A CURIOSITY IN THE PROOF OF GODELS THEOREM

David McGoveran

ABSTRACT: Since the publication of Goedel's famous paper in 1931 entitled "On Formally Undecidable Propositions of <u>Principla Mathematica</u> and Related Systems I", the proposal that classical mathematics be formulated as a formal axiomatic theory, and that the theory should be proved to be free of contradiction, has not been the subject of any serious and concerted effort within the mathematical and philosophical communities. The original proof of Goedel's theorem is re-examined. Certain assumptions concerning the independence of instantiations in the proof of Goedel's Incompleteness. Theorem will be shown to lead to the curious conclusion that the actual instantiations chosen by Goedel are contradictory.

1. INTRODUCTION

Formalism or formal exiomstics suffered a severe blow with the publication of "On Formally Undecidable Propositions of <u>Principia</u>

<u>Mathematics</u> and Related Systems I" [2][4][6]. The achievements in this landmark paper were numerous, from the development of a method for encoding

the system of Principia Mathematica in arithmetic (elementary number theory) to demonstrations that forty-five number-theoretic predicates are primitive recursive. The key theorem of the paper (Theorem VI) was to become famous and the construction of the proof, unique and unusual though it was, to be deemed flaxless.

Stated simply, the theorem shows that any system capable of representing elementary number theory would necessarily include undecidable propositions and thus, demonstrated that classical mathematics could not be proved to be consistent—i.e., free from contradiction. This result ended the search for a proof that classical mathematics could be completely axiomatized in a consistent formal system [4]. Since that time, numerous derivation of the result and its extensions have been published [1][3][7][9][10][1]].

We will point out the two lines of the proof which are not directly derived by some rule of the system from a previous line. These lines involve the choice of an instance of a free variable. It is assumed by the proof that these choices are free: if the universal is true then certainly any instance is true. By exploring the effect of different choices for the instance of the free variable, we will examine four possible choices which constitute the cases without significantly altering Goedel's construction.

In the sections which follow, we will examine Goedel's proof of Theorem VI in its original form (Section 2) with the exception that all details of the proof are laid bare. This constitutes case one: the two instances of a Universal chosen by Goedel are the instances chosen for this case. The "extra" lines of the proof are noted in the text with an asterisk next to the line number. Case Two is a reformulation of the proof with a simple change in notation (Section 3.1) which serves to treat

their relationship. In Case Three, the common variable is NEGATED. This serves to demonstrate that the reformulation is consistent with a change of variables (Section 3.2). The resulte of Goedel's formulation are shown to be obtainable in the reformulation through an INCONSISTENT choice of VARIABLES in the proof's two assumptions (Case Pour), thus demonstrating that the assumptions as originally formulated are logically inconsistent (Section 3.3) through the isomorphism established by Goedel.

The four formulations (or cases) given in this paper will use a consistent enumeration of the key expressions used in the proof for ease of comparison between sections and with the reference from which the proof was drawn [5]. We will also use the notation invented by Goedel, several points concerning which will aid the reader:

- (1) words in UPPERCASE refer to number theoretic statements;
- (2) Sb denotes the SUBSTITUTION in the formula which follows of the subscripted symbol in place of the superscripted symbol;
 - (3) Neg denotes NEGATION as does a bar;
 - (4) Gen denotes GENERALIZATION:
 - (5) x B y denotes x is a proof sequence of y;
 - (6) Bew denotes *it is provable that"; and
 - (7) %(x) denotes the NUMERAL for the formula x,

A more precise and complete understanding of the notation may be obtained by reading Goedel [5]. In summary, we will argue that these demonstrations are inexplicable in the light of the validity of Goedel's proof.

