the Cartesian product of the universe to the universe ($U^n \to U$). Logical symbols such as connectives ($\neg$, &, $\vee$, $\rightarrow$, etc.), quantifiers ($\forall$, $\exists$) and = (identity) are not interpreted, they have the fixed standard meaning[10].

Some formulas are true under every interpretation for any valuation of variables, i.e., they are valid in any interpretation structure. They are called *logically valid formulas* (also *logical truths* or *logical laws*). For instance, the formula $[\forall x\ P(x) \vee \forall x\ Q(x)] \rightarrow \forall x\ [P(x) \vee Q(x)]$ is a logical truth. Indeed, if the antecedent $[\forall x\ P(x) \vee \forall x\ Q(x)]$ of the implication is true under some interpretation over a universe U, then either the realization P^U of the symbol 'P' is equal to U or the realization Q^U of 'Q' is equal to U, or both. Which means that the set-theoretical union of the sets P^U and Q^U is equal to U ($P^U \cup Q^U = U$), and the consequent $\forall x\ [P(x) \vee Q(x)]$ is true under this interpretation as well. According to the definition of implication the whole formula is true; it cannot be false under any interpretation.

Summarizing: By $\mathbf{M} \models \varphi[e]$ we denote the fact that a formula φ is *satisfied* by the structure M and a valuation *e*. In other words, the formula φ is true under the interpretation over M, for the valuation *e*. If φ is true under M for *all* valuations *e* (of variables by elements of the universe), then M is a *model* of φ, or φ is *valid in* M; in symbols $\mathbf{M} \models \varphi$. Formula φ is *logically valid (logical truth),* if φ is true under every interpretation, denoted $\models \varphi$.

2.2. Hilbert's program

Before introducing Gödel's results I have to briefly describe the atmosphere in which they appeared. Paradoxes and conceptual problems of mathematics often stem from the infinite. This includes, for example, Zeno's paradoxes in Greek times, infinitesimals in the seventeenth century, and the paradoxes of set theory in the late nineteenth and early twentieth centuries. In any case, the problem appeared when mathematicians began to reason with infinite quantities.

The German mathematician David Hilbert (1862-1943) announced his program in the early 1920s. It calls for a formalization of all of mathematics in axiomatic form, and of proving the consistency of such formal axiom systems. The consistency proof itself was to be carried out using only what Hilbert called "finitary" methods. The special epistemological character of finitary reasoning then yields the required justification of classical mathematics. Although Hilbert proposed his program in this form only in 1921, it can be traced back until around 1900, when he first pointed out the necessity of giving a direct consistency proof of analysis. Hilbert first thought that the problem had essentially been solved by Russell's type theory in *Principia*. Nevertheless, other fundamental problems of axiomatics remained unsolved, including the problem of the "decidability of every mathematical question", which also traces back to Hilbert's 1900 address.

Within the next few years, however, Hilbert came to reject Russell's logicistic solution to the consistency problem for arithmetic. In three talks in Hamburg in the summer of 1921 Hilbert presented his own proposal for a solution to the problem of the foundation of mathematics. This proposal incorporated Hilbert's ideas from 1904 regarding direct consistency proofs, his conception of axiomatic systems, and also the technical developments in the axiomatization of mathematics in the work of Russell as well as the further developments carried out by him and his collaborators. What was new was the finitary way in which Hilbert wanted to carry out his consistency project.

He accepted Kant's finitist view in the sense that we obviously cannot experience infinitely many events or move about infinitely far in space. However, there is no upper bound on the number of steps we execute. No matter how many steps we may have executed,

[10] For details and precise definitions see, e.g., Mendelson (1997)

we can always move a step further. But at any point we will have acquired only a *finite* amount of experience and have taken only a finite number of steps. Thus, for a Kantian like Hilbert, the only legitimate infinity is a *potential* infinity, not the *actual* infinity. The Kantian element of Hilbert's view is what separates his formalism from earlier, implausible accounts. Hilbert's problem, as he saw it, lies in how infinite mathematics can be incorporated into the finite Kantian framework. Hilbert would say that finitist mathematical truths, intimately bound up with our perception, could be known *a priori,* with complete certainty. If we were content with finitist mathematics, this would be the end of the story. But Hilbert wanted more than this, and rightly so. He wanted to keep the extraordinary beauty, power and utility of classical mathematics, but he also wanted to do it in such a way that we could be fully confident that no more paradoxes would arise. This includes transfinite set theory, about which he declared: "No one shall drive us out of the paradise that Cantor has created for us".[11]

What Hilbert needs to do is to show that various parts of infinite mathematics will fit with one another and finite mathematics in such a way that no inconsistency could be derived. But what is involved in deriving things, in mathematical reasoning? Hilbert fixes on the symbols themselves. Here is the core of formalism: mathematics is about symbols. Hilbert's Kantian idea is now to study these symbols mathematically, not using the questionable infinity, but rather finite meaningful mathematics intimately linked to concrete *symbols* of classical mathematics itself. Hilbert was convinced that mathematical thinking could be captured by the syntactic laws of pure symbol manipulation[12].

Work on the program progressed significantly in the 1920s and many outstanding logicians and mathematicians took part in it, such as Paul Bernays, Wilhelm Ackermann, John von Neumann, Jacques Herbrand and, of course, Kurt Gödel.

2.3. Completeness of the proof calculus

The idea of finitist axiomatisation is simple: if we choose some basic formulas (axioms) that are decidedly true and if we use a finite method of applying some simple rules of inference that preserve truth, no falsehood can be derived from these axioms; hence no contradiction can be derived, no paradox can arise.

Logically valid formulas that are true under each interpretation are the most indisputably true formulas. Let us consider some logically valid formulas:

(1) $\varphi \rightarrow (\psi \rightarrow \varphi)$

(2) $(\varphi \rightarrow (\psi \rightarrow \xi)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \xi))$

(3) $(\neg\varphi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \varphi)$

(4) $\forall x\, \varphi(x) \rightarrow \varphi(c)$ (where c is a constant or a suitable variable, $\varphi(c)$ arises from $\varphi(x)$ by correct substituting c for x)

(5) $\forall x\, (\varphi \rightarrow \psi(x)) \rightarrow (\varphi \rightarrow \forall x\, \psi(x))$ (variable x does not occur free in the formula φ)

We can easily see that (1)–(5) are logically valid. For instance, (1) says that if φ is true then it is implied by any ψ, which is true due to the definition of mathematical notion of implication. The exact verification is however out of scope of the present article.

---

[11] Brown (1999), Hilbert (1926, p.170): "Aus dem Paradies, daß Cantor uns geschaffen, soll uns niemand vertreiben können".

[12] In advance we can state at this point that Gödel's Incompleteness results showed that this belief in the power of symbol manipulation was not realistic. Actually, Gödel's results delimitate the possibilities of mechanical symbol manipulation.

Now we have to choose some rules of derivation, which will produce new logical truths from the axioms (1)–(5). They are, for instance[13]:

i) *modus ponens*: from formulas φ and $(\varphi \rightarrow \psi)$ derive ψ; denoted $\varphi, (\varphi \rightarrow \psi) \vdash \psi$

ii) *generalization*: from a formula φ derive $\forall x\, \varphi$; denoted $\varphi \vdash \forall x\, \varphi$

The modus ponens rule is truth preserving: indeed, if ψ is true on the assumption that φ, and φ is true, then ψ must be true as well. The generalization rule is, however, obviously not truth preserving; but it preserves logical truth: if φ is logically valid, then it is satisfied by any structure and for any valuation of variables; hence $\forall x\, \varphi$ is logically valid as well.

