After Gödel

Hilary Putnam, *Department of Philosophy, Harvard University, Cambridge, MA 02138, USA*

Abstract

This paper describes the enormous impact of Gödel's work on mathematical logic and recursion theory. After a brief description of the major theorems that Gödel proved, it focuses on subsequent work extending what he did, sometimes by quite different methods. The paper closes with a new result, applying Gödel's methods to show that if scientific epistemology (what Chomsky calls our "scientific competence") could be completely represented by a particular Turing machine, then it would be impossible for us to *know* that fact.

Keywords: Kurt Gödel, Alfred Tarski, Noam Chomsky, Completeness, Incompleteness, Competence

1 Introduction

Some years ago, I heard a famous logician say that "Alfred Tarski was the greatest logician after Aristotle". Tarski was indeed a great logician, but for reasons I will explain in a moment, it seems clear to me that the title "greatest logician after Aristotle" belongs to Kurt Gödel and not Alfred Tarski (or, as one might be allowed to refer to him in Israel, "Alfred Teitelbaum" – Tarski changed his name to a more "Polish" one because of the anti-Semitism in the Polish universities). On another occasion, the famous mathematician David Mumford (who later "went into" computer science) said to me "As far as I am concerned, Gödel wasn't a mathematician. He was just a philosopher." I don't know if Mumford still thinks this, but if he does, he is wrong too.

Gödel is, of course, known best for the famous incompleteness theorems, but they are only a part of his contribution to logic. If they were the whole, or the only part of such fundamental importance, then the claim that Tarski's formalization of the notion of "satisfaction", and his use of that notion to show us how to define truth of a formula in a formalized language over a model, was at least as great a contribution might be tenable.[1] However, the very field for which Tarski is most famous, model theory, was launched by two theorems one of which bears Gödel's name: I refer to the Skolem-Löwenheim Theorem and the Gödel Completeness (or Completeness and Compactness) Theorem. Moreover, Church's Theorem was clearly known to Gödel before Church, as a careful reading of the footnotes to the famous paper on Undecidable Sentences makes clear. And without question, what I may call the "Gödel-Cohen Theorems" – that the Axiom of Choice and the Continuum Hypothesis are both independent of the other axioms of set theory (unless those axioms are inconsistent) are by far the most stunning results ever obtained in set theory. (Gödel proved that the AC and the CH are consistent with the axiom of ZF in the 30s and in 1962 Paul Cohen showed that their negations are likewise consistent. These result are extremely

[1] However, I have argued that the philosophical, as opposed to the mathematical, significance of Tarski's achievement has been both overestimated and misrepresented. See [3].

robust, in the sense that it is unlikely that any further axioms that one could add to set theory will both resolve the continuum problem and be found sufficiently "intuitively evident" to command acceptance by the mathematical community, although Gödel hoped the contrary.)

One could also mention many other contributions by Gödel, including important contributions to recursion theory. For example, in the course of proving the incompleteness theorems, Gödel showed that for any predicate $F(x)$ in a formal system which contains expressions for all the primitive recursive functions, there is a expression of the form $F(N)$, where f is a primitive recursive function and N is an integer,[2] such that the Gödel number of the formula $F(f(N))$ is equal to precisely $F(N)$ (i.e., the formula is true if and only if its own Gödel number has the property F). [I shall refer to this as "Gödel's Diagonal Lemma", although he did not give it a name.] This means that, for example, there is a formula which is true if and only if its own Gödel number is prime (just take $F(n)$ to be the formula Prime(n)), a formula which is true if and only if its own Gödel number is even, etc. This is not just *similar* to the fixed point theorem of recursion theory; it introduced precisely the technique of proof we need for that fixed point theorem (Kleene's recursion theorem).

What I want to do now is describe some of the ways logicians after Gödel have built upon and extended his results.

2 Diophantine equations

Gödel's own undecidable sentence can easily be put into the form "There does not exist a natural number n such that $f(n) = 0$", where $f(n)$ is primitive recursive. But primitive recursive functions are not a topic the average mathematician is particularly interested in. (Probably David Mumford was not when he made the remark I quoted earlier.)

