Internat. J. Math. & Math. Sci.
Vol. 7 No. 2 (1984) 283-289

FIXED POINTS AND SELF-REFERENCE

RAYMOND M. SMULLYAN

Department of Philosophy
Indiana University
Bloomington, Indiana 47405 U.S.A.

(Received May 15, 1982)

ABSTRACT. It is shown how Gödel's famous diagonal argument and a generalization of the recursion theorem are derivable from a common construation. The abstract fixed point theorem of this article is independent of both metamathematics and recursion theory and is perfectly comprehensible to the non-specialist.

KEY WORDS AND PHRASES. Diagonal ization, fixed point, self reference, equivalence relation.

1980 MATHEMATICS SUBJECT CLASSIFICATION CODE. 47H10

1. INTRODUCTION

In Gödel's seminal 1931 paper [1], he proves the incompleteness of a particular formal system--that of Principia Mathematica. He states, however, that his method is applicable not only to that particular system (as well as the Zermelo-Fraenkel system of set theory) but to an extensive variety of systems. Just what is this "extensive variety"? Various interpretations of this phrase have been given, and Gödel's argument has accordingly been generalized in many ways. Curiously enough, one of the ways that is the most direct and most easily accessible to the general mathematical reader is also the way that appears to be the least generally known. What is particularly curious is that the way in question is the very way indicated in the introductory section of Gödel's original paper! However, Gödel apparently did not regard this introductory section as anything more than a heuristic sketch, because it involved not only the notion of provability, but also the notion of <u>truth</u>, which had not yet been formalized (it was formalized later by Alfred Tarski [2]). Since the notion of truth has been formalized, it is now possible to reformulate Gödel's "heuristic" sketch as a precise theorem, which is what we do in §2 of this article. Gödel's argument uses an ingenous device known as <u>diagonalization</u>. A closely related technique underlines a basic result in Recursion Theory known as the Recursion Theorem (Kleene, [3]). We present a simple abstract version of this theorem in §3. Next, in §4, we consider a related fixed point result-a variant of the Mocking Bird Puzzle, which appeared in [4]. We conclude with a demonstration of how these apparently diverse results are all derivable from a common basic fixed point theorem.

2. AN ABSTRACT FORM OF GÖDEL'S DIAGONAL ARGUMENT

The axiom systems for which Gödel's argument goes through all possess the following features: There is a well defined set S of expressions called sentences, a well defined subset T of S whose elements are called true sentences, and a well defined subset P of T whose elements are called provable sentences (provable in the system, that is). Certain sets of positive integers are called nameable in the system; there are denumerably many such nameable sets, and they are arranged in a specified infinite sequence $A_1, A_2, A_3, \ldots$. We call a positive integer n an index of a nameable set A if $A_n = A$. (In general, a nameable set will have infinitely many different indices.) There is also given a function $S(x,y)$ which assigns to each ordered pair $<i,j>$ of positive integers a sentence $S(i,j)$ which is a true sentence (element of T) if and only if i is a member of the set A_j .

Each sentence X is assigned a positive integer $g(x)$ called the Gödel number of X . For any set W of sentences, by $g(W)$ we mean the set of Gödel numbers of all the sentences in W . For any positive integers i , j , we let $i*j$ be the Gödel number of the sentence $S(i,j)$. We let $d(x)$ be the number $x*x$. (The function d is sometimes called the diagonal function of the system. We note that for any i , the number $d(i)$ is the Gödel number of the sentence $S(i,i)$, which is true if and only if i belongs to A_i .) For any set A of numbers (positive integers) by $d^{-1}(A)$ is meant the set of all positive integers i such that $d(i) \in A$. Thus the statements $i \in d^{-1}(A)$ and $(i*i) \in A$ are equivalent. By the complement $\bar{A}$ of A is meant the complement with respect to the set of positive integers — thus $\bar{A}$ is the set of all positive integers not in A .

We let (S) be the system consisting of the sets S , T , P , the enumeration of the nameable sets, the function $S(x,y)$, and the Gödel numbering g of the sentences. We shall call the system rich if the set $g(P)$ of Gödel numbers of the provable sentences is one of the nameable sets. We shall call (S) complemented if the complement of every nameable set is again nameable. We shall call (S) diagonalizable if for every nameable set A , the set $d^{-1}(A)$ is also nameable in (S) .