To make the notion of a finite inference method perfectly precise, we define a *proof*:

A sequence of formulas $\varphi_1,\ldots,\varphi_n$ is a *proof*, if each formula φ_i is either

- an axiom or
- is derived from some previous members of the sequence $\varphi_1,\ldots,\varphi_{i-1}$ using a derivation rule i) or ii).

A formula φ is *provable* in the calculus (or *theorem* of the calculus, denoted $\vdash \varphi$) if it is the last member of a proof.

Since the axioms are logically valid (logical truths), and since *modus ponens* is a truth-preserving rule and generalization is a logical-truth-preserving rule, it is obvious that each φ_i of a proof $\varphi_1,\ldots,\varphi_n$ is a logically valid formula. Hence each provable formula, *theorem of the calculus,* is logically valid. We have defined a *sound proof calculus* (if $\vdash \varphi$, then $\models \varphi$).

In 1928 Hilbert and Ackermann published a concise small book *Grundzüge der theoretischen Logik,* in which they arrived at exactly this point: they had defined axioms and derivation rules of predicate logic (slightly distinct from the above), and formulated the problem of *completeness*. They raised a question whether such a proof calculus is *complete* in the sense that each logical truth is provable within the calculus; in other words, whether the calculus proves exactly all the logically valid FOPL formulas.

Gödel's *Completeness theorem* gives a positive answer to this question: the 1st-order predicate proof calculus (with appropriate axioms and rules, like those (1)—(5), i) and ii) above) is a complete calculus, i.e., all the FOPL logical truths are provable (if $\models \varphi$, then $\vdash \varphi$).

In FOPL syntactic provability is equivalent to being logically true: $\models \varphi \Leftrightarrow \vdash \varphi$.

There is even a stronger version of the Completeness theorem that Gödel formulated as well. We derive consequences not only from logically valid sentences but also from other sentences true under some interpretation. For instance, from the fact that all the even numbers are divisible by 2 and the number 6 is even we can derive that the number 6 is divisible by 2. In FOPL notation we have:

$$\forall x\, [P(x) \rightarrow Q(x,a)],\ P(b) \vdash Q(b,a).$$

But none of these formulas is a logical truth. Yet this derivation is correct, since the conclusion is *logically entailed* by the premises: whenever the premises are true, the conclusion must be true as well. In other words, the conclusion is true in all the models of the premises.

[13] More exactly, the above are *schemes* of axioms and rules. The system we demonstrate here is nowadays known as Hilbert calculus. There are other possibilities of choosing axioms and rules, of course. For instance, sequent (Gentzen) calculus or natural deduction have even fewer axioms (usually just one) and rather more natural rules of deduction.

To formulate the strong version of the Completeness theorem, we have to define the notion of a *theory* and a *proof in a theory.* In mathematics we often need to characterise some common features of particular distinct structures. For instance, the structure $\underline{\mathsf{N}} = \langle \mathsf{N}, \leq \rangle$, where N is the set of natural numbers and $\leq$ its usual linear ordering, can be characterised by a set O of the following formulas:

$$\forall x\, P(x,x) \qquad \varphi_1$$

$$\forall x \forall y\, [(P(x,y)\ \&\ P(y,x)) \to x=y] \qquad \varphi_2$$

$$\forall x \forall y \forall z\, [(P(x,y)\ \&\ P(y,z)) \to P(x,z)] \qquad \varphi_3$$

We say that the set O of formulas φ_1, φ_2, φ_3 is a *theory* of *partial ordering*. Particular formulas φ_1, φ_2, φ_3 are *special axioms* of the theory O and they characterize *reflexivity, anti-symmetry* and *transitivity*, respectively, of a partial ordering relation.

If the binary relation $\leq$ is assigned to the symbol P (and variables *x, y, z* range over N), each of these formulas is valid in the structure $\underline{\mathsf{N}}$. We also say that this structure is a *model of the theory O*.

Moreover, for any set S the structure $\underline{\mathsf{S}} = \langle \mathsf{P(S)}, \subseteq \rangle$, where $\mathsf{P(S)}$ is the power set of S and $\subseteq$ is the relation of the set-theoretical inclusion, is a model of the theory O as well.

Both the structures are also a model of $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$, where

$$\exists x \forall y\, P(x,y) \qquad \varphi_4$$

claims that among the elements of the universe at least one element exists such that it is in a relation P with all the elements of the universe. The formula

$$\forall x \forall y\, [P(x,y) \vee P(y,x)] \qquad \varphi_5$$

is satisfied by $\underline{\mathsf{N}}$, but not satisfied by $\underline{\mathsf{S}}$. We say that the set LO $\{\varphi_1, \varphi_2, \varphi_3, \varphi_5\}$ is a theory of *linear ordering,* and LO $\cup$ $\{\varphi_4,\}$ is the theory of linear ordering with the least element. The power set of a set S is not ordered linearly.

Now we can define: a (FOPL) *theory* is given by a (possibly infinite) set of FOPL formulas, the *special axioms*.

A *proof in a theory T* is a sequence of formulas $\varphi_1,\ldots,\varphi_n$ such that each φ_i is either

- a logical axiom or
- a special axiom of T, or
- is derived from some previous members of the sequence $\varphi_1,\ldots,\varphi_{i-1}$ using a derivation rule i) or ii).

A formula φ is *provable* in T iff it is the last member of a proof in T; we also say that the theory T proves φ, and the formula φ is a *theorem* of the theory (denoted T $|\!-\ \varphi$). A structure M is a *model of the theory* T, denoted M |= T, iff each special axiom of T is valid in M.

You may wonder whether the calculus is sound even in the stronger sense: whether each theorem of a theory, i.e., a formula provable in the theory, is logically entailed by the special axioms (denoted T |= φ); in other words, whether each theorem is valid in all the models of the theory. As said above, the generalization rule is not truth-preserving. From the fact that, e.g., an *x* is even, we cannot correctly derive that all the *x*'s are even. However, generalization of the form $\varphi(x)\ |\!-\ \forall x\, \varphi$, where *x* is free in φ, is intuitively correct in case that the *x* in φ is *any* element of the universe. Indeed, if φ is true in a model M of the theory for any *x,* $\forall x\, \varphi$ is true in the model M as well. In other words, the generalization rule preserves the validity of a formula in a model. For this reason not only formulas satisfied but also valid in the intended model are chosen for special axioms. On this condition, which is trivially met by sentences, the calculus is sound: if $\mathbf{T} \mid\!- \varphi$, then $\mathbf{T} \models \varphi$.

Another natural demand on special axioms is their mutual consistency. If the axioms contradicted each other, anything would be entailed by them, and any formula would be provable. The theory would be useless. Thus we define:

A theory T is *consistent* iff there is a formula φ which is not provable in T.

The *strong version of the Completeness theorem* claims that a formula φ is *provable in a (consistent) theory T if and only if φ is logically entailed by its special axioms*; in other words, iff φ is valid in every model of the theory; in (meta) symbols:

$$\mathbf{T} \models \varphi \Leftrightarrow \mathbf{T} \mid\!- \varphi.$$

The proof of the Completeness theorem is based on the following Lemma:
Each consistent theory has a model.

We need to prove that any formula φ that is logically entailed by T is also provable in T (if $T \models \varphi$ then $T \mid\!- \varphi$). We will show that if T does *not* prove φ then φ is *not* logically entailed by T (if not $T \mid\!- \varphi$, then not $T \models \varphi$). Indeed, if T does not prove φ then T extended by ¬φ, i.e., $\{T \cup \neg\varphi\}$, does not prove φ as well (¬φ does not contradict T), which means that $\{T \cup \neg\varphi\}$ is consistent. Hence according to the Lemma there is a model M of $\{T \cup \neg\varphi\}$; however, M is a model of T in which φ is not true, which means that φ is not entailed by T.

