4

Gödel's Undecidability Theorem

Author: Stephen F. Andrilli, Mathematical Sciences Department, La Salle University

Prerequisites: The prerequisites for this chapter are logic, properties of integers, and techniques of proof. See Chapter 1, and Sections 1.6 and 1.7 in particular, of *Discrete Mathematics and Its Applications*.

Introduction

In high-school geometry and in your discrete mathematics course, you encountered many types of mathematical proof techniques. These included such methods as proof by contrapositive, proof by contradiction, and proof by induction. After becoming familiar with these methods, many students think mathematicians have so much "power" at their disposal that, surely, any given mathematical statement must either already have been proven or disproven.

But this is not the case! For instance, consider the following simple statement.

> **Goldbach Conjecture:** Every even positive integer greater than 2 is the sum of two (not necessarily distinct) primes.

Note that $4 = 2 + 2$, $6 = 3 + 3$, $8 = 5 + 3$, $10 = 5 + 5$, $12 = 7 + 5$, and so on. Can we continue, expressing all larger even integers as the sum of two primes? The answer is not yet known, despite extensive work by many mathematicians.

Goldbach's Conjecture* is one of hundreds of mathematical "conjectures" whose truth value we have not been able to determine.

Now, of course, Goldbach's Conjecture is either true or false. Therefore, considering the huge battery of proof techniques available, it seems logical for us to predict with some confidence that eventually mathematicians will develop a proof that it is true or else find a counterexample to show that it is false. Until recently, the general attitude of most mathematicians was that every mathematical conjecture would, after enough effort, be resolved one way or the other. That is, it would either "become" a theorem or be identified as a false statement. However, as we will see later, the work of a brilliant 25-year-old German mathematician named Kurt Gödel** forever shattered this "belief", and led to a profound difference in the way we perceive mathematics as a whole.

Mathematical Systems: Examples from Geometry

To understand Gödel's work, we must first discuss the nature of general mathematical systems. Every statement in mathematics occurs in a *context*. This context might not always be stated directly, but is often implicit from previous statements. For example, a theorem in a high-school geometry can only be understood fully when we know what the terms in the theorem mean. These terms may have been *defined* previously in the text, or they may have been accepted as **undefined terms**. Undefined terms are terms that are so fundamental that we cannot fully describe them using other more basic terms. Terms such as "point","line", "equal", and "between" are typical examples of undefined terms used in many texts.

Sometimes we cannot prove a theorem without the help of previous theorems. We may even need to use some *axioms* (sometimes called *postulates*). **Axioms** are statements that we accept as true, because they are so fundamental that we cannot prove them from other more basic results. In high-school geometry, most axioms are based on common-sense notions. For example, a

* The Goldbach Conjecture appeared in a letter from Christian Goldbach (1690–1764) to Leonhard Euler in 1742. Goldbach was a historian and professor of mathematics at the Imperial Academy of St. Petersburg (now Leningrad). He tutored Czar Peter II, and later worked in the Russian Ministry of Foreign Affairs. Among his research areas were number theory, infinite series, and differential equations.

** Kurt Gödel was born in Brno, Czechoslovakia in 1906. During the 1930s he taught at the University of Vienna and did research at the Institute for Advanced Study in Princeton, New Jersey. Gödel spent time in a sanatorium for treatment of depression in 1934 and again in 1936. In 1940, to escape Nazi fascism, he emigrated to the United States, and continued his work at the Institute for Advanced Study. He died in Princeton in 1978.

typical axiom in many geometry texts is: Every two distinct points lie on a unique straight line.

In any high-school geometry text, the undefined terms, definitions, axioms, and theorems, taken together, form a **mathematical system**. Most high-school geometry courses introduce students to the system of **Euclidean geometry**, so named because most of the results are essentially those handed down from classical times in *The Elements* of Euclid.* Within such a mathematical system, we can make new definitions and prove additional theorems.

Example 1 After accepting the axiom "Every two distinct points lie on a unique straight line", and the undefined terms listed above, we can introduce a new term: "midpoint". The *midpoint* between two given points is defined as the (unique) point on the line connecting the given points that divides the line segment between the points into two equal parts.

