The Philosophical Quarterly Vol. 33 No. 131

THE LOGICAL PARADOXES AND THE LAW OF EXCLUDED MIDDLE

BY GRAHAM PRIEST

I will argue that a *prima facie* plausible, uniform way of solving the logical paradoxes fails. The proposed solution starts from the observation that in many cases the proof of the contradiction which is the paradox, goes *via* the proof of an assertion of the form $\varphi \leftrightarrow \neg\varphi$.[1] Now, the principal agent in the move from $\varphi \leftrightarrow \neg\varphi$ to $\varphi \wedge \neg\varphi$ is the *reductio* principle: $\varphi \rightarrow \neg\varphi / \neg\varphi$. Without this the proof of the contradiction is stymied.[2] This suggests that a uniform way of avoiding the paradoxes is to reject the *reductio* principle. This move, perhaps not so plausible at first, gains strength from the following considerations. The *reductio* principle is, under very weak assumptions, equivalent to the law of excluded middle.[3] (Indeed it is the law of excluded middle which provides the classical *rationale* for *reductio*.) And reasons for doubting the law of excluded middle are two-a-penny. Moreover, many have suggested that paradoxical sentences are neither true nor false, or some variant thereof (are meaningless, express no statement, etc.).[4] This failure of the principle of bivalence does not necessarily entail a failure of the law of excluded middle. However the step is but a very small, and a plausible one. The proposed solution then is this: block the paradoxes by ditching the law of excluded middle on the ground of some kind of infelicity in paradoxical assertions.[5] Further support comes from an as yet unpublished result in which Ross Brady proves that naive set theory, based on a version of the system **T–W**, is consistent. However, naive set theory based on **T–W** plus the law of excluded middle is inconsistent. (Russell's paradox is easily representable.)[6] Thus,

[1] The observation is made in, for example, A. Fraenkel, Y. Bar Hillel and A. Levy, *Foundations of Set Theory* (Amsterdam, 1973):

> . . . in most, if not all, antinomies the crucial contradiction has, or can be given, the form of an equivalence between a certain statement and its negation. (p. 207)

[2] As Fraenkel, Bar Hillel and Levy go on, in effect, to observe.

[3] Specifically, given only first degree entailments, rule *reductio* and the law of excluded middle are equivalent.

[4] See, for example, the "solutions" cited in S. Haack, *Philosophy of Logics* (Cambridge, 1978), Ch. 8.

[5] Just this position is taken by, for example, F. B. Fitch in his *Symbolic Logic* (New York, 1952), especially §§2.14, 18.8.

[6] **T–W** is the system **T** of A. Anderson and N. Belnap, *Entailment* (Princeton, 1975), minus the axiom **W**: $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$ (which gives rise to trivialising Curry paradoxes in the context of

provided one is prepared to accept **T–W**, or something weaker, as the correct logic of entailment (for which there are good independent grounds[7]) then jettisoning the law of excluded middle provides a uniform and provably adequate way of avoiding the logical paradoxes.[8]

However, this solution will not work. The reason is this: although no set theoretic paradox may be provable without the law of excluded middle, the case is different with the semantic paradoxes. Although some of these, such as the heterological paradox, go *via* an assertion of the form $\varphi \leftrightarrow \neg\varphi$, and hence use the law of excluded middle, the definability paradoxes, such as Berry's, Richards' and König's do not.

Take for example Berry's paradox. English has a finite vocabulary. Hence there are a finite number of noun phrases with less than 100 letters. Consequently there can be only a finite number of natural numbers which are denoted by a noun phrase of this kind. Since there is an infinite number of natural numbers there must be numbers which are not so denoted. Hence there must be a least. Consider the least number not denoted by an English noun phrase with less than 100 letters. By definition this cannot be denoted by a noun phrase with less than 100 letters, but we have just so denoted it. Contradiction.

Prima facie this argument appeals nowhere to the law of the excluded middle. But appearances can be deceptive. The rest of this paper is a formalisation of this argument, which proves that the law of the excluded middle is unnecessary for the establishment of this contradiction.

The proof is carried out in the language of first order arithmetic[9] augmented by a description operator ε, a (genuine) implication operator $\rightarrow$, a two-place predicate $\triangle$, and a one-place function symbol l. The only other point to note is that the language in question has only a finite number of variables. Thus the number of terms of any finite length is finite (as in English).

