REMARKS ON THE UNKNOWABLE

by
Harvey M. Friedman
Ohio State University
Gödel Centenary
April 28, 2006
revised May 22, 2006

I will talk about a specific candidate for unknowability. I have attempted to make this candidate as widely interesting as I can on short notice.

A number of formal results and conjectures have emerged, and it appears that the approach here has opened up some new lines of research.

I will focus attention on mathematical candidates for unknowability. There is, of course, the wider topic of the unknowability of propositions involving physical objects or other kinds of nonmathematical objects.

The kind of unknowability I will discuss concerns

the count of certain
natural finite sets of objects.

Even the situation with regard to our present strong formal systems is rather unclear. One can just profitably focus on that, putting aside issues of general unknowability.

Many of the ideas presented here are present in work of Chaitin, although in a different form. We haven't looked at the overlap. In particular, we propose that our transition systems is a particularly good vehicle for developing these ideas. Also the idea of exploiting special features of the standard axiom systems used for the foundations of mathematics, in this context, seems novel.

TRANSITION SYSTEMS

A 1 dimensional transition system, 1TS, is given by a quadruple (S,a,b,f), where

1. S is a finite set.
2. $a,b \in S$, $a \neq b$.
3. $f:S^3 \to S$.

The evolution of (S,a,b,f) is given by $H:N \times Z \to S$, where

H(0,x) = a if x < 0; b o.w.
H(t+1,x) = f(H(t,x-1),H(t,x),H(t,x+1)).

In H(t,x), $t \geq 0$ is time, and $x \in Z$ is position. H(t,x) is the state at time t and position x.

At time t = 0, the negative positions are in state a, and the nonnegative positions are in state b.

At time t+1, the state at position x is determined by f from the states at x-1,x,x+1 at time t.

Stabilization of (S,a,b,f) occurs at the least t such that for all $x \in Z$, H(t,x) = H(t+1,x).

A stable 1TS is a 1TS in which stabilization occurs.

Turing machines are naturally special cases of 1TS. The 1TS correspond to "systolic arrays". One can also think of 1TS's as a kind of discrete dynamical system.

The 1TS have an obvious extension to higher dimensions d. Here Z is replaced by Z^d, and $f:S^{2d+1} \to S$, as there are 2d neighbors of each lattice point in Z^d.

For d = 2, this corresponds to cellular automata. For specificity, focus on d = 1.

The size of a kTS is the number of states = |S|.

COUNT PROBLEMS

There is an obvious notion of isomorphism between 1TS. There is a bijection from S onto S' that sends a,b to a',b', and transforms f to f'.

We wish to count

$\Phi(1,n)$ = the number of stable 1TS of size n, up to isomorphism.

A very crude upper bound on the number of 1TS of size n is n^{n^3}. For n = 10, this is 10^{1000}. So obviously

$$\Phi(1,n) \leq 10^{1000} \leq 2^{3322}$$

and hence the count can be expressed in at most 1000 digits in base 10, or in at most 3322 digits in base 2.

Using merely that 1TS forms a complete model of computation, we can prove the following.

THEOREM 1. Let T be a recursively axiomatized formal system extending a weak fragment of arithmetic, that does not prove its own inconsistency. There exists n such that T cannot determine $\Sigma(1,n)$.

However, obviously the n in Theorem 1 can be arbitrarily large. We would like to focus on the situation for small n, so that the number of digits in $\Sigma(1,n)$ is manageable.

The key observation is a sharpening of Theorem 1.

THEOREM 2. Let T be a recursively axiomatized formal system extending a weak fragment of arithmetic, that does not prove its own inconsistency. Let n be such that there is a 1TS Φ of size n such that "Φ is stable $\Rightarrow$ Con(T)" is provable in T. Then T cannot determine $\Sigma(1,n)$.

COROLLARY 3. There exists a reasonably small positive integer c such that the following holds. ZFC cannot determine $\Sigma(1,c)$, unless ZFC proves its own inconsistency. This is true of any of the systems T extending ZFC by standard large cardinals (or even ZF + j:V $\prec$ V), unless T proves its own inconsistency.

PROBLEM: Give a small example of a c for which Corollary 3 holds for ZFC. What is the least c such that Corollary 3 holds for ZFC? Give a small example of a c for which Corollary 3 holds for the cited extensions of ZFC. What is the least c?

When I gave this talk at the Gödel Centenary, I expressed some confidence that Corollary 3 holds for c = 10, and gave the excuse of short notice for not being able to say anything definite. (The panel on Unknowability, in which I presented this talk was constructed shortly before the meeting).

I still cannot make any definite statement. I have a new excuse: I am trying to finish my BRT book by June 30, 2006.