Additionally, once we have defined terms like "right triangle" and "hypotenuse", we can prove theorems such as the familiar Pythagorean Theorem: In any right triangle, the square of the hypotenuse is equal to the sum of the squares of the other two sides. □

In this way, we can achieve the development of an entire mathematical system beginning with a few simple building blocks!

Non-Euclidean Geometries

One of the axioms that Euclid adopted in *The Elements* is equivalent to the following statement, commonly known as the **Parallel Postulate**:

> **Parallel Postulate:** If l is a line in a given plane, and P is a point of the plane not on l, then in this plane there is one and only one line going through P that is parallel to l.

This seems like a perfectly reasonable assumption — so reasonable, in fact, that even into the nineteenth century many mathematicians tried to prove that it was really a consequence of earlier axioms and definitions. Their efforts were ultimately unsuccessful. They next tried to deny this postulate to determine whether the resulting mathematical system contained a logical contradiction. But instead they obtained a set of new "strange" results!

We know now that it is possible to create several valid **non-Euclidean geometries** in which the Parallel Postulate is contradicted in different ways. For

---

* Euclid, who lived about 300 B.C., taught mathematics at Alexandria during the reign of the Egyptian Pharaoh Ptolemy Soter, and was most probably the founder of the Alexandrian School of mathematics. His classic work *The Elements* contains much of the basic geometrical knowledge accumulated up to that time.

example, we could decide to include the following axiom in our mathematical system instead of the Parallel Postulate:

> **Multi-Parallel Postulate:** If l is a line in a given plane and P is a point in the plane not on l, then in the plane there is more than one line going through P parallel to l.

Adopting this axiom leads to a type of geometry known as **hyperbolic geometry**, which was discovered independently before 1830 by Nikolai Lobachevsky* and János Bolyai**. We can now prove theorems that are consequences of this new postulate. For instance, in hyperbolic geometry we can prove that, with l and P as in the statement of the Multi-Parallel Postulate, there are an *infinite* number of lines through P parallel to l. We can also show that the sum of the angles of any triangle is less than 180°.

Another type of non-Euclidean geometry is created when we replace the Parallel Postulate with:

> **No-Parallel Postulate:** If l is a line in a given plane and P is a point in the plane not on l, then in the plane there is no line going through P parallel to l.

Such a geometry is called **elliptic geometry**, and was first created by Bernhard Riemann†. In this geometry, every pair of distinct lines meets in a point. Also, the angles of any triangle sum up to more than 180°.

---

* Nikolai Lobachevsky (1792–1856) taught at the University of Kazan, Russia, beginning in 1816. He later became Rector of the University. Lobachevsky published his first results on hyperbolic geometry in 1829. Another of his research interests was developing methods for approximating roots of algebraic equations.

** János Bolyai (1802–1860) studied at the Royal College of Engineering in Vienna and served in the army engineering corps. He was an expert swordsman and a superb violinist! Bolyai's father Farkas spent his lifetime to no avail trying to prove that the Parallel Postulate was a consequence of Euclid's earlier postulates. János continued this work, but realized that it might be impossible to prove. He began constructing a non-Euclidean geometry, a first draft of which was written in 1823. Bolyai's work went unnoticed at the time, and was largely overshadowed by Lobachevsky's publication. His contribution to non-Euclidean geometry was only recognized posthumously. Aside from his work in geometry, Bolyai also broke new ground in the theory of complex variables.

† Georg Friedrich Bernhard Riemann (1826–1866) was one of the most important mathematicians of the last century. He taught at Göttingen (University), and aside from his ground-breaking work in geometry, he developed many new concepts in complex and real analysis, and what is now topology and differential geometry. His discovery of elliptic geometry dates from about 1854 and was first published in 1867. Albert Einstein later based much of his theory of relativity on the geometrical ideas of Riemann.

Consistency of Mathematical Systems

As a result of these new observations in non-Euclidean geometry, mathematicians began to realize that abstract mathematical systems (with undefined and defined terms, axioms and theorems) could be created without any natural or obvious association with "reality" as we ordinarily perceive it. But this leads to a fundamental difficulty: once we choose an arbitrary set of axioms, how do we know that the system "makes sense"? That is, how can we be sure that none of the axioms are contradicted by the others?