However. in the late 1950's Martin Davis, Julia Robinson and myself proved that the decision problem for exponential Diophantine equations[3] is recursively unsolvable. Julia Robinson had already shown that the decision problem for ordinary Diophantine equations is equivalent to the decision problem for exponential Diophantine equations if and only if there exists a single ordinary Diophantine equation whose solutions (considered as functions of any one of the unknowns) have roughly exponential rate of growth, and a few years late Yuri Matyasevich proved that such an equation exists. This Davis-Matyasevich-Putnam-Robinson theorem showed that the decision problem for ordinary Diophantine equations is recursively unsolvable, thus providing a negative solution to Hilberts's Tenth Problem. In fact, we showed that for every recursively enumerable set S, there is a polynomial P with integral coefficients such that $P(n, x_1, x_2, \ldots x_k) = 0$ has a solution in natural numbers exactly when n belongs to S. Applied to Gödel's original paper on undecidable sentences, this yields the fact that the undecidable sentence can have the mathematically very familiar form, "the Diophantine equation $P = 0$ has no solution". To quote I don't know whom, "Who woulda thunk it?".

---

[2] I am deliberately ignoring the "use-mention distinction" to simplify exposition.

[3] Equations of the form $P = 0$, to be solved in natural numbers, where P is a polynomial with integral coefficients are called Diophantine equations "Exponential Diophantine equations" are equations of the form $P = 0$, to be solved in natural numbers, where P is an expression which is like a polynomial with integral coefficients except for having some variable exponents. E.g. the Fermat equation $x^n + y^n - z^n = 0$ is exponential Diophantine.

3 Model theory

In 1930, Gödel also showed that the standard axioms and rules of quantification theory (which was known at that time by the name the American philosopher Charles Pierce, the founder of Pragmatism, gave it – “first order logic”) are *complete*, in the sense that every valid formula, every formula which is true in all possible models, is a theorem. His proof also establishes that if every finite subset of an infinite set of formulas of quantification theory has a model, then the whole infinite set has a model. This is the Compactness Theorem, and is still of fundamental importance in model theory.

To illustrate the importance of the Compactness Theorem, let P be any set of axioms for Peano Arithmetic, or, for that matter, for any consistent system which extends Peano Arithmetic. Let $\underline{a}$ be a new individual constant, i.e., one not used in P. Consider the theory T with the following recursively enumerable set of axioms: the axioms of P *plus* “$\underline{a}$ is a natural number”, “$\underline{a} \neq 0$”, “$\underline{a} \neq 1$”, “$\underline{a} \neq 2$”, “$\underline{a} \neq 3$”, ... and so on *ad infinitum.* Let S be any finite subset of these axioms, and let N be the largest integer such that $\underline{a} \neq \mathrm{N}$ is a member of S. Then S obviously has a model – just take any model for P, and interpret $\underline{a}$ as denoting N+1. By the Compactness Theorem T has a model. In that model, the object denoted by “$\underline{a}$” is an infinite integer – and so is $\underline{a} + 1$, $\underline{a} + 2$, $\underline{a} + 3$. . . .. Moreover, a has a predecessor $\underline{a}-1$ in the model (otherwise, by a theorem of Peano arithmetic, and hence a theorem of P, it would be 0, violating the axiom “$\underline{a} \neq 0$”, and that predecessor has a predecessor, etc., and all of these “natural numbers” $\underline{a}-1$, $\underline{a}-2$, $\underline{a}-3$ . . . .. are likewise infinite integers. Thus we have the existence of non-standard models for mathematics! (By a “non-standard” model I mean a model in which, in addition to the integers 0, 1, 2, ... there are also, viewed from the outside, “infinite integers”. I say “viewed from the outside” because within the formal system itself there is no way to single out these non-standard elements, or even to say they exist: It is only in the metalanguage in which we prove the existence of the non-standard model that we can say that there are “foreign elements” in the model, elements other than the “real” natural numbers.) Abraham Robinson showed, in fact, that by using such models one can carry out Leibniz’s dream of a true calculus of infinitesmals, and the resulting branch of mathematics, which has been called Non-Standard Analysis, already has significant applications in many areas – to the theory of Lie groups, to the study of Brownian motion, to Lebesgue integration, etc. What gives the subject its power is that, because the “infinite numbers” all belong to a model for the theory of the standard (finite) integers and real numbers, a model within which they are not distinguished by any predicate of the language from the standard numbers, we are guaranteed from the start that they will obey all the laws that standard numbers obey.

