Turing's work and Hilbert's Tenth Problem

Kirsten Eisenträger

The Pennsylvania State University

AMS-ASL Special Session on the Life and Legacy of Alan Turing

January 5, 2012

Hilbert's Tenth Problem (H10)

Original problem was posed by Hilbert in 1900, tenth of his famous 23.

H10/$\mathbb{Z}$: Find an algorithm that decides, given a multivariate polynomial equation $f(x_1, \ldots, x_n) = 0$ with coefficients in the ring $\mathbb{Z}$ of integers, whether there is a solution with $x_1, \ldots, x_n \in \mathbb{Z}$.

- Matiyasevich (1970): No such algorithm exists!
- Matiyasevich's proof was based on work by Putnam, Robinson and Davis.
- We say that Hilbert's Tenth Problem is undecidable.
- Can also formulate the problem for a system of equations. This is equivalent since

$$f_1 = f_2 = 0 \iff f_1^2 + f_2^2 = 0.$$

The notion of algorithm

- To understand H10: need a precise definition of what we mean by "algorithm".
- In 1900: no precise notion of the term algorithm existed.
- In 1930s: several models of computation were proposed and were shown to be equivalent. One of these models was the Turing machine.
- The equivalence gave credibility to the Church-Turing thesis.
- Informally, the Church-Turing thesis asserts that the intuitive notion of algorithm equals that of Turing machine algorithm. Because of this, "algorithm" is taken to mean Turing machine algorithm.
- A Turing machine algorithm is equivalent to a finite-length computer program, when the computer is assumed to have unlimited memory.

Recursively enumerable and recursive sets

A set $Q \subseteq \mathbb{Z}$ is *recursively enumerable* (or *listable*) if there is an algorithm that prints the elements of Q when left running forever (in any order and with repetitions permitted).

A set $Q \subseteq \mathbb{Z}$ is *recursive* (or *computable* or *decidable*) if there is an algorithm that decides membership in Q. I.e. there is an algorithm that takes as input an integer a and answers YES if $a \in Q$ and NO if $a \notin Q$.

Immediate consequence: A set S is recursive if and only if S and its complement $\overline{S}$ are both recursively enumerable.

The Halting Problem

- The negative answer to Hilbert's Tenth Problem was proved by relating it to undecidability results in logic and computability theory from the 1930s.
- It used a recursively enumerable set that is not recursive. Such a set can be obtained from the *halting problem*.

The halting problem asks for an algorithm that takes as input a computer program p and an integer n, and outputs YES or NO, according to whether program p run on input n eventually halts (instead, for example, of entering an infinite loop).

Theorem (Turing, 1936). The halting problem is undecidable; that is, no Turing machine can solve it.

The Halting Problem, continued

Sketch of proof.

- Fix an encoding of programs as natural numbers and identify programs with their associated integers.
- Suppose that there were an algorithm for deciding if a program p halts on input n.
- Using this program we could build a new program H with the following property:
  For any n,

  H halts on input $n \iff$ program n does not halt on input n.

- Taking $n = H$, we obtain a contradiction: H halts on input H if and only if H does not halt on input H. $\square$

Turing's theorem has the following important corollary:

Corollary. There exists a recursively enumerable set that is not recursive.

Diophantine sets

A subset $Q \subseteq \mathbb{Z}^k$ is *diophantine* if there exists a polynomial $f(x_1, \ldots, x_k, y_1, \ldots, y_m)$ with integer coefficients such that

$$Q = \{\vec{x} \in \mathbb{Z}^k : \exists y_1, \ldots, y_m \in \mathbb{Z} : f(\vec{x}, y_1, \ldots, y_m) = 0\}.$$

Examples:

(1) $\mathbb{N}$ is diophantine over $\mathbb{Z}$:

$$x \in \mathbb{N} \iff \exists y_1, \ldots, y_4 \in \mathbb{Z} : y_1^2 + \cdots + y_4^2 - x = 0.$$

(2) The set of primes is diophantine over $\mathbb{Z}$.

The fact that the set of primes is diophantine follows from the proof of Hilbert's Tenth Problem.

The DPRM Theorem

Davis-Putnam-Robinson-Matiyasevich proved:

DPRM Theorem: A set $Q \subseteq \mathbb{Z}$ is recursively enumerable if and only if it is diophantine.

- One direction of this theorem is easy: if $Q \subseteq \mathbb{Z}$ is diophantine, then we can just write a program that loops through all elements $(a, y_1, \ldots, y_m) \in \mathbb{Z}^{m+1}$ and prints a if $f(a, y_1, \ldots, y_m) = 0$.
- The other direction is the hard direction.
- The DPRM Theorem was first conjectured by Davis in 1949.
- Davis also carried out the first steps towards its proof.

Undecidability Proof of H10

The DPRM Theorem immediately implies that there is no algorithm for Hilbert's Tenth Problem:

DPRM Theorem $\Rightarrow$ H10 is undecidable:
Let $Q \subseteq \mathbb{Z}$ be such that Q is recursively enumerable but not recursive.
DPRM Theorem $\Rightarrow$ Q is diophantine with defining polynomial $f(a, y_1, \ldots, y_m)$.
If there were an algorithm for Hilbert's Tenth Problem, apply this algorithm to f to decide membership in Q. But Q is not recursive, so such an algorithm cannot exist. $\square$

Outline of proof of DPRM Theorem

Davis carried out the first steps towards proving the DPRM Theorem. He showed:

Theorem: For every r.e. set S there is a polynomial $p(a, k, y, x_1, \ldots, x_n)$ such that a number a_0 belongs to S if and only if

$$(\exists y)\ (\forall k \leq y)\ (\exists x_1, \ldots, x_n)\ [p(a_0, k, y, x_1, \ldots, x_n) = 0].$$

- At first glance this result seems close to the final goal.
- Need to get rid of the universal quantifier $(\forall k \leq y)$ to get a diophantine definition for the r.e. set S.
- This turned out to be difficult.