It is important that any mathematical system we work in be **consistent**, that is, free from self-contradiction. It can be shown that Euclidean geometry is inwardly consistent, as are both of the non-Euclidean geometries discussed earlier. However, not all sets of axioms lead to consistent systems.

Example 2 Consider this set of axioms with undefined terms "point" and "line":

(i) There are exactly three points in the system.
(ii) There are fewer than three lines in the system.
(iii) Every pair of points lies on exactly one line.
(iv) Not all points lie on the same line.

Notice that "line" as used here is a general term. A "line" does not necessarily have to agree with our "common-sense" notion of an object that consists of an infinite number of points, but may only contain a finite set of points. These axioms are not consistent, since the only way that axioms (i), (ii), and (iii) can be satisfied simultaneously is to have all three points lying on the same line. But this contradicts axiom (iv). □

One way to be sure that none of the axioms in a given mathematical system contradict the others is to exhibit a **model** for the system that satisfies all of the axioms simultaneously.

Example 3 Consider the following collection of axioms, with undefined terms "point" and "line":

(i) At least two lines exist.
(ii) Every line contains at least two points.
(iii) Given any two distinct points, there is a line containing the first point but not the second point.

Notice also that the last axiom is "symmetric", because if we reverse the order of the points, axiom (iii) insists there is a line containing the original second point that does not contain the original first point.

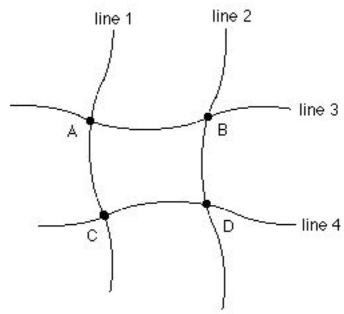

Figure 1. A model of four points and four lines.

The model of four "points" and four "lines" in Figure 1 shows that these axioms are consistent.

For example, given points A and B, line 1 goes through A but not B, while line 2 goes through B but not A. Exercise 3 asks you to find another model for this system using fewer than four points. □

Mathematicians are not interested in a system unless it is logically consistent. However, for a system that involves a large (possibly infinite) number of objects, a "model" of the system is often too difficult to construct. Such models would be extremely complicated, especially if we were trying to represent a system powerful enough to contain an entire branch of mathematics.

This can be seen even in a branch of mathematics as familiar as arithmetic. For instance, suppose we want to include an axiom that states that for every integer, there is a next largest integer. If we attempt to construct this system, then for every integer x that we list, we must list the integer y immediately following it as well. But then for this new integer y, we must list the integer z immediately following y, and so on. Thus, we would have to display an entire infinite sequence of integers, which is, of course, not physically possible. Mathematicians realized that a new approach was needed.

In 1910, two mathematicians, Alfred N. Whitehead* and Bertrand Russell** published a monumental treatise, *Principia Mathematica*. In this work,

---

* Alfred North Whitehead (1861–1947) was a philosopher and mathematician who held various academic positions in London from 1911 to 1924. From 1910–1913, he worked with Russell on *Principia Mathematica*. In 1924, he accepted a faculty post at Harvard University, and in the late 1920s he wrote a number of treatises on metaphysics.

** Bertrand Russell (1872–1970) was a philosopher and logician, who published numerous books on logic and the theory of knowledge. He was a student of Whitehead's before they collaborated on *Principia Mathematica*. Russell was active in many social causes, and was an advocate of pacifism (as early as World War I) and of nuclear disarmament.

they took a different approach to the consistency question. They attempted to derive all of arithmetic from principles of symbolic logic, which had been rigorously developed in the nineteenth century. (Symbolic logic includes such rules of inference as "modus ponens", "disjunctive syllogism", etc.) In this way, Whitehead and Russell could establish the consistency of arithmetic as a consequence of the basic laws of symbolic logic.

Resolving Logical Paradoxes

Whitehead and Russell's effort took on added importance because, at the turn of the century, mathematicians became more and more concerned about resolving **paradoxes** appearing in mathematical systems. Paradoxes are statements that are apparently self-contradictory. These paradoxes worried mathematicians because they seemed to imply that mathematics itself is ultimately inconsistent!