4 Kripke’s Proof

In the last twenty years various workers, the most famous being Paris and Harrington, have begun to use the existence of non-standard models to give independence proofs in number-theory itself [1]. The very existence of independent (or “undecidable”) propositions of elementary number theory was proved by Gödel in 1934 by syntactic, not model theoretic, means. The proposition proved independent by Paris and Harrington is a statement of graph theory (a strengthened version of Ramsey’s theorem).

What I want to tell you about now, however, is a different model-theoretic proof of the existence of undecidable sentences. This proof is due to Saul Kripke, and was written up by myself in a paper published in 2000 [4].

Because Kripke's theorem does not aim at establishing the independence of a statement which is nearly as complicated as the proposition Paris and Harrington wished to prove independent, the proof is much simpler than theirs. Also, because the independence proof is semantic rather than syntactic, we do not need to arithmetize the relation "x is a proof of y", as Gödel had to do for his proof. We do not need the famous predicate *Bew*(x) (x is the Gödel number of a theorem), or the famous self-referring sentence which is true if and only if its own Gödel number is not the Gödel number of a theorem. In short, Kripke gave us a different proof of the Gödel theorem, not just a different *version* of Gödel's proof.

While time does not permit me to give Kripke's proof in full – you can find it in the paper I mentioned, in the *Notre Dame Journal of Formal Logic*, I do want to give some idea of the remarkable ingenuity it displays, ingenuity that I am sure Gödel himself would have very much appreciated.

First, Kripke introduces what we might call "finite partial models" (the term is my own) of statements in number theory. To be specific: Consider a finite monotone-increasing series $\underline{s}$ of natural numbers, say

182, 267, 349, 518,......................., 3987654345

Such a series will be what I am calling a "finite partial model" for a formula A of number theory, say (∀x)(∃y)Rxy (with primitive recursive R), if the series fulfills the formula in the sense I will now explain. In our explanation we will identify a sequence with its Kleene "Gödel number" $\underline{s}$, where convenient. Following Kleene's convention, the successive members of the sequence will be denoted $(\underline{s})_0$, $(\underline{s})_1$, ...

We shall say $\underline{s}$ fulfills A if the second player (the "defending player") has a winning strategy in the game I shall describe.

The Game G

> The game is played as follows. The first player (the "attacking player") picks a number less than the length of the given sequence $\underline{s}$, say 3. The sequence $\underline{s}$ is examined to determine the third place in the sequence. The same player (the attacker) now picks a number less than the number in that place (less than 349, in the case of our example.) Let us suppose she picks 17. We assume the number picked by the first player was less than the length of the sequence (otherwise the first player has lost). If so, the second player (the "defending player") gets to look at the next number in the sequence (at $(\underline{s})_3$ or 518, in the case of the example). She must pick a number less than this number (less than $(\underline{s})_{n+1}$, if the first player picked the place $(\underline{s})_n$). Let us suppose she picks 56. We now evaluate the statement R(17, 56) (the statement R(n, m), where n is the number picked by the first player and m is the number picked by the second player). Since R is primitive recursive, this can be done effectively. If the statement is true the defending player has won; if false the attacking player has won.

The statement that a sequence $\underline{s}$ fulfils this statement A (that there is a winning strategy for the defending player) can itself be written out in number theory, as follows:

(I) $(\forall i \leq \text{length}(\underline{s}) - 1)(\forall n \leq (\underline{s})_{i-1})(\exists m \leq (\underline{s})_i)\mathrm{R}nm$