2. GOEDEL'S PROOF

Goedel's proof of Theorem VI depends upon the results proven earlier

in the paper. For the sake of brevity, we will omit the detailed demonstrations and simply state the results, accepting the proofs given by Goedel, since these details do not bear on the topic of this paper. To be explicit, Goedel leads up to Theorem VI by giving

- (1) a precise description of the system P whose axioms are those of Principis Mathematics with the adjunction of the Peano axioms,
- (2) an assignment of natural numbers to sequences of signs and sequences of sequences of signs of P (establishing an isomorphism between a subset of the natural numbers and the formulas of P),
- (3) a definition of primitive recursive functions and four theorems about them,
- (4) the proof that forty-five number-theoretic predicates are primitive recursive,
- (5) the proof of Theorem V--i.e., that every primitive recursive number-theoretic predicate is numeralwise representable in P, and
- (6) the definition of w-consistency.

Goedel then comes to the goal of the discussione: Theorem VI and its proof.

With the exception of Theorem V, Goedel uses the results presented in the paper exactly as formulated. Because the exact formulation of Theorem V will have a bearing on our discussion of the proof of Theorem VI, we repeat it here, though not the remainder of the discussions leading up to Theorem VI.

THEOREM V. For every recursive relation R(x ,...,x) there exists an l n-place RELATION SIGN r (with the FREE VARIABLES u , u ,..., u) such that for all n-tuples of numbers (x,...,x) we have

2.1 CASE ONE

The general result about the existence of undecidable propositions follows now in the formulation as given by Goedel. This will be referred to in the remainder of the paper as Case One.

THBUREN VI. For every w-consistent recursive class k of FORMULAS there are recursive CLASS SIGNS r such that neither v Gen r nor Neg(v Gen r) belong to Fig(k) (where v is the FREE VARIABLE of r).

PROOF (Case One): Let k be any recureive w-consistent class of FORMULAS, we define

Bw
$$(x) = (n)(n \le 1(x) \implies Ax(n Gl x) \lor (n Gl x) \in k \lor k$$

$$(Ep,q)\{0 < p,q < n \in Fl(n G) \times p Gl \times q Gl \times)\}\} \in I(x) > 0$$
 (II.5)

$$x B y = Bw (x) A [1(x)] G1 x = y$$
 (II.6)
 $k k$

$$Bes_{k}(x) = (Ey)y B_{k} x$$
 (11.6.1)

We obviously have

$$(x) [Rew (x) = x \in Flg(k)]$$
 (11.7)

and

$$(x)[Bev(x) \rightarrow Bev(x)]$$
 (II.8)

We now define the relation

$$Q(x,y) = x B (Sb(y))1.$$

$$k Z(y)$$
(11.8.1)

Since x B y (by II.5) and (II.6)) and Sb(y) (by Definition 17 and x 2(y)

31) are reducative, so is Q(x,y). Therefore, by Theorem V and (II.8) there is a RELATION SIGN q (with the FREE VARIABLES 17 and 19 such that

<u> ಕ್ರಾರ್</u>

₩a put

$$p = 17 \text{ Gen } q \tag{II.11}$$

(p is a CLASS SIGN with the FREE VARIABLE 19) and

$$r = Sb(q)$$

$$2(p)$$

$$(11.12)$$

(x is a recursive CLASS SIGN with the FREE VARIABLE 17).

Then we have

17 19 17 Bb(q) = Sb(r) (II.14)
$$z(x)z(p)$$
 $z(x)$

(by (II.12)). We now substitute p for y in (II.9) and (II.10)

19 17 19

x B [Sb(p)]
$$\rightarrow$$
 Bew [Neg(Sb(q))] *(II.14.3)

k $z(p)$ **k** $z(x)z(p)$

and take (II.13) and (II.14) into account to obtain

$$\begin{array}{c}
\overline{x} \xrightarrow{B} (17 \text{ Gen } r) \xrightarrow{} \text{Bew } \{\text{Reg(Sb(r))}\}. \\
x \xrightarrow{k} x \xrightarrow{} x \xrightarrow{}$$

From the contradictions of (II.15) and (II.16) the w-inconsistency of the system P is readily argued by invoking the law of the excluded middle.

A few comments on the nature of the proof of Theorem VI are in order. Note that the proof involves the choice of two assumptions: first, (II.S.1) defines a relation Q(x,y) with variables x and y assumed free; second, (II.11) and (II.14.1) assume that the particular substitution of p for y is not inconsistent with (II.3.1). An investigation of these assumptions will occupy the remaining sections of this paper as Cases Two through Four.