A typical paradox is **Russell's Paradox**. To understand this, let a **set** be described as a collection of objects (leaving "object" as an undefined mathematical term). Next, consider the set X of all sets having more than one element. Now X certainly contains more than one set. Thinking of these sets as elements of X, we see that X itself has more than one element. Hence, $X \in X$. Thus, it *is* possible to find sets that are elements of themselves. This leads us to the following.

> **Russell's Paradox:** Let S be the set of all sets that are not elements of themselves: i.e., $S = \{X \mid X \notin X\}$. Then either $S \in S$, or $S \notin S$. But either possibility leads to a contradiction. For, if $S \in S$, then by definition of S, $S \notin S$. On the other hand, if $S \notin S$, then by definition of S, $S \in S$.

In their *Principia Mathematica*, Whitehead and Russell devised methods to avoid Russell's Paradox and other similar paradoxes. They introduced the idea of "set types". They began with a fundamental collection of objects as above. These objects were considered to be of "type 0". Any set of objects of type 0 is considered to be a set of "type 1". Any set of objects of types 0 or 1 is considered to be a set of "type 2". In general, all objects in a set of "type n" must be of type less than n. Finally, only those collections of objects that have a "type" will be admitted into the system as sets. Such a scheme enabled Whitehead and Russell to avoid situations like Russell's Paradox where a set could be an element of itself.

In effect, what Whitehead and Russell did was to avoid mathematical paradoxes and "shift" the problem of the consistency of arithmetic to the consistency of a few basic logical principles. If these logical principles standing alone were consistent, then any mathematical system (in particular, arithmetic) derived solely from them would also be consistent. With this new method for checking consistency, many mathematicians felt it would only be a matter of time before other branches of mathematics would be proven consistent as well, and then, within those systems, eventually all statements could be proven or disproven.

Gödel's Undecidability Theorem

In 1931, Kurt Gödel published a paper "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems" in which he proved the following theorem.

Theorem 1 Gödel's Undecidability Theorem Any mathematical system containing all the theorems of arithmetic is an incomplete system. (That is, there is a statement within this system that is true, but can never be proven true.) ■

In other words, any system powerful enough to contain all of the rules of arithmetic, (e.g., the commutative and associative laws of addition and multiplication), must contain a mathematical statement that is true (has no counterexample), but for which *no proof can ever be found* using all of the results available in that system!

Another way of expressing Gödel's Undecidability Theorem is to say that, given a certain mathematical statement in this system, we may never learn whether it is true or false — and even if it is true, we may not ever have the resources to prove that it is true from within the system. Because of Gödel's Undecidability Theorem, we can no longer confidently predict that all open conjectures, such as Goldbach's Conjecture, will eventually be resolved one way or the other.

In his paper, Gödel produced a statement (let us call it G, for Gödel) that is true in any system similar to the system developed in *Principia Mathematica* but can never be proven true within that system. A natural question is: How do we know that G is true if we can't *prove* G is true? To explain this, we need to give some idea of Gödel's actual proof.

Gödel's actual paper is extremely complicated. The Undecidability Theorem is preceded by the establishment of several primitive objects called "signs", several variables of various "types" (as in *Principia Mathematica*), 45 definitions of functions (or relations), many other definitions, five axioms, one lemma, and seven theorems! Nevertheless, we can get the central thrust of Gödel's argument by studying the method he used for labeling items in *Principia Mathematica*. His technique is often referred to as *Gödel numbering*.

Gödel Numbering

Gödel assigned a distinct positive integer, a **Gödel number**, to each symbol, formula, and proof in the system. He assigned the first few odd integers to these seven basic symbols:

symbol	meaning	Gödel number
0	zero	1
f	successor	3
$\neg$	not	5
$\vee$	or	7
$\forall$	for all	9
(	left parenthesis	11
)	right parenthesis	13

Other useful symbols such as $\wedge$ (and), $\exists$ (there exists), and $=$ (equals) can be "built" by using an appropriate combination of the seven basic symbols above. However, for the sake of simplicity, we will depart from Gödel's actual numbering, and assign Gödel numbers to these symbols as follows:

symbol	meaning	Gödel number
$\wedge$	and	15
$\exists$	there exists	17
$=$	equals	19

Numerical variables (such as x, y, and z) that represent natural numbers are assigned larger odd values as their Gödel numbers — we assign 21 to x, and 23 to y.