Hilbert expected the Completeness theorem; this result was valuable but it was not a surprise. Hilbert, however, expected more. He wanted to avoid any inconsistencies in mathematics. Arithmetic of natural numbers is a fundamental theory of mathematics. Consider the set ω of natural numbers {0, 1, 2, …}. Often we say that there are infinitely many of them: no matter how far we count, we can always count one more. But Cantor's set theory actually says something much stronger: it says, e.g., that the power set P(ω) of all the subsets of ω is a set as well, and is larger than ω (uncountable), which means that an *actual infinite* exists. Platonists have no trouble with actual infinities while thinkers like Kant, and later intuitionists reject them outright, allowing only potential infinities. For Hilbert, statements involving the infinite are 'meaningless' but useful, justified by their enormous power and utility. He thought of these as 'ideal elements' that can be added to the meaningful, finite, true mathematics as supplements to make things run smoothly and to derive new things. There is, however, a necessary condition that these elements are added in a consistent way. Hilbert declares: 'there is one condition, albeit an absolutely necessary one connected with the method of ideal elements. That condition is a *proof of consistency,* for the extension of a domain by the addition of ideal elements is legitimate only if the extension does not cause contradictions'[14]. Hence Hilbert needed to find a *consistent* theory whose axioms characterise arithmetic of natural numbers *completely,* so that each arithmetic truth expressed in a formal language would be logically entailed by the axioms and thus derivable from them in a finite number of steps. Moreover, the set of axioms has to be fixed and initially well defined. Gödel's two theorems on incompleteness show that these demands cannot be met.

[14] Brown (1999, p.66)

3. Incompleteness.

3.1. Incompleteness of arithmetic, Gödel's first and second theorems

The results on incompleteness were announced by Gödel in 1930 and the work "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I" was published in 1931. This work contained a detailed proof of the Incompleteness theorem and a statement of the second theorem; both statements were formulated within the system of *Principia Mathematica*. In 1932 Gödel published in Vienna a short summary "Zum intuitionistischen Aussagenkalkül", which was based on a theory that is nowadays called Peano arithmetic. I will now present[15] these revolutionary results in terms of current systems of mathematical logic introduced above.

Now we are not interested just in *logical* truths, i.e., sentences true under *every interpretation* of the FOPL language, but in sentences characterizing arithmetic of natural numbers which are true under the *standard (intended) interpretation*, which is the structure **N**:

$$\mathbf{N} = \langle \mathrm{N}, 0, \mathrm{S}_\mathrm{N}, +_\mathrm{N}, *_\mathrm{N}, =_\mathrm{N}, \leq_\mathrm{N} \rangle$$

where N is the set of natural numbers, 0 is the number zero, S_N is the successor function (adding 1), $+_\mathrm{N}$ is the sum function (adding natural numbers), $*_\mathrm{N}$ is the multiplication function (on natural numbers), $=_\mathrm{N}$ is the relation of identity on natural numbers, and $\leq_\mathrm{N}$ is the relation "less than or equal" of linear ordering on natural numbers.

In order to be able to create formulas true in **N**, the alphabet of the arithmetic language has to contain a constant symbol $\underline{0}$, unary functional symbol S, binary functional symbols + and *, and binary predicate symbols =, ≤; the obvious intended interpretation associates these symbols with the respective elements of **N**. In this language we can express sentences like $\forall x \forall y\ (x+y) = (y+x)$, or $\exists x\ (\mathrm{S}(\mathrm{S}(x)) \leq \underline{0})$, the former being true in the structure **N**, the latter being false under this intended interpretation[16]. Actually, each sentence φ of the arithmetic language is either true or false under the intended interpretation. Hence if a theory is to characterise **N** completely, i.e., to demonstrate *all* the arithmetic truths, there must not be a sentence *independent* of T, i.e., neither provable nor refutable:

A theory T is complete if T is consistent and for each sentence φ it holds that *T proves either φ or ¬φ,* $\mathrm{T} \vdash \varphi$ or $\mathrm{T} \vdash \neg\varphi$; in other words, *each sentence φ is decidable in T.*

There are incomplete theories. Since according to the Completeness theorem, any consistent theory proves just the sentences entailed by the theory, to show that a theory is incomplete, we need to find an independent sentence φ that is neither entailed by the axioms of the theory, nor contradicts them (because then the theory would prove ¬φ); in other words, a sentence true in some but not all models of the theory. For instance, the theory O of partial ordering introduced in the previous chapter is not complete: there are partially ordered sets, like the set N ordered by $\leq_\mathrm{N}$, which are ordered linearly, and partially ordered sets that are not ordered linearly, like the power set of a set S ordered by set-theoretical inclusion. Hence the formula φ_5 — $\forall x \forall y\ [\mathrm{P}(x,y) \vee \mathrm{P}(y,x)]$ —is independent of the theory $\mathrm{O} = \{\varphi_1, \varphi_2, \varphi_3\}$. Also, the theory of linear ordering $\mathrm{LO} = \{\varphi_1, \varphi_2, \varphi_3, \varphi_5\}$ is incomplete. The sentence independent of LO is the sentence φ_4 — $\exists x \forall y\ \mathrm{P}(x,y)$. There are also complete theories, like, e.g., the theory of discrete ordering or the theory of a successor, however the proof of their completeness is rather non-trivial.

---

[15] For details, see Hájek (1996), Švejdar (2002)

[16] In the arithmetic language we use an infix notation when applying symbols like '+', '*', '≤', and instead of $\neg(\mathrm{S}(x) = a)$ we write $\mathrm{S}x \neq a$.

The FOPL proof *calculus* (with the logical axioms and rules described in the previous chapter) is a *complete calculus*: it proves all the logically valid formulas of FOPL. The calculus can be viewed as a theory without special axioms. This empty *theory* is *not a complete theory*: for instance, a simple formula like $\exists x\ P(x)$ is not decidable. Thus the calculus does not decide even simple arithmetic truths, for generally they are not logical truths. It might seem that the calculus decides at least all the logical truths, since they are provable. We will show as a consequence of Gödel's first incompleteness theorem that the problem of logical truth is also not decidable within FOPL.

To characterize arithmetic truths, we need some special axioms formulated in the arithmetic language. As an example we adduce a theory Q, called *Robinson's arithmetic*, given by the following seven axioms:

$\forall x \quad (Sx \neq \underline{0})$
$\forall x\forall y \quad (Sx = Sy \rightarrow x = y)$
$\forall x \quad (x + \underline{0} = x)$
$\forall x\forall y \quad (x + Sy = S(x + y))$
$\forall x \quad (x * \underline{0} = \underline{0})$
$\forall x\forall y \quad (x * Sy = (x * y) + x)$
$\forall x\forall y \quad (x \leq y = \exists z\ (z + x = y)\ )$

The theory Q characterizes basic arithmetic operations (a successor of any number is not equal to zero; adding zero to any number gives as a result the same number, etc.), and the structure **N** is a model of the theory; Q is however a weak theory. General simple statements like commutativity of adding or multiplying, i.e., sentences $\forall x\forall y\ (x+y = y+x)$ and $\forall x\forall y\ (x * y = y * x)$ cannot be proved in Q.