The underlying logic used in the proof is any quantified relevant logic stronger than first degree entailment (considered as a relation holding be-

naive set theory). The system **R** (Anderson and Belnap, *op. cit.*) is slightly stronger than **T**. Naive set theory based on **R–W** plus the law of excluded middle is known to be not only inconsistent but trivial. Both the consistency and the triviality of naive set theory based on **R–W** are, as far as I know, open problems. (A theory is *consistent* iff it contains both A and $\neg A$ for no formula A. It is *trivial* iff it contains every formula. For theories based on classical logic, inconsistency and triviality coincide. Saner logics do not collapse these notions into each other.)

[7] See my "Sense, Entailment and *Modus Ponens*", *Journal of Philosophical Logic*, 9 (1980), 415–35.

[8] Fitch's system is also provably consistent. However his system is inadequate since its implication operator is highly irrelevant. The system and its consistency proof also have other problems. See, Fraenkel, Bar Hillel and Levy, *op. cit.*, pp. 205–7.

[9] See, for example, S. Kleene, *Introduction to Metamathematics* (Amsterdam, 1952), §16.

tween formulas of arbitrary degree), but not containing the law of excluded middle (for example, **T–W** will do). Descriptions are supposed to behave as for Hilbert, i.e., all terms in the language have some denotation, and

$$(1)\quad \exists x\varphi \rightarrow \varphi(x/\varepsilon x\varphi)$$

is the description axiom. (At the cost of complicating the proof somewhat, we could use definite descriptions.) Identity is governed by the usual axioms including

$$(2)\quad (\varphi \wedge t_1 = t_2) \rightarrow \varphi(t_1/t_2)$$

The proof also uses a few arithmetic principles. Specifically I shall invoke without mention any true arithmetic equation, together with $x+1 \neq 0$, $x+0=x$, commutativity and associativity for $+$ and

$$(3)\quad x+y=x+z \rightarrow y=z.$$

All of these are true in the intended interpretation of the language (the natural numbers). Moreover each of these can be proved from appropriate versions of the Peano axioms using little more than substitutivity of identicals and transitivity of $\rightarrow$.[10]

The other axioms we need at present concern $\triangle$ and l. We need to have names for the symbols of the language in the language itself. We could add these in some way. However, it is simpler to take advantage of the standard device of using Gödel numbers as names. So let us suppose that we have a Gödel coding $\sharp$ such that for any term of the language t, $\sharp$ t is a number. We make no assumption about $\sharp$ other than that there is a two-way effective procedure between formula and code. If n is a number let $\underline{n}$ be its numeral. Neither $\sharp$ nor $\underline{\ \ }$ are part of the language of the proof, but for any term t of the language $\underline{\sharp t}$ is a numeral, and hence part of the language of the proof. The axiom scheme for $\triangle$ is:

$$(4)\quad \triangle\, \underline{\sharp\, t}\, x \leftrightarrow x = t$$

for all terms t. Thus $\triangle$ is the denotation predicate. In effect, it says that 't' denotes t and only t. The axiom scheme for l is

[10] The only point worth a mention is that induction has to be taken in rule form thus:

$$\begin{array}{c} \varphi(x/y) \\ \cdot \\ \cdot \\ \cdot \\ \varphi(x/0)\ \ \varphi(x/y') \\ \hline \forall x\varphi \end{array}$$

The proofs of all these things are then fairly standard. Those of Kleene, *op. cit.*, Theorem 25, p. 186, will do.

(5) $l\underline{\#\text{t}}=\underline{\text{n}}$

where the number of symbols in t is n. Since decoding is an effective procedure this set of axioms is recursive. Moreover we may suppose that l is the primitive recursive function whose value with argument y is the number of symbols in the term whose code is y, if there is one, and zero otherwise. Thus $ly\neq\underline{0}$ means, in effect, 'y is the code number of some term.'

Now let φ be the formula[11]

$$\forall y\neg(ly\neq\underline{0}\wedge\Delta yx\wedge ly<\underline{10}\times\underline{10})$$

(i.e. x is not denoted by a term with less than 100 symbols) and suppose that we can prove that

(α) $\exists x\varphi$

A contradiction is easily forthcoming. For by (1) and (α),

$$\forall y\neg(ly\neq\underline{0}\wedge\Delta y\varepsilon x\varphi\wedge ly<\underline{10}\times\underline{10})$$

and hence,

(β) $\neg(l\underline{\#\varepsilon x\varphi}\neq\underline{0}\wedge\Delta\ \underline{\#\varepsilon x\varphi}\varepsilon x\varphi\wedge l\ \underline{\#\ \varepsilon x\varphi}<\underline{10}\times\underline{10}$