Similarly, if we are given a statement A with four, or six, or however many alternating quantifiers in the prefix, we can define $\underline{s}$ fulfills A to mean that there is a winning strategy for the defending player in a game which is played very much like the game G: a game in which the attacking player gets to choose a new place in the sequence, each time it is her turn to play. The attacking player must also chose a number less than $(\underline{s})_{i-1}$, where $(\underline{s})_{i-1}$ is the number in the position she chose in the sequence.[4] Each time she plays, the attacking player has to chose a place which is to the right of the place in the sequence she chose before (unless it is her first turn to play) and not the last place in the sequence (unless she has no legal alternative, in which case she loses), and the defending player must then pick a number less than $(\underline{s})_i$ (less than the number in the next place in the sequence). The game ends when as many numbers have been chosen as there are quantifiers in the prefix of the formula. (We assume all formulas are prenex, and that quantifiers alternate universal, existential, universal, existential....) The numbers chosen are then substituted for the variables in the matrix of the formula A in order (first number chosen for x_1, second number chosen for x_2, etc., where x_1 is the variable in the first argument place, x_2 the variable in the second argument place, and so on). The resulting primitive recursive statement is evaluated and, as before, the defending player wins if the statement is true and the first (attacking) player wins if the statement is false. Once again, for any fixed formula A we can easily express the statement that $\underline{s}$ fulfills A arithmetically (primitive recursively in $\underline{s}$). And for any fixed recursively enumerable sequence of formulas $A_1, A_2 \ldots$, the statement that $\underline{s}$ fulfills A_n can be expressed as a primitive recursive relation between $\underline{s}$ and n, say *Fulfills*$(\underline{s}, A_n)$. Note that we can also speak (by an obvious extension) of an ordinary infinite monotone increasing sequence "fulfilling" a formula. (This means that if one picks any number less than a given number in the sequence to be the value of the first universal quantifier, it is always possible to pick a number less than the next place to the right in the sequence to be a value for the succeeding existential quantifier, so that no matter what number less than the number in an arbitrarily selected place still further to the right in the sequence one picks for the *next* universal quantifier, it is possible to pick a number less than the number in the place in the sequence immediately to the right of the last "universal quantifier place" chosen for the last existential quantifier so that the statement A comes out true.) And note that a statement which is fulfilled by an infinite monotone increasing sequence is true. (Since the restriction that one must pick numbers as values for the universal quantifiers which are bounded by the numbers in the sequence is, in effect, no restriction on the "attacking player" at all—-the numbers in the sequence get arbitrarily large, so she can pick any number she wants by going out far enough in the sequence!)

In Kripke's proof, one confines attention to sequences with the following two properties (call them good sequences):

(1) The first number in the sequence is larger than the length of the sequence.

[4] N.B. the number in the ith position is called "$(\underline{s})_{i-1}$" and not "$(\underline{s})_i$" because Kleene – whose notation I am employing—calls the first position "$(\underline{s})_0$" and not $(\underline{s})_1$".

(2) Each number in the sequence after the first is larger than the square of the number before. (This is to ensure that the sum and product of numbers $\leq (\underline{s})_i$ are $\leq (\underline{s})_{i-1}$.)

Finally (this is the last of the preliminaries!) let P1, P2, P3, ... be the axioms of Peano arithmetic.

We will say that a statement is n-fulfillable if there is a good sequence of length n which fulfills the statement. The following is the statement which Kripke showed to be independent of Peano arithmetic:

(II) *For every n and every m, the conjunction of the first m axioms of Peano arithmetic is n-fulfillable.*

If we think of good sequences that fulfill a formula A as "finite partial models" for A, then what this says is that the first m axioms of Peano arithmetic have finite partial models that are as big as you like. It is easy to see that (II) is true in the standard model of Peano arithmetic.

Kripke's construction of a non-standard model in which (II) is *false* is actually very short and elegant. Thus we have the remarkable result that post-Gödelian model theory can be used to replace recursion theory in the proof of the Gödel Incompleteness Theorem!

5 Prime numbers

Here is yet another result which "milks" Gödel's theorem. This is actually an easy corollary of the results about Diophantine equations that I mentioned earlier, but number theorists expressed amazement when I published it in 1960:

If $P(x_1, x_2, \ldots x_k)$ is a polynomial, let us refer to the positive integers n such that $P(x_1, x_2, \ldots x_k) = n$ for some integers $x_1, x_2, \ldots x_k$ as the *positive range* of P. Then the theorem I proved is that there *exists a polynomial whose positive range is exactly the prime numbers* [5].

The proof is sufficiently simple for me to give it in full.

The primes are a recursively enumerable set. So by the Davis-Matyasevich, Putnam, Robinson theorem, there is a polynomial P with integral coefficients such that

(1) The equation $P(n, x_1, x_2, \ldots x_k) = 0$ has a solution in integers when and only when n is a prime number.

It is a theorem of number theory that every positive integer is the sum of four squares. So, I claim that the following equation has an integral solution with positive n when and only when n is a prime number:

(2) $n = \left(y_1^2 + y_2^2 + y_3^2 + y_4^2\right) \cdot \left[1 - P(y_1^2 + y_2^2 + y_3^2 + y_4^2, x_1, x_2, \ldots x_k)\right]^2$
where P is the polynomial in (1).