The theory Q proves only syntactically simple sentences. Syntactical complexity of a sentence is determined by a number of alternating quantifiers. More precisely: We say that an arithmetic formula φ is formed from a formula ψ by a *bounded quantification*, if φ has one of the following forms (the binary predicate symbol '<' is being interpreted as the "less than" relation, i.e., $x < y$ abbreviates $(x \leq y)\ \&\ \neg(x=y)$):

$\forall v\ (v < x \rightarrow \psi)$, $\exists v\ (v < x\ \&\ \psi)$, $\forall v\ (v \leq x \rightarrow \psi)$, $\exists v\ (v \leq x\ \&\ \psi)$,

where *v, x* are distinct variables. Quantifiers of the above form are called *bounded quantifiers.* A formula φ is a *bounded formula* if it contains only bounded quantifiers. A formula φ is a *Σ-formula*, if φ is formed from bounded formulas using only conjunction, disjunction, existential quantifier and any bounded quantifiers.

There is an interesting, rather non-trivial fact valid of Robinson's arithmetic: Q is *Σ-complete*; it proves all the Σ-sentences that should be provable, i.e., the Σ-sentences true in **N**: if σ is a Σ-sentence such that **N** $\models \sigma$, then Q $\vdash \sigma$.

If we extend the theory Q by the scheme of *induction axioms*:

$[\varphi(\underline{0})\ \&\ \forall x\ (\varphi(x) \rightarrow \varphi(Sx))] \rightarrow \forall x\ \varphi(x)$,

we obtain the theory PA called *Peano arithmetic*. Note that we added a *scheme* of *infinitely many* axioms which can be obtained by substituting a formula for φ. Yet this theory is "reasonable", it conforms to finitism: we added a "geometrical pattern" of formulas. Thus in PA we have a finite number of structural relations in which the formulas stand to each other, and proving in the theory does not involve a procedure that would make reference to actual infinity.

The structure **N** is a *standard model* of PA. Each number $n \in$ N is denoted by a term of the arithmetic language, namely the term SS…S$\underline{0}$ (the n^{th} successor of the constant $\underline{0}$), called the n^{th}–*numeral.* We use an abbreviated notation: $\underline{n}$.

Peano arithmetic is rather a strong theory and many laws of arithmetic are provable in it; however, it is not a complete theory: there is a sentence φ that is true in **N** but not provable in PA. And, of course, ¬φ is not provable as well, because ¬φ is not true in **N** and PA proves only sentences true in its models**.** You might attempt at adding some more "geometrical patterns of formulas" as axioms, so that to complete the theory. Providing a finite number of such 'structural relations' (i.e. axioms or axiom schemas) could be found, Hilbert's goal would be completed. Unfortunately it is not possible. Incompleteness is not a special feature of Peano arithmetic: any "reasonable" theory of arithmetic is incomplete. Though there is a *naïve* complete theory of arithmetic (called *true arithmetic*), it cannot meet the finitistic demands on involving only such procedures that make no reference either to an infinite number of structural properties of formulas or to an infinite number of operations on formulas. To state these results more precisely, we have to define what is meant by a reasonable theory:

A theory T is *recursively axiomatized* if there is an algorithm[17] that for any formula φ decides whether φ is an axiom of the theory or not.

We also need a notion of arithmetic soundness: A theory T is *arithmetically sound,* if all arithmetic sentences provable in T are valid in **N**.

Gödel's first theorem on incompleteness: let T be a theory that contains Q (i.e., the language of T contains the language of arithmetic and T proves all the axioms of Q). Let T be recursively axiomatized and arithmetically sound. Then T is an *incomplete theory,* i.e., there is a sentence φ that is not decidable in T: T proves neither φ nor ¬φ.

Note: Actually, Gödel proved the theorem on the assumption of the theory being *omega-consistent.* Omega-consistency is a technical concept which applies to a theory T if, for no property P, (i) T proves the general proposition that there exists some natural number with the property P, but (ii) for every specific natural number n, T proves that n does not have the property P. This is mostly of technical interest, since all true formal theories of arithmetic, i.e., theories the axioms of which are true in **N**, are omega-consistent. Note that omega-consistency implies consistency, but not vice versa. Later J. Barkley Rosser strengthened the theorem and proved that the assumption on T being arithmetically sound (or **N** being a model of T) could be weakened by the assumption on consistency of T.

It should be clear now what Gödel proved: it is not possible to find a recursively axiomatized consistent theory, in which all the true arithmetic sentences about natural numbers could be proved. Feasibility of a theory certainly involves the ability to recognize whether a formula is an axiom or not, i.e., the axioms of the theory have to be recursively defined; otherwise we could not execute the proof. Hence, completeness of arithmetic and recursive axiomatization are two distinct goals which cannot be met both together.

In what follows we just outline the main ideas of the proof. First we have to introduce Gödel's method of *arithmetization of metamathematics.* Well-formed formulas are sequences of symbols, proofs are sequences of formulas, and the set of these sequences is countable. Hence it is possible to define an unambiguous *numbering* of all the formulas and proofs (expressed in the language of a recursively defined theory T). Gödel defined a one-to-one

[17] We use the notion of algorithm in an intuitive way here: a finite procedure that for any input formula φ gives a "Yes / No" output in a finite number of steps. Due to Church's thesis it can be explicated by any *computational model*, e.g., Turing machine.

mapping gn (Gödel's numbering) assigning to each formula φ and to each proof d (in the theory T) a natural number gn(φ), gn(d), respectively. Moreover, Gödel's definition of the mapping is *effective*: there is an *algorithm* that calculates the value of gn at each formula or proof, and there is also an algorithm that to each Gödel's number calculates its inverse syntactic image.

The technique of numbering is not important. Any well-defined effective one-to-one mapping can serve the goal. Therefore we will use notation $<\varphi>$ for a code of a formula φ. However, what matters is the fact that due to unambiguous coding of syntactic objects by natural numbers formulas and other syntactic objects can be identified with natural numbers, and sets of formulas can be considered as sets of natural numbers. Thus, for instance, we can ask whether a set of formulas is recursive.

To remember basic notions of the theory of recursive functions, we briefly recapitulate: (partial) *recursive function*s are exactly those functions that are algorithmically computable. A set S is *recursively enumerable* if there is a partial recursive function *f* such that S is a domain of *f*: Dom(*f*) = S. A set S is a *(general) recursive set* if its characteristic function is a (total) recursive function.

We will also need the notion of a *set definable by a formula:* Let D = $\langle$D, …$\rangle$ be an interpretation structure for a language L. We say that a formula $\varphi(x_1,\ldots,x_k)$ of L, where variables $x_1,\ldots,x_k$ are free in φ, *defines a set A in* D if *A* is the set of those *k*-tuples of elements $a_1,\ldots,a_k$ of D (i.e., $A \subseteq D^k$) for which the formula φ is satisfied: D $\models \varphi(x_1,\ldots,x_k)[e]$ for a valuation *e* that assigns elements $a_1,\ldots,a_k$ to variables $x_1,\ldots,x_k$.

Note that for any arithmetic formula $\varphi(x_1,\ldots,x_k)$ the definability of a set A in the standard model **N**, i.e., the condition **N** $\models \varphi(x_1,\ldots,x_k)[e]$, where *e* assigns numbers $n_1,\ldots,n_k$ to variables $x_1,\ldots,x_k$, is equivalent to **N** $\models \varphi(\underline{n}_1,\ldots,\underline{n}_k)$.