Natural numbers after 0 can be represented in this system as follows:

$$\underbrace{f0}_{\text{one}},\quad \underbrace{ff0}_{\text{two}},\quad \underbrace{fff0}_{\text{three}},\quad \text{etc.}$$

Formulas of this type that contain a sequence of symbols are assigned Gödel numbers by creating a product of powers of successive primes 2, 3, 5, 7, 11, ...

Example 4 Find the Gödel number for the sequence $fff0$.

Solution: The symbols in the sequence

$$fff0$$

have Gödel numbers 3, 3, 3, 1, respectively, and therefore we assign the Gödel number

$$2^3 3^3 5^3 7^1$$

to the entire sequence $fff0$. □

Example 5 Find the Gödel number of the formula $\exists x(x = f0)$.

Solution: The symbols in the formula

$$\exists x(x = f0)$$

have Gödel numbers 17, 21, 11, 21, 19, 3, 1, 13, respectively. Hence we would assign the Gödel number

$$2^{17}3^{21}5^{11}7^{21}11^{19}13^{3}17^{1}19^{13}$$

to the formula $\exists x(x = f0)$. □

By the way, the formula $\exists x(x = f0)$ in the last example asserts that there is some natural number that is the successor of zero. Since '1' is the successor of zero, we would consider this to be a true proposition. With the method of this example, we can assign to every logical proposition (and hence every theorem) an associated Gödel number.

We can generalize this even further by associating a Gödel number with a *sequence* of propositions. If there are k propositions in a sequence with Gödel numbers $n_1, n_2, \ldots, n_k$, then we assign the Gödel number $2^{n_1}3^{n_2}5^{n_3}\ldots p_k^{n_k}$ (where p_k is the kth prime in the list 2, 3, 5, 7, 11, ...) to this sequence.

Example 6 Find the Gödel number of the sequence

$$S = \left\{ \begin{array}{c} x = f0 \\ \neg(f0 = 0) \\ \neg(x = 0) \end{array} \right\}.$$

Solution: We first find the Gödel number of each individual proposition in the sequence. Now,

the Gödel number of $x = f0$ is $k_1 = 2^{21}3^{19}5^{3}7^{1}$,
the Gödel number of $\neg(f0 = 0)$ is $k_2 = 2^{5}3^{11}5^{3}7^{1}11^{19}13^{1}17^{13}$, and
the Gödel number of $\neg(x = 0)$ is $k_3 = 2^{5}3^{11}5^{21}7^{19}11^{1}13^{13}$.

Now, the Gödel numbers of these individual propositions are placed as exponents on the primes 2, 3, 5, respectively, to obtain the Gödel number

$$2^{k_1}3^{k_2}5^{k_3}$$

(an incredibly large number) for the entire sequence S. □

Note that, in Example 6, if the first two statements in the sequence S have already been proven (or accepted as axioms), then the third statement follows logically from the first two. Hence, if $x = f0$, we can consider the sequence S to be a *proof* of the statement $\neg(x = 0)$. In this way, every proof in this mathematical system can be assigned a Gödel number.

In this manner, every basic symbol, variable, sequence of symbols and variables (e.g., proposition), and sequence of propositions (e.g., proof) that can be expressed in a system such as *Principia Mathematica* has its own *unique* Gödel number. Conversely, given any Gödel number, we can reconstruct the unique sequence of symbols corresponding to it. (This follows from the Fundamental Theorem of Arithmetic, which asserts that each positive integer has a *unique* expression as a product of primes, where the primes are listed in increasing order.) In other words, there is a one-to-one correspondence between valid sequences of symbols and valid Gödel numbers in our *Principia Mathematica*-like system.

A Few Additional Definitions

Suppose we consider all the possible formulas in our system having exactly one variable (say, x), and put them in increasing order according to Gödel number. Let $R(n)$ represent the nth such Gödel formula.