Proof. First, suppose n is a prime number. Then there are $x_1, x_2, \ldots x_k$ such that $P(n, x_1, x_2, \ldots x_k) = 0$, by (1). Choose such $x_1, x_2, \ldots x_k$ and let y_1, y_2, y_3, y_4 be such that $n = y_1^2 + y_2^2 + y_3^2 + y_4^2$ (by the Four Squares theorem there are such y_1, y_2, y_3, y_4). Then $P(y_1^2 + y_2^2 + y_3^2 + y_4^2, x_1, x_2, \ldots x_k) = 0$ and the second factor in (2) is equal to 1. So (2) reduces to $n = (y_1^2 + y_2^2 + y_3^2 + y_4^2)$, which is correct, by the choice of y_1, y_2, y_3, y_4.

Second, suppose that (2) is true for natural numbers n, y_1, y_2, y_3, y_4, and $x_1, x_2, \ldots x_k$. If $P(n, x_1, x_2, \ldots x_k) \neq 0$, then $[P(y_1^2 + y_2^2 + y_3^2 + y_4^2, x_1, x_2, \ldots x_k)]^2 \geq 1$, and $(1 - [P(y_1^2 + y_2^2 + y_3^2 + y_4^2, x_1, x_2, \ldots x_k)]^2)$ is zero or negative. In either case $n = y_1^2 + y_2^2 + y_3^2 + y_4^2 \cdot (1 - P(y_1^2 + y_2^2 + y_3^2 + y_4^2, x_1, x_2, \ldots x_k)]^2)$ is non-positive. And if $P(y_1^2 + y_2^2 + y_3^2 + y_4^{2,}, x_1, x_2, \ldots x_k) = 0$, then by (1), $y_1^2 + y_2^2 + y_3^2 + y_4^2$ is a prime number and since $(1 - [P(y_1^2 + y_2^2 + y_3^2 + y_4^{2,}, x_1, x_2, \ldots x_k)]^2 = 1$, (2) reduces to $n = y_1^2 + y_2^2 + y_3^2 + y_4^2$, so n is a prime. Thus the prime numbers are all and only the *positive* integers taken on as values by the polynomial $(y_1^2 + y_2^2 + y_3^2 + y_4^2) \cdot [1 - P(y_1^2 + y_2^2 + y_3^2 + y_4^2, x_1, x_2, \ldots x_k]^2$. ∎

Since this conference has been titled "Models of Computation", let me mention that this last result, generalized to say that *every* recursively enumerable set is the positive range of some polynomial with integral coefficients, yields a very simple "model of computation": in this model a "program" is simply a polynomial with integral coefficients, and one recursively enumerates a set simply by plugging in k-tuples of integers for the variables in the polynomial, calculating the numerical value of the resulting expression, and writing down the result whenever that numerical value is positive. I think the mathematical interest of this model is clear; alas, it is also clear that as a method of computation it is also extremely impractical!

6 An application of Gödel's method: can our "scientific competence" be simulated by a Turing Machine? And if yes, can we know that fact?

I shall close by describing some recent research of mine. What provoked this research is actually a conversation I had with Noam Chomsky more than twenty years ago. I asked Noam whether he thought our *total* competence – not just the competence of the "language organ" that he postulates, but also the competence of the "scientific faculty" that he also postulates – can be represented by a Turing Machine (where the notion of "competence" is supposed to distinguish between our true ability and our "performance errors"). He said "yes". I at once though that a Gödelian argument could be constructed to show that if that were right, then we could never *know* – not just never prove mathematically, but never know even with the help of empirical investigation – *which* Turing Machine it is that simulates our "scientific competence". One such Gödelian argument was given in a paper titled "Reflexive Reflections" that I published exactly twenty years ago [6], but I have always been dissatisfied with one of the assumptions that I used in that proof (I called it a "criterion of adequacy" for formalizations of the notion of justification), namely the assumption that no empirical evidence can justify believing p if p is *mathematically false*.[5] Finally I am ready to show you a proof whose assumptions seem to me unproblematic.