3. REFORMULATION

An examination of the relation Q(x,y) as defined in (II.B.1) reveals that it is the negation of a recursive relation which we define as

$$R(x,y) = x \beta [Sb(y)].$$
 (III.4.1)

Clearly,

$$Q(x,y) = R(x,y). \tag{III.4.2}$$

Purthermore, it is clear from Theorem V, the expression (II.9), and the expression (II.18) that the RELATION SIGN q assumes

$$\begin{array}{c} 19 \\ \text{Neg[Sb(y])}, \\ \mathbb{Z}(y) \end{array}$$

3.1 CASE TWO

In this section we will follow the proof construction of Theorem VI (Case One) with R(x,y) in place of Q(x,y) and therefore q' (of as yet unknown relationship to q) in place of q. As we shall see in this section and the next, our reformulation leads to consistent results. We proceed with the construction from (II.8.1), as this is the first notational change to be encountered.

CONSTRUCTION (Case Two): Let k be any recursive w-consistent class of FORMULAS. We define (III.5) through (III.8) precisely as *II.5) through (II.8), respectively. We now define the relation

$$\mathbb{E}(x,y) = x B \{Sb(y)\}\}.$$
 (III.0.1)

Since $x \to y$ (by (III.5) and (III.6)) and 9b(y) (by Definition $x \to y$

17 and 31) are recursive, so is R(x,y). Therefore, by Theorem V and

(III.6) there is a RELATION SIGN q^{τ} (with the FREE VARIABLES 17 and 19) such that

and

We put

$$p = 17 \text{ Gen } q^*$$
 (111.11)

(p is a CLASS SIGN with the FREE VARIABLE 19) and

$$x = \mathcal{B}b(q^{1})$$

$$\mathbf{z}(p)$$
(111.12)

(r is a recursive CLASS SIGN with the FREE VARIABLE 17).

Then we have

19
Sb(p) = Sb({17 Gen q'}) = 17 Gen Sb(q') = 17 Gen r
$$Z(p)$$
 $Z(p)$ $Z(p)$ (111.13)

(by (llI.ll) and (Ill.12)); furthermore

$$Sb(q^{*}) = Sb(x)$$

$$g(x)g(p) = g(x)$$
(111.14)

(by (III.12)). We now substitute p for y in (III.9) and (III.10)

and take (III,13) and (III.14) into account to obtain

$$x \rightarrow B$$
 (17 Gen x) $\rightarrow Bew (Sb(r)).$ (III.)5)

Obviously, this reformulation does not prove Theorem VI. Note that this construction is essentially that used by Goedel with one exception. Lines III.9 and III.10 follow Theorem V exactly and do not embed a NEGATION as did the construction Goedel used ((II.9) and (II.10)). It is curious that something as simple as following the details of Theorem V would undo the famous result of the basic proof schema. Furthermore, this reformulation and the results ((III.15) and (III.16)) are obviously unchanged when q' is replaced by q.

3.2 CASE THREE

Although made a bit more complicated, the reformulation and the results of Section 3.1 remain unchanged when q' is replaced with Neg q as follows.

CONSTRUCTION (Case Three): Let k be any recursive w-consistent class of FORMULAS. We define (IV.5) through (IV.8) precisely as (II.5) through (II.8), respectively. We now define the relation

$$R(x,y) = x B \{Sb(y \}\},$$

$$k \quad 2(y)$$

Since x B y (by (IV.5) and (IV.6)) and Sh(y) (by Definition 17 and z(y)

31) are recursive, so is R(x,y). Therefore, by Theorem V and (IV.8) there is a RELATION SIGN q^* (with the PREE VARIABLES 17 and 19) such that q^* =Neg q and

ងរាជី

With eome minor rearrangement, these expressions yield the Goedel expressions (11.9) and (11.10):

and

Note that (IV.8.2) leads to (IV.10) and (IV.8.3) leads to (IV.9). We put

$$p = 17 \text{ Gen}(q^*) = 17 \text{ Gen (Neg q)}$$
 (1V.11)

(p ie a CLASS SIGN with the FREE VARIABLE 19) and

19

$$I = Sb(q') = Sb(Neq(q))$$
 (IV.12)
 $I(p) = I(p)$

It is a recursive CLASS SIGN with the PRES VARIABLE 17).