Gödel became interested in what would happen if the particular sequence of the form $fff \cdots f0$ representing the natural number n were substituted in place of the variable x in the nth Gödel formula $R(n)$. This produces a new formula, which we label as

$$\mathrm{SUBST}[n, R(n)],$$

where "SUBST" stands for "substitution". (Gödel actually used $[R(n); n]$ in his paper to represent this new formula.) This new formula may or may not be *provable* within our mathematical system. That is, there may or may not be some sequence of previously established propositions that logically proves the new formula $\mathrm{SUBST}[n, R(n)]$.

Suppose, for a particular n, the Gödel formula $\mathrm{SUBST}[n, R(n)]$ is provable. We express this by writing

$$\mathrm{PR}(\mathrm{SUBST}[n, R(n)]),$$

where "PR" stands for "provable". (Gödel expressed this as $\mathrm{Bew}[R(n); n]$.)

The Heart of Gödel's Argument

Consider the set Z of all integers n for which $\mathrm{PR}(\mathrm{SUBST}[n, R(n)])$ is true — that is, for which there exists a finite sequence of propositions to prove the statement obtained when the sequence $fff \cdots f0$ for n is substituted for the variable x in the nth Gödel formula. (In his paper, Gödel actually used the *complement* of this set as the set Z.) Now, a given integer n is either in Z or it is not. Consider the formula $\neg(x \in Z)$ (i.e., it is not true that $x \in Z$). It is not immediately obvious that this formula involving x has a Gödel number, but Gödel showed that there is a way of expressing the formula $x \in Z$ in terms of more basic symbols in the system. Therefore, let us suppose that $\neg(x \in Z)$ has a Gödel number, and that it is the qth Gödel formula (in the ordering of single-variable Gödel formulas discussed earlier). Then, $\neg(x \in Z)$ is $R(q)$.

Finally, let us consider $\text{SUBST}[q, R(q)]$ — the new formula obtained by putting the appropriate sequence of symbols of the form $fff \cdots f0$ for q into the formula $R(q)$. We will call this new formula G.

Now, either G is true or $\neg G$ is true. However,

Theorem 2 Neither G nor $\neg G$ is a provable statement.

Proof: We give a proof by contradiction. Suppose first that G is a provable statement. Then, $\text{PR}(\text{SUBST}[q, R(q)])$ is true — and so by definition of Z, we have $q \in Z$. But upon actually substituting q into $R(q)$, we obtain the statement $\neg(q \in Z)$, which is supposed to be provable. Since we cannot have both $q \in Z$ and $\neg(q \in Z)$ provable in our system, we have a contradiction.

On the other hand, suppose $\neg G$ is a provable statement. Then G is certainly not true, and therefore not provable. Hence, $\text{PR}(\text{SUBST}[q, R(q)])$ is not true, and so by definition of Z, it follows that $q \notin Z$. That is, $q \notin Z$ is a provable statement. But, since G is not provable, substituting q into $R(q)$ gives us the unprovable statement $\neg(q \in Z)$. Since we cannot have $q \notin Z$ both provable and unprovable, we have a contradiction. ■

We have seen that either G or $\neg G$ is true but that neither of these two statements is provable. Thus, in our *Principia Mathematica*-like system, there is at least one statement that *is true, but is not provable!*

Applications of Gödel's Undecidability Theorem

Gödel's Undecidability Theorem has profound implications. It tells us that even such a fundamental branch of mathematics as arithmetic has its limitations. It is not always possible to find a proof for every true statement within that system. Thus, we may never be able to establish whether certain conjectures within the system are true. You might wonder whether we could enlarge our system to a more powerful one in the hopes of finding proofs for certain statements. But, we could then apply Gödel's Undecidability Theorem to this larger system to show that it also has limitations.

Rudy Rucker writes in [4] that "Gödel once told the philosopher of science Hao Wang that the original inspiration for his ... theorem was his realization that 'truth' has no finite description." That is, no mathematical system with a finite number of axioms such as we have described here can contain a full depiction of reality. No mathematical system like *Principia Mathematica* can be both consistent and complete.

In addition to the Undecidability Theorem, Gödel also demonstrated in his paper that a system like *Principia Mathematica* is not even powerful enough to establish its own consistency! That is, the consistency of the system cannot be proven using only the principles available within the system itself.