The second ingredient of the incompleteness proof is the *Σ-completeness* of the theory Q. We are going to prove that PA and any (recursively axiomatizable) stronger 1[st]-order theory are incomplete. On the other hand there is a class of Σ-formulas such that each Σ-sentence true in **N** is provable in Q. An important property of the class of Σ-formulas is its exact correspondence to algorithmic computability: Σ-formulas define just all the algorithmically computable, i.e., recursively enumerable sets of natural numbers. Now the set Thm(T) of Gödel's numbers of those formulas that are provable in T (theorems of T) is definable by a Σ-formula[18], which Gödel denoted by *Bew*(x) – "beweisbar", we will use *Pr*(x). Hence: φ is provable in T if and only if $<\varphi> \in$ Thm(T), i.e., the sentence *Pr*($\underline{<\varphi>}$) is valid in **N,** which is equivalent to **N** $\models$ *Pr*($\underline{<\varphi>}$), where $\underline{<\varphi>}$ is the numeral denoting Gödel's number of φ.

The third ingredient is *Gödel's diagonal lemma*: For any formula $\psi(x)$ of the arithmetic language with one free variable there is a sentence φ such that $\varphi \equiv \psi(\underline{<\varphi>})$ is provable in Q.

Hence the equality Q $\vdash \varphi \equiv \psi(\underline{<\varphi>})$ with one unknown sentence φ has always, for any ψ, a solution, and the solution is independent of coding. We could say in a rather metaphoric way that φ says "I have a property ψ". The proof of the lemma also contains the self-reference element, and it is not difficult. However, particular non-trivial applications of the self-reference lemma are targets of importance. What matters is a proper choice of the formula $\psi(x)$ so that the equality Q $\vdash \varphi \equiv \psi(\underline{<\varphi>})$ had a solution φ with some interesting properties.

---

[18] This fact follows from the recursive axiomatization of T. Roughly: since the set of axioms is algorithmically computable, so is the set of proofs and so is the set Thm(T) of formulas provable in T. Hence Thm(T) is a recursively enumerable set, which implies that Thm(T) is definable by a Σ-formula.

The authors of particular non-trivial applications are thus known as the authors of self-reference formulas.

An interesting self-reference application has been proposed by Alfred Tarski. He raised a question whether it is possible to reproduce Epimenides Liar paradox in arithmetic, i.e., to find a sentence claiming "I am not true". Though it is possible to define Truth_n (in the standard model N) for some subsets of formulas, a uniform definition of Truth for *all* arithmetic formulas is impossible. There is *no* arithmetic formula $\mathrm{Tr}(x)$ that would define the set Th(N)[19] of coding numbers of formulas true in N (*true arithmetic*)**.** Tarski statement can be formulated more generally:

Let T be any consistent theory containing Q. Then there is no arithmetic formula $\mathrm{Tr}(x)$ such that for any arithmetic formula φ it holds that $T \vdash \varphi \equiv \mathrm{Tr}(\underline{<\varphi>})$.

Proof: Suppose that $\mathrm{Tr}(x)$ exists. Then according to the diagonal lemma for the formula $\neg\mathrm{Tr}(x)$ there is a sentence ω such that $Q \vdash \omega \equiv \neg\mathrm{Tr}(\underline{<\omega>})$. Since the theory T contains Q, it also holds that $T \vdash \omega \equiv \neg\mathrm{Tr}(\underline{<\omega>})$. Since the *disquotation scheme*[20] is valid for the formula $\mathrm{Tr}(x)$, we have $T \vdash \omega \equiv \mathrm{Tr}(\underline{<\omega>})$. It follows from both the equivalences that T proves $\omega \rightarrow \neg\omega$ and $\neg\omega \rightarrow \omega$, which means that T proves $\omega \;\&\; \neg\omega$. This contradicts the assumption on consistency of T.

Tarski statement is known as the impossibility to define Truth in a theory. In particular it means that there is no formula $\mathrm{Tr}(x)$ such that for any sentence φ the following equivalence would hold: $N \models \varphi$ if and only if $N \models \mathrm{Tr}(\underline{<\varphi>})$.

The diagonal self-reference lemma can be, of course, also applied to a formula $\psi(x)$ that is known to exist. Gödel's sentence claims "I am not provable", Rosser's sentence says that "each my proof is preceded by a smaller proof of my negation". Hence the last idea of Gödel's incompleteness proof is: apply the diagonal lemma on the formula $\neg Pr(x)$. We obtain *Gödel's diagonal formula* ν such that $Q \vdash \nu \equiv \neg Pr(\underline{<\nu>})$). Thus we have:

ν iff $<\nu> \notin \mathrm{Thm}(T)$ iff ν is not provable in T.

This reminds us of the Liar paradox: the sentence claiming "I am not true" is neither true nor false. Gödel was inspired by such diagonal paradoxes. However, we have to keep in mind that there is a substantial distinction: whereas Epimenides' sentence cannot be even expressed in the language of arithmetic[21], the formula ν can be constructed as a well-formed formula of the language.

Now ν is independent of T. It is true in N but not provable in T: indeed, if it were provable in T, then the formula $Pr(\underline{<\nu>})$ would be true in N. This formula is however a Σ-formula, which means that it is provable in Q and thus in T as well. Now $Pr(\underline{<\nu>}) \equiv \neg\nu$, which means that $\neg\nu$ is provable in T. We have derived that both ν and $\neg\nu$ are provable, which means that T is inconsistent. But it is not, because it has a model N**.** We have to refute the assumption that ν is provable. Hence $\neg Pr(\underline{<\nu>})$ is true in N and ν is true in N**,** but ν is not provable: T is not complete, it does not demonstrate all the truths of arithmetic.

To make these results more comprehensive, we now briefly recapitulate the main steps of the whole argument:

[19] So we use the notation Th(N) – **th**eory (for true arithmetic) and Thm(T) – for the set of **th**eore**m**s of T.

[20] The scheme is known as "It is snowing if it is snowing": the sentence ω is true in N iff its code is an element of the set defined by $\mathrm{Tr}(x)$: $\omega \equiv \mathrm{Tr}(\underline{<\omega>})$.

[21] The set Th(N) of numbers that encode true sentences of arithmetic is not definable by any formula φ.

1. A theory is *adequate* if it encodes finite sequences of numbers and defines sequence operations such as concatenation. An arithmetic theory such as Peano arithmetic (PA) is adequate (so is, e.g., a set theory).
2. In an adequate theory T we can encode the syntax of terms, sentences (closed formulas) and proofs. This fact means that we can ask which facts about provability in T are provable in T itself. Let us denote the code of φ as <φ>.
3. *Self-Reference (diagonal) lemma*: For any formula φ(*x*) (with one free variable) in an adequate theory there is a sentence ψ such that ψ iff φ(<ψ>).
4. Let Th(N) be the set of numbers that encode true sentences of arithmetic (i.e. formulas true in the standard model of arithmetic), and Thm(T) the set of numbers that encode sentences provable in an adequate (sound) theory T. Since the theory is sound, the latter is a subset of the former: Thm(T) $\subseteq$ Th(N). It would be nice if they were the same; in that case the theory T would be complete.
5. No such luck if the theory T is recursively axiomatized, i.e., if the set of axioms is *computable* in the following sense: there is an algorithm that given an input formula φ the algorithm computes a Yes / No answer to the question whether φ is an axiom or not. Computability of the set of axioms and completeness of the theory T are two goals that cannot be met together, because:
   5.1. The set Th(N) is *not even definable* by an arithmetic sentence (that would be true if its number were in the set and false if not): Let *n* be a number such that $n \notin$ Th(N). Then by the Self Reference (3) there is a sentence φ such that <φ> = *n*. Hence φ iff <φ> $\notin$ Th(N) iff φ is not true in N iff not φ – contradiction. There is no such φ. Since undefinable implies uncomputable there will never be a program that would decide whether an arithmetic sentence is true or false (in the standard model of arithmetic).