Then we have

(by (IV.11) and (IV.12)); furthermore

(by (IV.12)). We now substitute p for y in (IV.9) and (IV.10)

19 17 19
$$x \in [Sb(p)] \rightarrow Bew[Sb(Neg q)] \qquad *(TV,14,2)$$

$$k = 2(x) \qquad k \qquad 2(x)2(p)$$

and take (IV.13) and (IV.14) into account to obtain, after some rearranging,

$$x \in [17 \text{ Gen } r] \rightarrow \text{Bew } \{\mathcal{E}b(r)\},$$
 (IV.15)

It is easy to see that the substitution of Neg q for q' yields (IV.9) and (IV.10) equivalent to (II.9) and (II.10) in Goedel's proof. Note that, up to (IV.10), this construction is different from that due to Goedel only in the choice of the recursive relation to be used. Nevertheless, we have recovered the expressions given by Goedel through a suitable choice of q', demonstrating that we can counter the differences in the construction. In particular, note that the relationship between Q(x,y) and R(x,y) is precisely that between q and Neg q. Thus we have shown that the y of Goedel's recursive relation Q(x,y) necessitates Neg q if one is to obtain (II.9) and (II.10). Even so, we have not as yet fully recovered Goedel's result, but have shown our construction to be a valid one even with a consistent change of definition for the RELATION SIGN as would be expected.

3.3 CASE FOUR

We now proceed to demonstrate how to obtain the results of Section 2.1 (Goedel's proof of Theorem VI) using the reformulation given in Section 3.2. Once again we will manipulate the two instantiations of the proof—namely, the choice of Q(x,y) and p.

CONSTRUCTION (Case Pour): Let x be any recursive w-consistent class of PORMULAS. We define (V.5) through (V.6) precisely as (II.5) through (II.6), respectively. We now define the relation

$$R(x,y) = x B [Sb(y)]$$
 (V.8.1)
k $Z(y)$

Since x B y (by V.5) and (V.6)) and Sb(y) (by Definition 17 and x z(y) are recursive, so is R(x,y). Therefore, by Theorem V and (V.8) there

is a RELATION SIGN q^{\dagger} (with the FREE VARIABLES 17 and 19) such that $q^{\dagger}\text{-Neg}$ q and following the results of Section 3.3 ((IV.6.2) and (IV.6.3)) we obtain as before

and

$$\begin{array}{c} 19 \\ x \ B \ [Sb(y]) \end{array} \longrightarrow \begin{array}{c} 19 \\ Bew \ [Meg(Sb(q))] \end{array}))), \qquad (V.10)$$

In contrast to the choice of Section 3.3, we now put

$$p = 17 \text{ Gen g} \tag{V.11}$$

(p is a CLASS SIGN with the FREE VARIABLE 19) and

$$r = Sb(q)$$

$$Z(p)$$

$$(V.12)$$

(r is a recursive CLASS SIGN with the FREE VARIABLE 17).

Then we have

(by (V.11) and (V.12)): furthermore

17 19 17
Sb(q) * Sb(r) (V.14)
$$\Sigma(x)\Sigma(p)$$
 $\Sigma(x)$

(by (V.12)). We now substitute p for y in (V.9) and (V.10)

and take (V.13) and (V.14) into account to obtain

$$\begin{array}{c} x & B & (17 \text{ Gen } x) \rightarrow \text{ Bew [Sb(x))}, \\ k & z(x) \end{array}$$
 (V.15)

Note that we obtain (V.9) and (V.10) by introducing R(x,y) in place of Q(x,y) and that this forces the use of Neg q explicitly. With the single exception of this choice in the method of constructing (V.9) and (V.10), the construction is that of Goedel's original proof of Theorem V. Thus we argue that in the original construction of (I.9) and (I.10), y necessitates Neg q, but that this fact is hidden by the definition of Q(x,y) resulting in an unacknowledged negation. The existence of this negation is made obvious when $\{V.9\}$ and $\{V.10\}$ are derived as in the above.