Theorem A (An Abstract Form of Gödel's Theorem).

Every rich, complemented diagonalizable system must contain a true sentence which is not provable in the system.

The conclusion of Theorem A says nothing more nor less than that there is at least one element of T not in P . As will be seen from the proof, there are numbers i , j such that $S(i.j)$ is such a sentence; thus i is a member of the set A_j , but the sentence $S(i,j)$ is not provable in the system. Also, if we are given a number p which is an index of the set $g(P)$ and if given any n , we can effectively find an index n' of the complement of A_n , and if given any n we can effectively find a number $n^{\#}$ which is an index of the set $d^{-1}(A_n)$ (and these "effective" conditions do indeed hold for the various systems to which Gödel's arguments have been applied), then we can actually find numbers i , j such that the sentence $S(i,j)$ is true but not provable in the system. As an instructive illustration of this, suppose $p = 2$, and for every n , $n' = 2^n$ and $n\# = 3^n$. Now the reader has enough information to actually find numbers i, j (either the same or different) such that i is a member of A_j, but the sentence $S(i,j)$ (which expresses this fact) is not provable in the system. (This is an instructive puzzle rather like that posed in [5]. There are infinitely many solutions, and at least two in which i, j are both less than 600.)

Gödel Sentences

Before proving Theorem A (which, incidentally, solves the above puzzle), we will consider a basic property of diagonalizable systems.

We call a sentence X a Gödel sentence for a set A of positive integers if either X is true and its Gödel number is in A, or X is false and its Gödel number is not in A.

Theorem 1 — [The Diagonal Lemma]

If the system (S) is diagonalizable, then for any nameable set A, there is a Gödel sentence for A.

Proof

Take any nameable set A_n. By the hypothesis of diagonalizability, there is a number $n^\#$ such that $A_{n^\#} = d^{-1}(A_n)$. This means that for all x, $x \in A_{n^\#}$ if and only if $(x * x) \in A_n$. And so for every x, the sentence $S(x,n^\#)$ is true if and only if $(x * x) \in A_n$. Therefore $S(n^\#,n^\#)$ is true if and only if $(n^\# * n^\#) \in A_n$. But $n^\# * n^\#$ is the Gödel number of the sentence $S(n^\#,n^\#)$! And so $S(n^\#,n^\#)$ is a Gödel sentence for A_n.

Proof of Theorem A

Assume hypothesis. Since the system is rich, there is a number p such that $A_p = g(P)$. Then by complementation, there is a number p' such that $A_{p'} = \overline{g(P)}$. Since the system is diagonalizable, then by Theorem 1, there is a Gödel sentence G for the set $A_{p'}$ (namely $S(p'^\#,p'^\#)$, where $p'^\#$ is any index of $A_{p'}$). This sentence G is true if and only if $g(G) \in A_{p'}$, and so G is true if and only if $g(G) \in \overline{g(P)}$. Also $g(G) \in \overline{g(P)}$ if and only if $G \notin P$. And so G is true if and only if G is not provable in the system. Therefore, either G is true and not provable in the system, or G is not true but provable in the system. The latter alternative is ruled out by the assumption that $P \subseteq T$. Therefore G is true but not provable in the system.

Exercise

Show that for any number k which is an index of the set $d^{-1}(\overline{g(P)})$, the sentence $S(k,k)$ is true but not provable in the system. What about the sentence $S(p^{\#'},p^{\#'})$, where $p^\#$ is an index of $d^{-1}(A_p)$, and $p^{\#'}$ is an index of the complement of $A_{p^\#}$; can its truth be determined? Can its provability be determined?

Tarski's Theorem

The diagonal lemma has another important consequence: Consider a system (S) which is complemented and diagonalizable without necessarily being rich. Can it be determined whether the set $g(T)$ (the set of Gödel numbers of the true sentences) is nameable in the system? As was shown by Tarski [2], $g(T)$ is not nameable in the system. Here is the argument.