In recent years, the implications of Gödel's Undecidability Theorem in computer science have been studied widely, especially in the area of artificial intelligence. If we imagine giving to a computer an initial set of formulas (axioms) and a set of logical principles (operations) for calculating new formulas (theorems), then Gödel's Undecidability Theorem implies that, no matter how powerful the computer, there are always some true statements that the computer will never be able to derive. Although computers have been useful to mathematicians in proving computationally difficult results, Gödel's Undecidability Theorem destroys the myth many people have that any difficult mathematical problem will eventually be resolved by a powerful enough computer.

Computer scientists have drawn an even closer parallel between mathematical systems and computers. The initial state (register contents, data) of a computer is analogous to the initial set of formulas (axioms) in a mathematical system. Similarly, the logical operations (via hardware and software) that a computer is allowed to perform are analogous to the principles of logic permitted in the mathematical system. Extending this analogy, the mathematician (and pioneer computer scientist) Alan Turing* extended Gödel's results to prove that there is no general algorithm that can always correctly predict whether a randomly selected computer program will run or not.

In 1961, the British philosopher J. Anthony Lucas tried to use Gödel's Undecidability Theorem to show that a machine will never be able to "think"

* Alan Turing (1912–1954) was a British mathematician and logician, who received his Ph.D. from Princeton in 1938. In his landmark 1937 paper "On Computable Numbers", he proposed the notion of a general computer (later known as a *Turing machine*), that could execute finite mathematical algorithms, and proved there exist mathematical problems that can not be solved by such a machine. Turing also worked on the enormous problems of breaking secret codes during World War II with the Foreign Office of the British Department of Communications. He was the subject of a recent Broadway play (starring Derek Jacobi) by Hugh Whitemore entitled "Breaking the Code". After the war, Turing worked at the National Physical Lab at Teddington on the design of a large (automatic) computer, and became Deputy Director of the Computing Lab at the University of Manchester in 1949. Turing died from a dose of potassium cyanide poisoning, self-administered, but possibly accidental.

in quite the same manner as a human, and hence that the achievement of an artificial intelligence is impossible. However, Lucas' argument is refuted in Douglas Hofstadter's book [2]. In this book, Hofstadter also considers a number of other related artificial-intelligence questions, such as whether the human mind is essentially a computer system (i.e., a mathematical system). If so, are there some truths that the unaided human mind will never fathom?

Suggested Readings

1. M. Gardner, "Douglas R. Hofstadter's 'Gödel, Escher, Bach' ", *Scientific American*, 1979, pp. 16–24. (A quick tour through Hofstadter's book [2].)

2. D. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, 20th Anniversary Edition, Basic Books, New York, 1999. (Pulitzer Prize winning *tour de force* that relates the work of Gödel to the work of the artist M.C. Escher and the composer Johann Sebastian Bach.)

3. E. Nagel and J. Newman, *Gödel's Proof*, New York University Press, New York, 1958. (The classic reference on Gödel's Undecidability Theorem.)

4. R. Rucker, *Mind Tools: The Five Levels of Mathematical Reality*, Houghton-Mifflin, Boston, 1987. (See especially the section on Gödel's Undecidability Theorem, pp. 218–226.)

5. R. Smullyan, *Forever Undecided: A Puzzle Guide to Gödel*, Knopf, New York, 1987. (Understanding the concepts behind Gödel's Undecidability Theorem by solving successively more difficult puzzles.)

6. J. van Heijenoort, *From Frege to Gödel, A Source Book in Mathematical Logic, 1879–1931*, Harvard University Press, Cambridge, MA, 2002. (Contains a complete English translation of Gödel's 1931 paper, as well as Gödel's abstract of the paper, and a related note that he wrote about the paper.)

Exercises

1. Verify that Goldbach's Conjecture is true for all the even integers from 14 to 50.

2. Another unproven conjecture in number theory is the following: Let f: $\mathcal{N} \longrightarrow \mathcal{N}$ be defined by

$$f(n) = \begin{cases} n/2 & n \text{ even} \\ 3n+1 & n \text{ odd;} \end{cases}$$

then, for every n, there is an integer i such that $f^i(n) = 1$. Verify that this conjecture is true for $n = 22$ and $n = 23$.