Although the proof proceeds as in Goedel's construction, note that this construction assumes that it is q (instead of Neg q) that is necessitated by p (V.11) and thus by y also (V.14.1). In this way we see that the details of the proof assume both q and Neg q necessitated by y. This is an instance of the classic definition of an embedded contradiction in a proof schema.

Even if the NEGATION of q and q are not held to be contradictory, the problem of explaining this eituation does not go away. In particular, if Goedel's construction is correct and if Cape Four is also a valid reformulation of Goedel's Proof [8] correct, then either O(x,y) along with q as Goedel has used it, or R(x,y) along with Neg q and q may be used in constructing the proof. Case Four could be rewritten such that the

choice of R(x,y) is explicit, but the choice of Neq q is implicit. Then the only visible differences in the reformulations (Case One and Case Four) would be the use of R(x,y) in place of Q(x,y) of the proof. But this can not be, for R(x,y) contradicts Q(x,y) as given. Thus the leomorphism established in Theorem V falls, since contradictory recursive relations produce the same results. Given this embedded contradiction, it is surprising that Goedel should be able to generate an acceptable proof of the w-inconsistency of the system P.

4. SUMMARY

Goedel's theorem is based on (a) an isomorphism between the proofs and sentences of PM+Feano and a set of number-theoretic propositions; (b) the proof of certain relations between some number-theoretic propositions; (c) using the isomorphism (mentioned in (a)) to transform these into a statement about the incompatibility of various assumptions (w-consistency and decidability) in P.

We have shown that additional assumptions are involved in the proof as well: independence of instantiations and that certain variables are free. Since the proof establishes an embedding, the choice of variables (without a semantice being involved) revolves around affirmation (q'=q) and negation (q'=Neq q). As there are two lines in the proof schema at which one must select either q or Neq q for the instantiation, this leads to four cases: two in which the choices are the same (both being either q or Neq q) and two in which they are contrary (q and Neq q or Neq q and q). Goedel's result is obtained only in the cases in which the two are contrary. It is not obvious as to why this should be the case, if Goedel's result is valid.

It might be arqued that because (II.3) and (II.10) have been

demonstrated true for all positive integers, subject to the restrictions of TYPE, at the number theoretic level, that the question of consistency with (II.8.1) cannot be posed at the number-theoretic level, but only at the logical level, using the isomorphism. But we claim here that the question of consistency with (II.8.1) can be posed at the number-theoretic level, since (II.8.1) makes a number theoretic statement which assumes a certain TYPE. Based on the discussion presented, it is clear that the "restrictions of TYPE" between the statements are inconsistent. Indeed, they must be if the isomorphism is to hold since the inconsistency has been demonstrated for the logical level.

Even if the proof were reproduced purely in terms of number-theoretic statements, this can have no bearing on consistency or decidability of PM+Peano as numbers do not speak of such things. It is only through the isomorphism that an interpretation is possible. If the logical details of the proof are not faithful, then the interpretation is not reliable. The theorem must say something both as a number theoretic argument and as a logical argument at each step. The discussion that has been presented here feals only with the validity of the logical argument and, therefore, its interpretation.

This investigation is dependent upon applying the isomorphism to some of the intermediate steps of the proof under the image. Assuming that the isomorphism is faithful, then the intermediate steps must also admit of a meaningful interpretation under the image. If intermediate steps of the proof have no relevance to the "logical" interpretation - then the conclusion of the proof can have no implications concerning the structure of the image: we can have NO faith in the inverse map used in the arguments presented by Goedel at the conclusion of the proof (remarks 1 and 2, pages 608-609 [5]).