Suppose $g(T)$ were nameable in the system. Then by complementation $\overline{g(T)}$ would also be nameable. Then by the diagonal lemma, there would be a Gödel sentence X for $\overline{g(T)}$, and we would have $X \in T$ if and only if $g(X) \in \overline{g(T)}$, but $g(X) \in \overline{g(T)}$ if and only if $X \notin T$, and so we would have the absurdity that X is in T if and only if X is not in T. Therefore $g(T)$ is not one of the nameable sets.

It should be of interest to note that this result of Tarski provides an alternative proof of Gödel's theorem: Suppose (S) is rich, complemented, and diagonalizable. By Tarski's result, the set $g(T)$ is not nameable in the system, but by richness, the set $g(P)$ <u>is</u> nameable in the system. Therefore P, T must be different sets. Since $P \subseteq T$, then T must contain a sentence not in P, which alternatively proves Theorem A.

3. AN ABSTRACT RECURSION THEOREM

Consider now a denumerable set of any objects whatsoever arranged in an infinite sequence $E_1, E_2, \ldots, E_n, \ldots$. Let Σ be a collection of functions from the positive integers to the positive integers. Σ is said to be closed under composition if for any functions f, g in Σ, there is a function h in Σ such that for all (positive integers) x, $h(x) = f(g(x))$. We shall also consider a function $F(x,y)$ from the set of ordered pairs of positive integers to the positive integers.

<u>Theorem 2</u>

Suppose the following three conditions hold:

C_1: Σ is closed under composition.

C_2: The function $F(x,x)$ is in Σ.

C_3: For any $f \in \Sigma$, there is a positive integer a such that for all x, $E_{F(a,x)} = E_{f(x)}$.

<u>Conclusion</u>: For any $f \in \Sigma$ there is at least one positive integer i such that $E_i = E_{f(i)}$.

<u>Proof</u>

Take any function f in Σ. By C_2, the function $F(x,x)$ is in Σ, and so by C_1, the function $f[F(x,x)]$ is in Σ. Then by C_3, there is a number a such that for all x, $E_{F(a,x)} = E_{f(F(x,x))}$. Taking a for x, it follows that $E_{F(a,a)} = E_{f(F(a,a))}$. We take $F(a,a)$ for i, and so $E_i = E_{f(i)}$.

<u>Discussion</u>

In applications to Recursion Theory, Σ is the class of recursive functions of one argument. This class is closed under composition, so C_1 holds. For one form of the recursion theorem, we take E_i to be the i^{th} partial recursive function of one argument (in a standard enumeration). By a result known as the <u>iteration theorem</u> there is a recursive function $F(x,y)$ satisfying C_3, and condition C_2 is automatic (because for a recursive function $G(x,y)$, the function $G(x,x)$ is recursive). Then by Theorem 2, for any recursive function $f(x)$ there is an i such that the partial recursive function E_i is the same as the partial recursive function $E_{f(i)}$; this is one form of the Recursion Theorem.

4. A MOCKING BIRD PUZZLE

We next consider a variant of a problem posed in [4].

We are given a collection of birds. Given any birds B, C, if a spectator calls out the name of C to B, the bird B responds by calling back the name of some bird $B(C)$. (Thus each bird B induces a function from birds to birds.) If $B(C) = C$, then we say that B is fixated on C. We call B egocentric if B is fixated on itself. We are given that the set of functions induced by the birds is closed under composition (more explicitly, for any birds B, C there is a bird D such that for every bird X, $D(X) = B(C(X))$). We are also given that there is a bird M (called a mocking bird) such that for every bird B, $M(B) = B(B)$. The problem is to prove that every bird is fixated on at least one bird, and that at least one bird is egocentric. For the solution, take any bird B. By closure under composition, there must be a bird C such that for every bird X, $C(X) = B(M(X))$ (M is a mocking bird). Then taking C for X, $C(C) = B(M(C))$. Also $C(C) = M(C)$. Therefore $C(C) = B(C(C))$, and so B is fixated on the bird $C(C)$.

Since the mocking bird M is one of the birds, it is also fixated on some bird E. Thus $M(E) = E$, but also $M(E) = E(E)$. Therefore $E(E) = E$, and so E is egocentric. (Incidentally, removing the parentheses from "M(E)" tells its own tale.)