In order to think about Chomsky's conjecture in a rigorous Gödelian way, let "COMPETENCE" abbreviate the empirical hypothesis that a particular Turing Machine T_k perfectly simulates the competence of our "scientific faculty", in Chomsky's sense. To make this concrete, we may take this to assert that, for any scientific proposition **p**, and for any evidence **u** (expressible in the language that T_k uses for expressing scientific propositions and expressing evidence), T_k sooner or later prints out $\mathit{Justified}_u(\mathbf{p})$ if and only if it is justified to accept **p** when **u** is our total relevant

[5] Although some would regard this assumption as self-evident, those of us who allow 'quasi-empirical' methods in mathematics should not accept it. On the latter, see [2].

evidence. In addition, COMPETENCE asserts that whatever hypotheses and evidence we can describe in an arbitrary natural language can also can be represented in T_k's language – i.e., that that language has the full of expressive power of the "Mentalese" that Chomskians talk about. What I shall show is that if this is true, then COMPETENCE cannot be both true and justified.

My proof will, of course, proceed via *reductio ad absurdum.* So let us assume from now on that 'e' is evidence that justifies accepting COMPETENCE" and let "$T_k \Rightarrow Justified_e$ (**p**)" symbolize, "Machine T_k sooner or later prints out $Justified_e$ (**p**)" (the "⇒" can be read simply as 'says'). (Note that "e" and "k" are constants throughout the following argument.)

Applying Gödel's Diagonal Lemma to the predicate $\neg(Justified_e(\mathbf{p}))$", we can construct a sentence in the language of T_k that "says" that acceptance of itself is not justified on evidence e. That sentence has the following GÖDEL PROPERTY:

$Justified_e$(GÖDEL)≡ $Justified_e(\neg(T_k \Rightarrow Justified_e(\text{GÖDEL}))$.

Where does this property come from? Well, by the Diagonal Lemma, "GÖDEL)≡ $\neg(T_k \Rightarrow Justified_e(\text{GÖDEL}))$" is provable (that is the sense in which GÖDEL "says" $\neg(T_k \Rightarrow Justified_e(\text{GÖDEL}))$ – here I am just mimicking Gödel's proof. But since we know this equivalence, the left side is justified just in case the right side is justified.

Here are the Axioms we shall assume concerning these notions:

Axiom A1) $Justified_e(\mathbf{p})$ iff $Justified_e(Justified_e(\mathbf{p}))$[6]

Axiom A2) It is never both the case that $Justified_e(\mathbf{p})$ and $Justified_e(\neg\mathbf{p})$.

Finally, since COMPETENCE is the statement that T_k perfectly represents out competence, we have: If $Justified_e(\mathbf{p})$ and COMPETENCE is true, $T_k \Rightarrow Justified_e(\mathbf{p})$). Since we are assuming we know that this is the case, we have:

Axiom A3) If $Justified_e(\mathbf{p})$ and $Justified_e$(COMPETENCE), then $Justified_e(T_k \Rightarrow Justified_e(\mathbf{p}))$

7 An anti-Chomskian incompleteness theorem: "COMPETENCE" can't be both true and justified

To guide the reader, here is an outline of the proof we shall give: [In both parts of the proof we will assume that COMPETENCE is both true and justified by the evidence e]

Part I: We will prove (by assuming the opposite and deriving a contradiction) that it is not the case that $T_k \Rightarrow Justified_e(\text{GÖDEL}$*) [i.e., the machine* T_k*, which we assumed to be the one that simulates our epistemic competence, does not tell us that we are justified in accepting the Gödelian sentence, on the evidence which, we assumed, justifies our acceptance of COMPETENCE].*

Part II. The Conclusion of Part I can be expressed thus in our notation:

a) $\neg(T_k \Rightarrow \text{Justified}_e(\text{GÖDEL}))$

[6] It suffices here to assume this for case in which p itself contains (via use or mention) no more than two occurrences of "Justified"). Unlike Chomsky, I shall not assume that we have the competence to understand arbitrarily long sentences.

Since this proof can be known to us if we have empirically justified COMPETENCE, and a proof is a justification, it follows immediately that:

b) $\mathit{Justified}_e(\neg(\mathrm{T}_k \Rightarrow \mathrm{Justified}_e(\text{GÖDEL})))$.

Then the rest of Part II will derive a contradiction from b), thus showing that the assumption that COMPETENCE is both true and justified must be false.