But by (4),

$$\Delta\ \underline{\#\ \varepsilon x\varphi}\varepsilon x\varphi.$$

Moreover $\varepsilon x\varphi$ has 52 symbols. Hence by (5),

$$l\ \underline{\#\ \varepsilon x\varphi}=\underline{52}$$

Thus,

$$l\ \underline{\#\ \varepsilon x\varphi}\neq\underline{0}\wedge l\ \underline{\#\ \varepsilon x\varphi}<\underline{10}\times\underline{10}$$

whence

$$l\ \underline{\#\ \varepsilon x\varphi}\neq\underline{0}\wedge\Delta\ \underline{\#\ \varepsilon x\varphi}\varepsilon x\varphi\wedge l\ \underline{\#\ \varepsilon x\varphi}<\underline{10}\times\underline{10}$$

contradicting (β).

Now (α) is certainly true in the intended interpretation of the language. However, it might be thought that the truth of (α) depends in some way on the law of excluded middle. So it is worth showing that (α) is provable from something entirely unproblematical, namely the following. Let n be any number. Since there are only a finite number of variables in the language there are only a finite number of terms of the language with less than n symbols. Let these be t_i, $1\leqslant i\leqslant k_n$. We take as an axiom

[11] $x<y$ is defined in the usual way as $\exists z(x+z'=y)$.

$$(\gamma)\quad ly<\underline{n}\wedge ly\neq\underline{0}\leftrightarrow \bigvee_{1\leq i\leq k_n} y=\underline{\sharp\ t_i}$$

for all n. (γ) is easily seen to be a true primitive recursive identity and is our final axiom.

The proof of (α) from (γ) proceeds as follows. Let $t_1 \ldots t_k$ be all the terms with fewer than 100 symbols. Let t be $t_1 + \ldots + t_k + \underline{1}$. Then for any $1\leq i\leq k$,

$$t_i=t\rightarrow t_i=t_i+t_1+\ldots+t_{i-1}+t_{i+1}+\ldots+t_k+\underline{1}$$

$$\rightarrow t_i+\underline{0}=t_i+t_1+\ldots+t_{i-1}+t_{i+1}+\ldots+t_k+\underline{1}$$

$$\rightarrow \underline{0}=t_1\ldots+t_{i-1}+t_{i+1}\ldots+t_k+\underline{1}$$

by (3). Hence, by transitivity and contraposition

$$t_i\neq t$$

But (4) gives

$$\Delta\ \underline{\sharp\ t_i}\,t\leftrightarrow t=t_i$$

Whence

$$\neg\Delta\ \underline{\sharp\ t_i}\,t$$

Thus

$$(\delta)\quad \bigwedge_{l\leq i\leq k} \neg\Delta\ \underline{\sharp\ t_i}\,t$$

Now

$$ly\neq\underline{0}\wedge\Delta yt\wedge ly<\underline{10}\times\underline{10}\rightarrow\Delta yt\wedge \bigvee_{1\leq i\leq k} y=\underline{\sharp\ t_i}$$

by (γ). But

$$\Delta yt\wedge \bigvee_{1\leq i\leq k} y=\underline{\sharp\ t_i}\rightarrow \bigvee_{1\leq i\leq k}(\Delta yt\wedge y=\underline{\sharp\ t_i})$$

by distribution, and

$$\bigvee_{1\leq i\leq k}(\Delta yt\wedge y=\underline{\sharp\ t_i})\rightarrow \bigvee_{1\leq i\leq k}\Delta\ \underline{\sharp\ t_i}\,t$$

by (2). Thus by transitivity and contraposition,

$$\neg\bigvee_{1\leq i\leq k}\Delta\ \underline{\sharp\ t_i}\ t\rightarrow\neg(ly\neq\underline{0}\wedge\Delta yt\wedge ly<\underline{10}\times\underline{10})$$

Hence, by (δ) and de Morgan's Laws,

$$\neg(ly\neq\underline{0}\wedge\Delta yt\wedge ly<\underline{10}\times\underline{10})$$

So

$$\forall y \neg (ly \neq \underline{0} \wedge \Delta yt \wedge ly < \underline{10} \times \underline{10})$$

which implies (α).

Thus a contradiction is provable, without the use of the law of excluded middle, from things which appear uncontroversially true. This shows two things. First, in combination with Brady's result, it shows that even if standard number theory can be developed in naive set theory based on **T–W**, this theory cannot define its own denotation (and hence satisfaction) relation.[12] Secondly, it shows that jettisoning the law of excluded middle is not sufficient to avoid all the logical paradoxes.

University of Western Australia

[12] The same goes for Fitch's system.