   5.2. The set Thm(T) is definable in an adequate theory, say Q: for any formula φ the number <φ> is in Thm(T) iff φ is provable, for: the set of axioms is recursively enumerable, i.e., computable, so is the set of proofs that use these axioms and so is the set of provable formulas and thus so is the set Thm(T). Since computable implies definable in adequate theories, Thm(T) is definable. Let *n* be a number such that $n \notin$ Thm(T). By the Self Reference (3) there is a sentence φ such that <φ> = *n*. Hence φ iff <φ> $\notin$ Thm(T) iff φ is not provable. Now if φ is false then φ is provable. This is impossible in a sound theory: provable sentences are true. Hence φ is true but improvable.

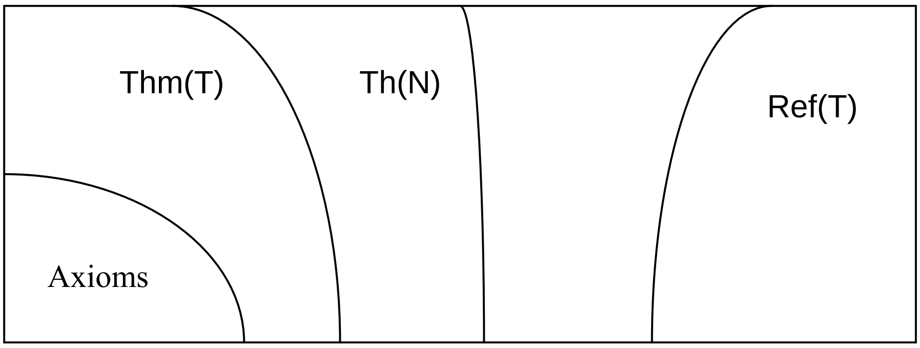
Now you may wonder: if we can algorithmically generate the set Thm(T)*,* can't we obtain all the true sentences of arithmetic? Unfortunately, we cannot. No matter how far shall we generate we will never reach all of them; there is no algorithm that would decide every formula, and there will always remain independent true formulas. We define:

A *theory T is decidable* if the set Thm(T) of formulas provable in T is (generally) recursive.

If a theory is recursively axiomatized and complete, then it is decidable. However, one of the consequences of Gödel's incompleteness theorem is:

No recursively axiomatized theory T that contains Q and has a model N, is *decidable*: there is no algorithm that would decide every formula φ (whether it is provable in the theory T or not). For, if we had such an algorithm, we could use it to extend the theory so that it were complete, which is impossible if the theory T is consistent (according to Rosser's improvement of Gödel's first theorem).

Denoting Ref(T) the set of all the sentences refutable in the theory T (i.e. the set of all the sentences φ such that T |– ¬φ), it is obvious that also this set Ref(T) is not recursive. Now we can illustrate mutual relations between the sets Thm(T), Th(N), and Ref(T) by the following figure[22]:



If the (consistent) theory T is recursively axiomatized and complete, the sets Thm(T), Th(N) coincide, and Ref(T) is a complement of them.

Another consequence of the Incompleteness theorem is the undecidability of the problem of logical truth: The FOPL proof *calculus* is a theory without special axioms. Though it is a complete calculus (all the logically valid formulas are provable), as an "empty" *theory* it is *not decidable*: there is no algorithm that would decide any formula φ whether it is a theorem or not (which equivalently means whether it is a logically valid formula or not). The *problem of logical truth is not decidable* in FOPL. For, Q is an adequate theory with a finite number of axioms. If $Q_1,\ldots Q_7$ are its axioms (closed formulas), then a sentence φ is provable in Q iff $(Q_1 \,\&\, \ldots \,\&\, Q_7) \rightarrow \varphi$ is provable in the FOPL calculus[23], and so $(Q_1 \,\&\, \ldots \,\&\, Q_7) \rightarrow \varphi$ is a logically valid formula. If the calculus were decidable so would be Q, which is not.

Alonzo Church proved that the proof calculus is *partially decidable*: there is an algorithm, which at an input formula φ that *is* logically valid outputs the answer *Yes.* If however the input formula φ is not a logical truth the algorithm may answer *no* or it can never output any answer.

Gödel discovered that the sentence ν claiming "I am not provable" is equivalent to the sentence τ claiming "There is no <φ> such that both <φ> and <¬φ> are in Thm(T)". The latter is a formal statement that the system is consistent. Since ν is not provable, and ν and τ are equivalent, τ is not provable as well. Thus we have:

Gödel's Second Theorem on incompleteness. In any consistent recursively axiomatizable theory T that is strong enough to encode sequences of numbers (and thus the syntactic notions of "formula", "sentence", "proof") the consistency of the theory T is not provable in T.

The second incompleteness theorem shows that there is no hope of proving, e.g., the consistency of the first-order arithmetic using finitist means provided we accept that finitist means are correctly formalized in a theory the consistency of which is provable in PA. As Georg Kreisel remarked, it would actually provide no interesting information if a theory T proved its consistency. This is because inconsistent theories prove everything, including their consistency. Thus a consistency proof of T in T would give us no clue as to whether T really is consistent.

[22] See Švejdar 2002

[23] Here we use a Theorem of deduction: $Q_1 \,\&\, \ldots \,\&\, Q_n \mid\!- \varphi$ iff $Q_1 \,\&\, \ldots \,\&\, Q_{n-1} \mid\!- Q_n \rightarrow \varphi$

One of the first to recognize the revolutionary significance of the incompleteness results was John von Neumann (Hungarian-born brilliant mathematician) who even almost anticipated Gödel’s *second* theorem on incompleteness. Others were slower to absorb the essence of the problem and to accept its solution. For example, Hilbert’s assistant Paul Bernays had difficulties with technicalities[24] of the proof that were cleared up only after repeated correspondence. Gödel’s breakthrough even drew sharp criticism, which was due to prevailing conviction that mathematical thinking can be captured by laws of pure symbol manipulation, and due to inability to make the necessary distinctions involved, such as that between the notion of *truth* and *proof*. Thus, for instance, the famous set-theorist Ernst Zermelo interpreted the latter in a way that leads to a pure contradiction with Gödel’s results.

3.2. Research after Gödel

Let me close this section by making a remark that Gödel incompleteness theorems, especially the celebrated 2nd Incompleteness Theorem, not only is a brilliant result which logicians are proud of and which can be reflected philosophically; it also plays the role of a useful technical tool for proving theorems about meta-mathematics of axiomatic theories. Therefore I would like to mention the fact that an interesting research inspired by Gödel continued also after Gödel[25].

Since no reasonable axiomatic theory T can prove its own consistency, a theory S capable of proving the consistency of T can be viewed as considerably stronger than T. Of course, considerably stronger implies non-equivalent. The Levy Reflection Principle, which is non-trivial but also not so difficult to prove, says that Zermelo-Fraenkel set theory ZF proves consistency of each of its finitely axiomatized sub-theories. So by Gödel 2nd Incompleteness Theorem, full ZF is considerably stronger than any of its finitely axiomatized fragments. This in turn yields a simple proof that ZF is not finitely axiomatizable. The same, with a similar but a little bit more complicated proof, is true of PA. Also ZF proves the consistency of PA.

As to the research after Gödel, I want to mention the Gentzen consistency proof, Pudlak’s extensions of Gödel 2nd Incompleteness Theorem and the connections of Gödel Theorem to modal logic.

Gerhard Gentzen, around 1940, raised the following question: once we know that consistency of Peano arithmetic PA cannot be proved in PA itself but can be proved in ZF, what exactly of all the set-theoretical machinery is necessary to prove consistency of PA? Gentzen’s answer was: all we need is to know that the (countable) ordinal ε_0, defined as the limit or ordinals 1, ω, ω^ω, ... is well founded, i.e., it is not a majorant of an infinite decreasing sequence of ordinals.