Part I of the proof:

1. $\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL) (ASSUMPTION OF PART I – to be refuted)
2. $\mathit{Justified}_e$(GÖDEL). Reason: For the "reductio" proof of our Incompleteness Theorem we are assuming both COMPETENCE and that this belief is justified by evidence **e**. But COMPETENCE, when written out in our notation, is just: "On any empirical evidence **e**, and for any scientific proposition **p**, $\mathit{Justified}_e(\mathbf{p})$ iff $\mathrm{T}_k \Rightarrow \mathit{Justified}_e(\mathbf{p})$", and by 1), $\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL), so $\mathit{Justified}_e$(GÖDEL)).
3. It is justified$_e$ to believe $\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL). Reason: we just showed that $\mathit{Justified}_e$(GÖDEL). Then by the assumption (that we are making throughout) that $\mathit{Justified}_e$(COMPETENCE) and Axiom A3), we are justified$_e$ in believing $\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL).
4. We are justified$_e$ in believing "$\neg(\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL))". Reason: By (2), we are justified$_e$ in believing $\mathit{Justified}_e$(GÖDEL). By the GÖDEL PROPERTY, $\mathit{Justified}_e$(GÖDEL) iff $\mathit{Justified}_e(\neg(\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL))), and since this equivalence is known to us, and we are justified$_e$ in believing the left side, we are justified$_e$ in believing the right side – that is, we are justified$_e$ in believing $\mathit{Justified}_e(\neg(\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL))) – that is, $\mathit{Justified}_e(\mathit{Justified}_e(\mathit{Justified}_e(\neg(\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL))))). So, by Axiom A1), we are justified$_e$ in believing $\neg(\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL)).

But 3) and 4) violate our consistency axiom, A2)! Thus (still assuming that COMPETENCE is both true and justified on evidence e) we conclude that the assumption of our sub-proof is false: it is not the case that $\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL). [But this is reasoning that we can easily go through if we have discovered COMPETENCE to be true! So, without any additional empirical evidence, we have justified$_e$:

"$\neg(\mathrm{T}_k \Rightarrow \mathrm{Justified}_e$(GÖDEL))"

So we have found that $\mathit{Justified}_e(\neg(\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL))). To complete the proof of our Incompleteness Theorem, we therefore now need a proof that this too leads to a contradiction. Here it is:

Part II:

1) $\mathit{Justified}_e(\neg(\mathrm{T}_k \Rightarrow \mathit{Justified}_e$(GÖDEL))). (PROVED IN PART I)
2) By the GÖDEL PROPERTY, 1) is equivalent to $\mathit{Justified}_e$(GÖDEL), and we know this equivalence, so it suffices to show that $\mathit{Justified}_e$(GÖDEL) leads to a contradiction. So assume $\mathit{Justified}_e$(GÖDEL). Then (using also the assumption that we are making throughout, that $\mathit{Justified}_e$(COMPETENCE)

and Axiom A3), *Justified*$_e$ (T_k $\Rightarrow$*Justified*$_e$(GÖDEL)). But we assumed *Justified*$_e$($\neg$(T_k $\Rightarrow$*Justified*$_e$(GÖDEL))), and this violates our consistency axiom A2). This completes the proof of our Anti-Chomskian Incompleteness Theorem.

References

[1] Jeff Paris and Leo Harrington, "A Mathematical Incompleteness in Peano Arithmetic", in *Handbook of Mathematical Logic*, J. Barwise, ed. (North Holland, 1977), pp. 1133–1142.

[2] H. Putnam, "What Is Mathematic Truth", in my *Philosophical Papers, vol. 1, Mathematics, Matter and Method* (Cambridge University Press, 1975), pp. 60–78.

[3] H. Putnam, "A Comparison of Something With Something Else", collected in my *Words and Life* (Harvard University Press, 1994) and my *Representation and Reality* (MIT Press, 1988).

[4] H. Putnam, "Nonstandard Models and Kripke's Proof of the Gödel Theorem," *Notre Dame Journal of Formal Logic*, vol. 41, no. 1 (2000), pp. 53–58.

[5] H. Putnam, "An Unsolvable Problem in Number Theory," *The Journal of Symbolic Logic 25* (1960), 220–232.

[6] H. Putnam, "Reflexive Reflections", *Erkenntnis* 22 (1985), pp. 143–153 [collected in my *Words and Life* (Harvard University Press, 1994).]

Received 15 June 2004