If one claims that the isomorphism does not apply to the abatements in the expansion presented — i.e., by claiming that "PREDICATED ON" has no bearing on their validity, then one must also give up the derivation of statements II.13... Furthermore, it would then be nonsensical to claim that the numeral Sb(Neg) is the same as Neg(5b) as Goedel claims. Nonetheless, it is the isomorphism which demands the validity of this portion of the derivation.

If one claims that the details of the proof can not be expanded and that it is improper to ask whether or not the proof scheme is valid - i.e., that lines of the proof such as II.S.l be accepted "springing full-grown from the head of Zeus" (alias Goedel), then this defeats the very notion of the proof even as defined by Goedel in I.44. We shall then have a nonsense proof about nonsense.

The choice of q' as either q or Neg q is clearly the difference between the reformulations. In particular, why is Goedel's result so clearly dependent upon choosing q' as q in one instance and q' as Neg q in the other? It must at least be shown how (or perhaps why) the alternate proof schemas fail to be the correct ones. Otherwise we are left with the problem of reconciling two valid proof schemas which yield inconsistent results. Such a conclusion would seem to imply that we have NO valid proof schemas nor means for judging what is and what is not valid. Surely, we ere not free to make contradictory assumptions in a single system as has been made the choice "Neg q" where Goedel chooses "q".

Either the proof consists of consistently interpretable number theoretic expressions or it does not. If it does, then the proof sequence itself is NOT a valid proof schema as it contains incompatible assumptions. If it does not, then the relevance of the proof must be reevaluated.

In short, we have demonstrated that the original proof of Goedel's Incompletness Theorem (1931) contains what appears to be a contradiction in its assumptions and that the choice of instantiation based on these assumptions is critical to the success of the proof. Note that we say not objecting here to the notion that some choice of instantiation for the free variables of Goedel's proof schema might yield the famous result. Rather, we seek to understand the validity of the particular choices, since we have demonstrated that these choices can be expressed as a contradiction by a simple change of notation.

5. ACKNOWLEDGEMENTS

Prof. H. Piette Noyee (Stanford University) provided encouragement and guidance, and insisted that the original thoughts be given serious attention. S. Peferman (Stanford University) provided references and pointed out typographical errors. C. Clarke provided an intelligent (and somewhat proverbial) brick wall for me to butt my head against. C. Gefwert, G. Chew, and others gave valuable advice and encouragement. Mary provided considerable patience and understanding.

REFERENCES

- 1. Barwise, J., etd., <u>Handbook of Mathematical Logic</u>, North-Holland Publishing Company, Amstedam, The Netherlands, 1977
- 2. Crossley, J.N., et.al., What is Mathematical Logic, Oxford University Press, Oxford, England, 1979
- 3. Davis, M., Computability and Unsolvability, McGraw-Hill Book Company, New York, 1958
- 4. Delong, H., A <u>Profile of Mathematical Logic</u>, Addison-Wesley Publishing Company, Reading, Massachusetts, 1970

- 5. Goedel, R., *On Formally Undecidable Propositions of <u>Principia Mathematics</u> and Related System T* in <u>From Frege to Godal</u> edited by J. van Reijenoort, Harvard University Press, Cambridge, Massachusetts, copyxight 1967, pages 592-616.
- 6. Heack, S., <u>Philosophy of Logics</u>, Cambridge University Press, Cambridge, England, 1980
- 7. Hintikks, J., The Philosophy of Mathematics, Oxford University Press, Oxford, England, 1969
- 8. Jeroslaw, R.G., *Redundancies in the Eilbert-Bernays derivability conditions for Goedel's second incompleteness theorem*, J. Symbolic Logic, 38, 359-367
- 9. Kleene, S.C., <u>Introduction to Meta-Mathematics</u>, North-Eolland Publishing Company, Amsterdam, The Netherlands, 1980
- 10. Rogers, H., Theory of Recursive Functions and Effective Computability, McGraw-Hill Book Company, New York, 1967
- 11. Tareki, A., Logic, Semantics, Meta-Mathematics, Esckett Publishing Company, Indianapolis, Indiana, 1983