Diophantine definition of exponentiation

Julia Robinson aimed to show that exponentiation was diophantine, i.e. that the set of all triples

$$\{(a, b, c) \in \mathbb{N}^3 : \ c = a^b\}$$

was a diophantine set.

Robinson was able to prove that the diophantineness of exponentiation would follow from the existence of a 2-variable diophantine relation of exponential growth.

She introduced the following hypothesis:

Julia Robinson Hypothesis (JR): There exists a diophantine set J of pairs (a, b) such that

- If (a, b) belongs to J then $b < a^a$;
- for every natural number k, there is a pair (a, b) in J for which $b > a^k$.

For example, the set of pairs (a, b) such that $b = 2^a$ satisfies both conditions.

Diophantine definition of exponentiation, continued

Julia Robinson proved:

Theorem (Robinson, 1952): Assume that JR holds. Then exponentiation is diophantine.

Robinson also studied *exponential diophantine equations*, which are equations in which some of the exponents are variables as well.

A set that is definable by an exponential diophantine equation is called an *exponential diophantine set*.

Not hard to see: If exponentiation is diophantine then all exponential diophantine sets are diophantine.

Recursively enumerable sets and exponential diophantine sets

In 1961 Davis, Putnam and Robinson proved the analogue of the DPRM theorem for exponential diophantine equations:

Theorem (Davis-Putnam-Robinson, 1961). Every recursively enumerable set is exponential diophantine.

- So to finish the proof of the DPRM Theorem: had to show that Julia Robinson's hypothesis was true.
- So had to find a 2-variable diophantine relation of exponential growth.
- This was done by Matiyasevich, using properties of Fibonacci numbers.

Matiyasevich's contribution

Matiyasevich proved the following theorem:

Theorem (Matiyasevich, 1970). Let F_n be the n-th Fibonacci number. The relation $m = F_{2n}$ is diophantine.

- The Fibonacci numbers grow exponentially and satisfy the conditions of Julia Robinson's hypothesis JR.
- So Matiyasevich's theorem completed the proof of the DPRM Theorem and proved the undecidability of Hilbert's Tenth Problem.

Generalizations of H10

H10/R: Find an algorithm that decides, given a multivariate polynomial equation $f(x_1, \ldots, x_n) = 0$ with coefficients in R, whether it has a solution with $x_1, \ldots, x_n \in R$.

- Biggest unsolved question: $R = \mathbb{Q}$.
- Can rephrase the problem for $\mathbb{Q}$ in more geometric terms: H10/$\mathbb{Q}$ is equivalent to the existence of an algorithm that decides whether an algebraic variety defined over $\mathbb{Q}$ has a $\mathbb{Q}$-rational point.
- Even more difficult: Hilbert's Tenth Problem for arbitrary number fields K.

Generalize definition from before: A subset $Q \subseteq R^k$ is *diophantine over R* if there exists a polynomial $f(x_1, \ldots, x_k, y_1, \ldots, y_m)$ with coefficients in R such that

$$Q = \{\vec{x} \in R^k : \exists y_1, \ldots, y_m \in R : f(\vec{x}, y_1, \ldots, y_m) = 0\}.$$

Decidability and Undecidability Results

Decidable: H10 for

- finite fields
- p-adic fields
- real closed fields, algebraically closed fields (Tarski)

Undecidable: H10 for

- some rings of integers of number fields (Denef, Lipshitz, Pheidas, Shlapentokh, Poonen)
- function fields of curves over finite fields (Pheidas, Shlapentokh, Videla, E.)
- $K(t)$, K formally real (Denef)
- $\mathbb{C}(t_1, t_2, \ldots, t_n)$ ($n \geq 2$) and finite extensions (Kim and Roush, E.)
- p-adic function fields (Kim and Roush, E., Moret-Bailly)
- large subrings of $\mathbb{Q}$ (Poonen, E.-Everest)
- large subrings of number fields (Poonen-Shlapentokh, E.-Everest-Shlapentokh)

Open Problems

- H10 for $\mathbb{Q}$ and for number fields in general.
- H10 for rings of integers in arbitrary number fields. By a recent result of Mazur and Rubin, H10 is undecidable for arbitrary rings of integers if the Shafarevich-Tate conjecture holds.
- H10 for $\mathbb{C}(t)$ and $\overline{\mathbf{F}}_p(t)$.

Approach for Undecidability Proofs in Characteristic 0

When R is a domain of characteristic zero we can use the following theorem:

Theorem: If $\mathbb{Z}$ is diophantine over R, then H10/R is undecidable.

Proof is by reduction:

Assume by contradiction that H10/R is decidable. From the algorithm for H10/R we can get an algorithm for H10/$\mathbb{Z}$ as follows:

- Given a polynomial $f(x_1, \ldots, x_n)$ over $\mathbb{Z}$, we can use the algorithm for R to test whether f has a solution $x_1, \ldots, x_n$ in R.
- Can use the diophantine definition of $\mathbb{Z}$ to add extra equations saying that the variables x_i take integer values.
- Since H10/$\mathbb{Z}$ is undecidable, an algorithm for it does not exist, and hence H10/R must be undecidable as well.