5. A GENERAL FIXED POINT THEOREM

Now we come to the finale: Theorems 1, 2, as well as the solution of the Mocking Bird Puzzle, are all derivable from a common construction.

It is usual to call an element x a fixed point of a function f if $f(x) = x$. The first thing we shall do is to generalize this notion: Consider a set N, a function f from N into N and an equivalence relation $\equiv$ on N. We shall say that an element n of N is a fixed point of f with respect to the equivalence relation if $f(n) \equiv n$. Under this extended sense of "fixed points" we will see that Theorems 1 and 2 are indeed fixed point theorems.

Our general setup is this: We consider a set N, an equivalence relation $\equiv$ on N, and a collection Σ of functions from N into N. For any function $F(x,y)$ from ordered pairs of elements of N to elements of N, we shall say that F enumerates Σ with respect to $\equiv$, if for every $f \in \Sigma$ there is some $a \in N$ such that for all $x \in N$, $F(a,x) \equiv f(x)$. Lastly, given any function h from N into N (but not necessarily one of the functions in Σ), we shall say that h is admissible (with respect to Σ and $\equiv$), if for every $f \in \Sigma$, there is some $f' \in \Sigma$ such that for all x in N, $f'(x) \equiv f(h(x))$. (We might note that for the special case that $\equiv$ is the identity relation, the statement that every element of Σ is admissible is equivalent to the statement that Σ is closed under composition.)

Theorem F — (A General Fixed Point Theorem)

A sufficient condition that every element of Σ has a fixed point with respect to $\equiv$ is that there is a function $F(x,y)$ with the following two properties:

P_1: $F(x,y)$ enumerates Σ (with respect to $\equiv$) .
P_2: $F(x,x)$ is admissible (with respect to Σ , $\equiv$) .

Proof

Suppose $F(x,y)$ is a function satisfying P_1 , P_2 . Take any $f \in \Sigma$. By P_2 , there is some $f' \in \Sigma$ such that for all $x \in N$, $f'(x) \equiv f(F(x,x))$. By P_1 , there is some $a \in N$ such that for all x , $F(a,x) \equiv f'(x)$. Hence for all x , $F(a,x) \equiv f(F(x,x))$. Taking a for x , $F(a,a) \equiv f(F(a,a))$, and hence $f(F(a,a)) \equiv F(a,a)$. Letting $b = F(a,a)$, $f(b) \equiv b$, and so b is a fixed point of f with respect to $\equiv$.

Applications

(1) The application of Theorem F to the Mocking Bird Puzzle is quite obvious: We clearly take N to be the set of birds, Σ to be the set of functions induced by the birds, $\equiv$ to be identity, and for any birds X , Y we define $F(X,Y)$ to be $X(Y)$. The rest should be transparent.

(2) To obtain Theorem 2 as a corollary of Theorem F, we take N to be the set of positive integers. We define $i \equiv j$ if $E_i = E_j$. The collection Σ and the function $F(x,y)$ are already given. Condition C_3 of the hypothesis of Theorem 2 says that for every $f \in \Sigma$ there is some $a \in N$ such that for all $x \in N$, $F(a,x) \equiv f(x)$ — in other words that $F(x,y)$ enumerates Σ with respect to $\equiv$. Conditions C_1 , C_2 jointly imply that for every $f \in \Sigma$, the function $f(F(x,x)) \in \Sigma$, and so taking $f'(x)$ to be $f(F(x,x))$, $f' \in \Sigma$, and of course $f'(x) \equiv f(F(x,x))$ (since $f'(x) = f(F(x,x))$). This means that $F(x,x)$ is admissible (with respect to Σ , $\equiv$) . And so C_1 implies property P_1 , and C_2 , C_3 jointly imply property P_2. Therefore by Theorem F, for every $f \in \Sigma$, there is some $i \in N$ such that $f(i) \equiv i$, which means that $E_{f(i)} = E_i$.