Pavel Pudlak in 1980’s showed that Gödel 2nd Incompleteness Theorem holds also for very weak fragments of PA, and if carefully (re)formulated, it holds for the Robinson’s arithmetic too. He also proved a quantitative version of Gödel 2nd Incompleteness Theorem, saying that statements of the form "there is no proof of contradiction the length of which is less than *n*”, while provable in T if T satisfies usual requirements, only have proofs the lengths of which grow very rapidly with *n*.

Provability logic is a modal propositional logic where the usual modal operator $\Box$ (Box) is interpreted as formal provability in some fixed axiomatic theory extending arithmetic. Then e.g. $\neg\Box$False (not provable False) is a modal formula which can be read “contradiction is unprovable in T”. Its arithmetical counterpart is the formalized consistency statement which, by Gödel 2nd Incompleteness Theorem, cannot be proved in T. So $\neg\Box$False is an example of

[24] The technical device used in the proof is now known as Gödel numbering
[25] These remarks were formulated by Vítězslav Švejdar. I am deeply grateful for them.

a non-tautology of provability logic. Since $\neg\Box$False is a formula equivalent to $\Box$False $\rightarrow$ False (provable False implies False), this example also shows that the scheme $\Box$A $\rightarrow$ A is not an acceptable axiom scheme for provability logic. On the other hand, the formula $\neg\Box$False $\rightarrow$ $\neg\Box\neg\Box$False ("if a contradiction is not provable then the statement that a contradiction is unprovable is unprovable") is an example of a tautology of provability logic. Provability logic was investigated in parallel by several researchers (in Amsterdam, Italy, U.S.A., Sweden). Its arithmetical completeness theorem was proved by R. Solovay in 1975. Provability logic is interesting for both mathematicians and philosophers; it combines metamathematical investigations with modal-logical tools like Kripke semantics. Out of several papers dealing with provability logic I can recommend Švejdar's (2000), from which some explanations and symbolism above are taken.

There are some extensions of provability logic obtained by employing additional "modalities". For example, interpretability logic uses, besides $\Box$ for "proves", a binary modality operator ► for "interprets". It is designed for research on (*syntactic*) *interpretability* of axiomatic theories. The concept of syntactic interpretability is distinct from the concept of semantic interpretation introduced above. Slightly simplified, a theory T is said to be *interpretable* in a theory S iff the language of T can be translated into the language of S in such a way that S proves the translation of every theorem of T. Of course, there are some natural conditions on admissible translations here, such as the necessity for a translation to preserve the logical structure of formulas. This concept, together with weak interpretability, was introduced by Alfred Tarski in 1953. However, intensive research has been pursued in Prague; one of the important results that stimulated the interest in interpretability was Petr Hájek's observation that Zermelo-Fraenkel set theory ZF and Gödel-Bernays set theory GB, though identical as to provability of set sentences, differ in interpretability.

Connections of Gödel theorems to interpretability logic are given by the following two facts. First, there is a generalization of the Second Incompleteness Theorem saying that a consistent theory T can interpret no theory S such that S is an extension of T and S proves consistency of T. Second, if S proves consistency of T then S interprets T. This second fact can be obtained by formalizing Gödel completeness theorem in S.

4. Concluding remarks

Now I would like to mention some properties of arithmetic models. From the Compactness theorem[26] it can be easily derived that there are *non-standard models*[27] of a recursively-formalized arithmetic. A non-standard model is one that constitutes a structural interpretation of the formal theory that is *admittedly different* from the intended one. By structural interpretation I mean interpretation where isomorphic models count as the same interpretation.

The existence of non-standard models can be also derived from the stronger version of the Completeness theorem. Roughly: a formula φ is provable in a theory T iff φ is logically entailed by its special axioms; T $\models \varphi$ iff T $\vdash \varphi$. Now the sentence ν is not provable in T, hence ν is not valid in every model of T. It is however valid in the standard model N, which is a model of T. Every model isomorphic to N is also a model of T; ν is however not valid in every model of T. Hence T must have a non-standard model.

[26] If a formula φ is logically entailed by a theory T (T $\models \varphi$), then there is a finite subset F of T such that φ is entailed by F (F $\models \varphi$).

[27] See a nice exposition on non-standard models by Haim Gaifman (2003)

From the point of view of capturing the intended interpretation, i.e., characterizing the set of natural numbers completely, the existence of non-standard models counts as a failure of the formal language to capture the semantics fully. The special axioms of the theory do not "implicitly define" the intended model, the consistency problem becomes crucial. Ordinary mathematical practice amounts to a study of the 'intended interpretation'. But if mathematics is not only a "science of quantity" but a fully formalized discipline that draws conclusions logically implied by any given set of axioms, and if a mathematical inference in no sense depends upon any special meaning that may be associated with the terms and formulas, the question whether the given set of axioms is internally consistent so that no contradictory theorems can be derived, becomes crucial. If the axioms are simultaneously true of some sequences of numbers, they cannot be incompatible. But the models of arithmetic are composed of infinite number of elements, which makes it impossible to encompass the models in a finite number of observations; hence the truth of the axioms themselves is a subject of doubt. Using the axiom of induction we can only check that a finite number of objects are in the agreement with the axiom. But the conclusion involves an extrapolation from a finite to an infinite set of objects. Hence Hilbert sought an 'absolute' proof of consistency. Unfortunately no such absolute proof will ever be at hand.

You may pose another question: *Which of the models is the standard one*? Which sequence of objects constitutes the subject matter of the inquiry, "what is it all about"? It can be characterized by a *minimality* condition[28]: it is the smallest model, included as an initial segment in any other model. If the model is non-standard, then it will be revealed by a proper initial segment that is closed under the successor function. Formally, the characterization is expressed by the inductive scheme:

(I) $[\mathrm{P}(\underline{0}) \mathbin{\&} \forall x\,(\mathrm{N}(x) \rightarrow (\mathrm{P}(x) \rightarrow \mathrm{P}(\mathrm{S}x))] \rightarrow \forall x\,[\mathrm{N}(x) \rightarrow \mathrm{P}(x)]$,

where $\mathrm{N}(x)$ stands for 'x is a natural number', and 'P( )' stands for any predicate. Any wff of the language can be substituted for 'P( )'. The concept of the natural number sequence is however *not* language dependent. The absoluteness of the concept can be secured, if we help ourselves to the standard power set of some infinite set; for then we can treat 'P' as a *variable* ranging over that power set. In other words, we shift the system into the 2nd order. But this is highly unsatisfactory. Quoting from Gaifman (2003):

> […] it bases the standard number sequence on the much more problematic shaky concept of the standard power set. It is, to use a metaphor of Edward Nelson, like establishing the credibility of a person through the evidence of a much less credible character witness. The inductive scheme should be therefore interpreted as an open-ended meta-commitment:
>
> (II) Any non-vague (crisp) predicate, in whatever language, can be substituted for 'P' in (I).
>
> As Van McGee expresses it, if God himself creates a predicate, then this predicate can be substituted for 'P'. One, who has reservations about actual infinities, can still doubt the conception of the standard number sequence, but these doubts do not gain additional support from the existence of non-standard models.

The second-order theories (of real numbers, of complex numbers, and of Euclidean geometry) do *have complete axiomatizations*. Hence these theories have no true but unprovable sentences. The reason they escape the incompleteness is their inadequacy: they can't encode and computably deal with finite sequences. The price we pay for the 2nd-order completeness is high; the second-order calculus is *not (even partially) decidable*. We cannot

[28] Now I refer to H.Gaifman (2003).

algorithmically generate all the 2nd-order logical truths, thus not all the logical truths are provable, the 2nd-order calculus is *not semantically complete*.