But it still appears to me that c = 10 is an appropriate challenge, although it may be more difficult than I thought.

Under the weaker assumption of consistency, the coding involved blows up more easily. We still have:

THEROEM 4. There exists a positive integer c such that the following holds. ZFC cannot determine $\Sigma(1,c)$, unless ZFC is inconsistent. This is true of any of the standard systems T extending ZFC by large cardinals (or even ZF + $j:V \prec V$), unless T is inconsistent.

Let us say that we have c which Theorem 1.2 holds for ZFC. Let us say, for the sake of argument, that c = 10. Then we know that ZFC does not determine all of the at most 3322 base 2 digits for $\Sigma(1,c)$. Therefore, there exists $1 \leq i \leq 3322$ such that ZFC does not determine the i-th base 2 digit of $\Sigma(1,c)$. Can be name such an i? Does ZFC at least determine the position of the leading digit?

But what about 'absolute unknowability', since we may in the future "know" mathematical propositions that are in no way captured by the current systems?

Reasonable bounds for c can be obtained just from the fact that the usual systems have modest "entropy". Such bounds should be worse than what can be obtained for ZFC and related systems.

THEOREM 5. (Informal). Let T be a consistent formal system extending a weak fragment of arithmetic, not proving its own inconsistency, with "low entropy". There exists a "modest" n such that T cannot determine $\Sigma(1,n)$.

The "entropy" of existing systems has always been "low", and appears to have not increased significantly over many decades despite the emergence of many new and very strong axioms.

In order for us to "know" mathematics, we seem to need to derive it from clear, simple, and basic principles. The formulation of such principles seems very slow, never involving significant "raising of entropy".

One can attempt to model the emergence of new axioms by the brain. Although there may emerge an enormous number of new

brains, new axioms can be argued to have to meet such high standards that their validation must be strongly present in every suitably intelligent brain. This may be argued to be incompatible with "high entropy".

Such considerations point to the possible unknowability of $\Sigma(1,n)$ for modest n, even for n = 10.

It is possible to extend Theorem 1 to systems with full induction such as extensions of PA and extensions of ZF, just under the assumption of consistency (rather than unprovability of consistency). One may well get worse estimates on c.

ADDED MAY 22, 2006:

It is possible to extend Theorem 1 to any recursively axiomatized system T containing a weak fragment of arithmetic.

LEMMA 6. Let $f:N \to N$ be a presented elementary recursive function, and $S \subseteq N$ be a presented r.e. set. Assume that EFA proves $(\forall n)(\exists n(n) = 0 \to f(n) \in S)$. There exists e such that EFA refutes: $e \in S \to g(e) = 0$.

LEMMA 7. There is a presented elementary recursive function $g:N \to N$ such that EFA proves $\exists n(n) = 0 \to g(n)$ is the Godel number of a 1TS that stabilizes.

LEMMA 8. Let $f:N \to N$ be a presented partial recursive function. There exists a Godel number α of a 1TS such that EFA proves

$f(\alpha) = 0 \to \alpha$ stabilizes.

THEOREM 9. Let T be a consistent recursively axiomatized system extending EFA. There exists n such that T cannot determine $\Sigma(1,n)$.

We sketch the proof. Fix T and assume that for all n, T determines $\Sigma(1,n)$.

We define a partial recursive function $f:N \to N$ as follows. Let α be the Godel number of a 1TS, and let n be its number of states. Find the proof with least Godel number of a determination in T of $\Sigma(1,n)$, and call the determination #(n). Now look for #(n) stabilizing 1TS with n states, up

to isomorphism, in the obvious effective way. If we find them, then set f(□) = 1 if □ is among them; f(□) = 0 if □ is not among them. If we don't find them, then f(□) is undefined. If □ is not the Godel number of any ITS, then set f(□) = 2.

By Lemma 8, let □ be (the Godel number of) a 1TS be such that EFA proves

f(□) = 0 □ □ stabilizes.

Let □ have n states. Let □ be the proof with least Godel number of a determination in T of □(1,n). Let #(n) be the determination in □. Then T proves that we will find #(n) stabilizing ITS with n states, up to isomorphism, in the obvious effective way. Furthermore, T proves that this lists all of the 1TS with n states, up to isomorphism, that stabilize. In particular, T proves the f(□) is defined.

According to T, if □ is among these, then f(□) = 1 and □ stabilizes. Also, according to T, if □ is not among these, then f(□) = 0 and □ does not stabilize. Hence T refutes

f(□) = 0 □ □ stabilizes.

Since EFA proves the above, T is inconsistent. Contradiction. QED

Obviously the proof of Theorem 9 applies equally well to quite general computational setups beyond that of 1TS's.