(3) The application to Theorem 1 is perhaps the least obvious, and (in our view) the most interesting: For Theorem 1, we again take N to be the set of positive integers. Let us call two sentences <u>equivalent</u> if either they are both true (both in T) or both false (both outside T). And for any two positive integers i , j , we define $i \equiv j$ to mean that either i , j are both Gödel numbers of equivalent sentences, or that neither i nor j is a Gödel number of a sentence. This defines the relevant equivalence relation on N . For each i , define $f_i(x) = x * i$, and let Σ be the collection of all functions f_i , as i ranges over N . Then define $F(i,j) = f_i(j)$. It is trivial that the function $F(x,y)$ enumerates Σ with respect to the identity relation, hence also with respect to the equivalence relation $\equiv$.

Now, the hypothesis of diagonalizability implies (in fact is equivalent to) the statement that the function $F(x,x)$ is admissible. To see this, we first note that $d(x) = F(x,x)$ (since $d(x) = x * x = f_x(x) = F(x,x)$). Now suppose the system is diagonalizable. Then for any i there is some j such that $A_j = d^{-1}(A_i)$. Now take any $x \in N$. Then $x \in A_j$ if and only if $d(x) \in A_i$, hence $x \in A_j$ if and only if $F(x,x) \in A_i$. Therefore the sentences $S(x,j)$ and $S(F(x,x),i)$ are equivalent sentences, which means that $x * j \equiv F(x,x) * i$, and hence $f_j(x) \equiv f_i(F(x,x))$, which means that $F(x,x)$ is admissible. (Conversely, the admissibility of $F(x,x)$ implies diagonalizability, as the reader can verify by reversing the steps of the above argument, but we do not need this fact in what follows.) And so by Theorem F, if (S) is diagonalizable, then for every i, the function f_i has a fixed point n with respect to $\equiv$. But the statement that n is a fixed point of f_i (with respect to $\equiv$) implies (in fact is equivalent to) the statement that n is the Gödel number of a sentence which is a Gödel sentence for A_i, because suppose it is the case that $n \equiv f_i(n)$. Then $n \equiv n * i$. Since $n * i$ is a Gödel number — namely of the sentence $S(n,i)$ — then n is also a Gödel number of some sentence X which is equivalent to $S(n,i)$. But $S(n,i)$ is true if and only if $n \in A_i$. Therefore X is true if and only if its Gödel number n belongs to A_i, which means that X is a Gödel sentence for A_i. (Conversely, if X is a Gödel sentence for A_i, then its Gödel number is a fixed point for f_i with respect to $\equiv$, as the reader can easily verify.)

We now see how Theorem F generalizes both Theorem 1 and Theorem 2 (as well as the Mocking Bird Puzzle). We have found a host of variants and generalizations of Theorem F, about which a full-length monograph [6] is currently in preparation. A number of related fixed point theorems and their applications to other forms of Godel's argument (including purely syntactic forms, which do not employ the notion of truth) can be found in [7].

Here is one generalization of Theorem F which we would like to leave with the reader as a puzzle: Suppose we have a collection Σ of binary relations on a set N. Suppose we have a function $d(x)$ from N into N satisfying the following two conditions: (1) For every $R \in \Sigma$, there is at least one $i \in N$ such that $R(i,d(i))$; (2) For every $R \in \Sigma$ there is some $R' \in \Sigma$ such that for all x, y in N, $R'(x,y)$ implies $R(d(x),y)$. The first problem is to prove that for every R in Σ, there is at least one $i \in N$ such that $R(i,i)$. Then show how this generalizes Theorem F.

REFERENCES

1. GÖDEL, KURT [1931]. Über formal unentscheidbar Sätze der Principia Mathematica und verwandter Systeme I, Monatshafte für Mathematik und Physik 38, 173-198.

2. TARSKI, ALFRED [1933]. Der Wahrheitsbegriff in den formalisterten Sprachen, Studia Philosophica 1 (1936), 261-405.

3. KLEENE, STEPHEN C. [1938]. On notation of ordinal numbers, Journal Symbolic Logic 3, 150-155.

4. SMULLYAN, RAYMOND M. [1978]. In: Mathematical Intelligencer 1, No. 2, 111.

5. SMULLYAN, RAYMOND M. On Insolubilia, Mathematical Intelligencer (to appear).

6. SMULLYAN, RAYMOND M. Fixed Points, Diagonalization and Self-Reference (in prep.)

7. SMULLYAN, RAYMOND M. The Lady Or the Tiger? — and other logical puzzles. Alfred A. Knopf, New York, 1982.