The consequences of Gödel's two theorems are clear and generally accepted. First of all, the formalist belief in identifying *truth* with *provability* is destroyed by the first incompleteness theorem. Second, the impossibility of an absolute (acceptable from the finitist point of view) consistency proof is even more destructive for Hilbert's program. The second Gödel's theorem makes the notion of a finitist statement and finitist proof highly problematic. If the notion of a finitist proof is identified with a proof formalized in an axiomatic theory T, then the theory T is a very week theory. If T satisfies simple requirements, then T is suspected of inconsistency. In other words, if the notion of finitist proof means something that is non-trivial and at the same time non-questionable and consistent, there is no such thing.

Though it is almost universally believed that Gödel's results destroyed Hilbert's program, the program was very inspiring for mathematicians, philosophers and logicians. Some thinkers claimed that we should be formalists anyway[29]. Others, like Brouwer, the father of modern constructive mathematics, believe that mathematics is first and foremost an activity: mathematicians do not discover pre-existing things, as the Platonist holds and they do not manipulate symbols, as the formalist holds. Mathematicians, according to Brouwer, make things. Some recent intuitionists seem to stay somewhere in between: being ontological realists they admit that there are abstract entities we discover in mathematics, but at the same time being semantic intuitionists they claim that these abstract entities "do not exist" unless they are well defined by a constructive[30] formal proof, as a sequence of judgements[31].

Possible impact of Gödel's results on the philosophy of mind, artificial intelligence, and on Platonism might be a matter of dispute. Gödel himself suggested that the human mind cannot be a machine and that Platonism is correct. Most recently Roger Penrose has argued that "the Gödel's results show that the whole programme of artificial intelligence is wrong, that creative mathematicians do not think in a mechanic way, but that they often have a kind of insight into the Platonic realm which exists independently from us"[32]. Gödel's doubts about the limits of formalism were certainly influenced by Brouwer who criticized formalism in the lecture presented at the University of Vienna in 1928. Gödel however did not share Brouwer's intuitionism based on the assumption that mathematical objects are created by our activities. For Gödel as a Platonic realist mathematical objects exist independently and we discover them. On the other hand he claims that our intuition cannot be reduced to Hilbert's concrete intuition on finitary symbols, but we have to accept abstract concepts like *well defined mathematical procedures* that have a clear meaning without further explication. His proofs are constructive and therefore acceptable from the intuitionist point of view.

In fact, Gödel's results are based on the two fundamental concepts: *truth for formal languages* and *effective computability.* Concerning the former, Gödel stated in his lectures in Princeton that he was led to the incompleteness of arithmetic via his recognition of the non-definability of arithmetic truth in its own language. In the same lectures he offered the notion of *general recursiveness* in connection with the idea of *effective computability*; this was based on a modification of a definition proposed by Herbrand. In the meantime, Church was making a proposal of his thesis, which identified the effectively computable functions with the λ-definable functions. Gödel was not convinced by Church's thesis, because it was not based on a conceptual analysis of the notion of *finite algorithmic procedure.* It was only when Turing,

[29] See, e.g., Robinson, A. (1964) 'Formalism 64', or more recently Detlefsen, M. (1992) 'On an Alleged Refutation of Hilbert's Program Using Gödel's first incompleteness theorem'

[30] The notion of constructive proof is central for intuitionistic logic.

[31] The above is a slightly reformulated remark made by Peter Fletcher in an e-mail correspondence.

[32] See, Brown (1999. p. 78)

in 1937, offered the definition in terms of his *machines* that Gödel was ready to accept the identification of the various classes of functions: λ-definable, general recursive, Turing computable.

Pursuit of Hilbert's program had thus an unexpected side effect: it gave rise to the realistic research on the theory of recursive functions and algorithms. John Von Neumann, for instance, along with being a great mathematician and logician, was an early pioneer in the field of modern computing, though it was a difficult task because computing was not yet a respected science. His conception of computer architecture has actually not been surpassed till now. Gödel's first theorem has another interpretation in the language of computer science. In first-order logic, theorems are recursively enumerable: you can write a computer program that will eventually generate any valid proof. You can ask if they satisfy the stronger property of being recursive: can you write a computer program to definitively determine if a statement is true or false? Gödel's theorem says that in general you cannot; a computer can never be as smart as a human being because the extent of its knowledge is limited by a fixed set of axioms, whereas people can discover unexpected truths.

The greatness of a great thinker is measured by the influence his work has on future generations. One can fairly say that Gödel's results changed the face of meta-mathematics and influenced all aspects of modern mathematics, artificial intelligence and philosophy of mind.

Marie Duží
VSB-Technical University Ostrava
Department of Computer Science FEI
17. listopadu 15
708 33 Ostrava
marie.duzi@vsb.cz

References

Brown, J. (1999): *Philosophy of Mathematics.* Routlege, London, New York.
Dawson, J.W. (1984): Kurt Gödel in Sharper Focus. *The Mathematical Intelligencer,* 4.
Feferman, S. (1960): Arithmetization of metamathematics in a general setting, *Foundations of Math.*, vol. 49, pp. 35-92.
Feferman, S., ed. (1986): Kurt Gödel: *Collected Works*. Oxford University Press.
Gaifman, H. (2003): *Non-Standard Models in a Broader Perspective*. Manuscript.
Hájek, P (1996): Matematik a logik. In Malina, Novotný.
Köhler, E. (1991): Gödel und der Wiener Kreis. In: *Jour fixe der Vernunft*, edited by Kruntorad, P. with the aid of Haller, R. and Hochkeppel, W, Verlag Hölder-Pichler-Tempsky, Vienna.
Köhler, E. (2002): Gödels Jahre in Princeton. In: *Kurt Gödel. Wahrheit & Beweisbarkeit*, Volume I, edited by Köhler, E., Weibel, P., Stöltzner, M., Buldt, B., Klein, C., DePauli-Schimanovich-Göttig, W., Vienna.
Köhler, E. (2002a): Gödel's Platonismus. In: *Kurt Gödel. Wahrheit & Beweisbarkeit*, Volume II, edited by Buldt, B., Köhler, E., Stöltzner, M., Weibel, P., Klein, C., DePauli-Schimanovich-Göttig, W. Vienna.
Malina, J., – Novotný, J., ed. (1996): *Kurt Gödel*. Nadace Universitas Masarykiana, Brno.
Mendelson, E. (1997): *Introduction to Mathematical Logic*. Chapman & Hall.
Nagel, E. – Newman, J.R. (1958): *Gödel's Proof.* New York University Press.

Švejdar, V. (2000): On Provability Logic. *Nordic Journal of Philosophical Logic,* Vol. 4, No. 2, December 1999.
<URL http://www.hf.uio.no/filosofi/njpl/vol4no2/provlog/index.html>
Švejdar, V. (2002): *Logika, neúplnost, složitost a nutnost.* Academia Praha.

Acknowledgements

I am deeply grateful to Vítězslav Švejdar for his valuable comments on the manuscript of the paper, which improved the quality of the survey. Vítězslav also contributed the concluding remarks on provability and interpretability logics. I am also grateful to Bjorn Jespersen and Andrew Holster for their careful proof reading and improving the English, and to Eckehard Köhler for providing relevant references.

This work has been supported by the program "Information Society" of the Czech Academy of Sciences, project No. 1ET101940420 "Logic and Artificial Intelligence for multi-agent systems".