3. Exhibit a model for the axioms in Example 3 containing fewer than four points.

In Exercises 4–6, find a finite model for the system having the given set of axioms. Let "point" and "line" be undefined terms. (Remember, a "line" does not necessarily have to be straight or contain an infinite number of points.)

4. (i) There are exactly four points in the system.
(ii) Not all points are on the same line.
(iii) Through any three points there is exactly one line.

5. (i) There are at least two lines in the system.
(ii) Each line contains at least two points.
(iii) Every pair of points is on a line.
(iv) No point is on more than two lines.

6. (i) There is at least one line and a point not on that line.
(ii) Each line contains exactly three points.
(iii) Each pair of points lies on exactly one line.
Hint: Use seven points and seven lines.

7. In Exercise 5, can you find a model containing exactly 4 points?

In Exercises 8 and 9, show that the given set of axioms is inconsistent.

8. (i) There are exactly four points in the system.
(ii) Each pair of points lies on exactly one line.
(iii) Every line contains exactly three points.

9. (i) There is at least one line in the system.
(ii) Every two points lie on exactly one line.
(iii) Every point lies on exactly two distinct lines.
(iv) Every line contains exactly three points.

10. Explain why the following statement is a paradox: "This sentence is false."

11. Suppose there are two types of people living on a certain island: truth-tellers (who never lie) and liars (who never tell the truth).
a) An inhabitant of the island says, "I am lying." Explain the paradox.
$\star$ b) An inhabitant of the island says, "You will never know that I am telling the truth." Is this a paradox?

12. Using the Gödel numbering scheme established earlier in this chapter, give the Gödel number for each of the following statements:

a) $x = ff0$.

b) $\neg((x = 0) \vee (y = 0))$.

c) $\forall x \exists y(\neg(fx = y))$.

d) $\forall x((x = f0) \vee \neg(x = f0))$.

13. Give the Gödel number for each of the following proofs (i.e., sequence of statements):

a) $$S = \left\{ \begin{array}{c} x = f0 \\ y = fx \\ y = ff0 \end{array} \right\}$$

b) $$T = \left\{ \begin{array}{c} \exists y(x = fy) \\ \forall y(\neg(0 = fy)) \\ \neg(x = 0) \end{array} \right\}$$

14. In each case, state the proposition that has the given Gödel number.

a) $2^5 3^{11} 5^{21} 7^{19} 11^{23} 13^{13}$.

b) $2^9 3^{21} 5^{17} 7^{23} 11^{11} 13^{21} 17^{19} 19^3 23^{23} 29^{13}$.

c) $2^{11} 3^{21} 5^{19} 7^1 11^{13} 13^7 17^5 19^{11} 23^{23} 29^{19} 31^1 37^{13}$.

15. In each case, state the sequence of propositions that has the given Gödel number.

a) $2^{k_1} 3^{k_2} 5^{k_3}$, where $k_1 = 2^{21} 3^{19} 5^3 7^3 11^1$, $k_2 = 2^5 3^{11} 5^{23} 7^{19} 11^{21} 13^{13}$, and $k_3 = 2^5 3^{11} 5^{23} 7^{19} 11^3 13^3 17^1 19^{13}$.

b) $2^{k_1} 3^{k_2} 5^{k_3}$, where $k_1 = 2^9 3^{23} 5^{17} 7^{21} 11^5 13^{11} 17^{21} 19^{19} 23^{23} 29^{13}$, $k_2 = 2^{23} 3^{19} 5^3 7^1$, and $k_3 = 2^{17} 3^{21} 5^5 7^{11} 11^{21} 13^{19} 17^3 19^1 23^{13}$.

$\star$**16.** Suppose we enlarged our *Principia Mathematica*-like system to include Gödel's statement G as an axiom. If the new mathematical system is consistent (that is, if G is true), what does Gödel's Undecidability Theorem tell us about the new system?

Computer Projects

1. Write a program that takes as input an even positive integer greater than 2 and writes it as a sum of two primes (Goldbach's Conjecture).

2. Write a program that implements the function defined in Exercise 2.

3. Write a program that takes a list of primes and gives as output the string of symbols that constitutes the corresponding Gödel statement or recognizes that the list of primes is unacceptable.