2004

On diagonal argument, Russell absurdities and an uncountable notion of lingua characterica

King, James Douglass

Lethbridge, Alta. : University of Lethbridge, Faculty of Arts and Science, 2004

ON DIAGONAL ARGUMENT, RUSSELL ABSURDITIES
AND AN UNCOUNTABLE NOTION OF *LINGUA CHARACTERICA*

JAMES DOUGLASS KING
Bachelor of Arts, University of Lethbridge, 1996

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the Requirements for the Degree

MASTER OF ARTS

Department of Philosophy
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

ABSTRACT

There is an interesting connection between cardinality of language and the distinction of *lingua characterica* from *calculus rationator*. *Calculus*-type languages have only a countable number of sentences, and only a single semantic valuation per sentence. By contrast, some of the sentences of a *lingua* have available an uncountable number of semantic valuations. Thus, the *lingua*-type of language appears to have a greater degree of semantic universality than that of a *calculus*. It is suggested that the present notion of *lingua* provides a platform for a theory of ambiguity, whereby single sentences may have multiply – indeed, *uncountably* – many semantic valuations. It is further suggested that this might lead to a pacification of paradox. This thesis involves Peter Aczel's notion of a universal syntax, Russell's question, Keith Simmons' theory of diagonal argument, Curry's paradox, and a 'Leibnizian' notion of language.

ACKNOWLEDGEMENTS

taking the sting out of administrative procedures which, at other institutions, I usually find arcane. The departments of Philosophy and of Mathematics and Computer Science have proved themselves friendly and fertile communities of academic exchange. In general, and speaking as one who has ample experience at other universities, the academic community of the University of Lethbridge is a place of high excellence and collegiality, where all of my best flourishing has happened.

TABLE OF CONTENTS

SYMBOLS

Throughout this thesis, all references to a formal system of logic will mean an ω-consistent extension of the Peano Arithmetic. The following symbolic conventions are also used throughout this thesis.

L A first-order expressively strong system of formal logic.

FORM   The set of the well-formed formulas of L.

NN   The set of the natural numbers.

$\lceil ... \rceil$ Gödel quotation – A total function from FORM 1-1 into NN.

$\lfloor ... \rfloor$ Gödel disquotation – A partial function from NN 1-1 onto FORM. (This is the *inverse* of the Gödel-quote function.)

C A Curry sentence. ($\lceil C \rceil = \lceil C \to Q \rceil$, for some arbitrarily chosen formula, Q.)

$C(n)$ Another Curry sentence. ($n = \lceil C(n) \rceil = \lceil C(n) \to Q \rceil = \lceil \lfloor n \rfloor \to Q \rceil$)

$\vdash$ 'Turnstyle' – derivability/theoremhood predicate.

$\nvdash$ Negated 'turnstyle'.

SOME COMMENTS ON UNIVERSAL SYNTAX

In the first two paragraphs of his paper, 'Schematic consequence', Peter Aczel

makes some interesting comments. (For 'AC', read '*Aczel's comments*'.)

> (AC) Nowadays we are well aware that there are many different logics.
> There are computer systems which are meant to be used to
> implement many logics. But there is no generally accepted
> account of what a logic is. Perhaps this is as it should be. We
> need imprecision in our vocabulary to mirror the flexible
> imprecision of our thinking. There are a number of related
> phrases that seem to have a similar imprecision e.g. formal
> system, language, axiom system, theory, deductive system, logical
> system, etc.... These are sometimes given technical meanings,
> often without adequate consideration of the informal notions.
>
> When a logic has been implemented on a computer system
> the logic has been represented in the logical framework that the
> computer system uses. The logical framework will involve a
> particular approach to syntax, which may differ from the approach
> to syntax taken when the logic was first presented. This means
> that in order to represent the logic, as first presented, in the
> framework a certain amount of coding may be needed and *the*
> *question will arise whether the logic as first presented is indeed*
> *the logic that has been represented in the framework. To make*
> *this question precise it is necessary to have a notion of a logic*
> *that abstracts away from particular approaches to syntactic*
> *presentation. [emphasis added]* [ACZ5, p.261]

Owing to practical constraints, I will discuss neither of Aczel's general questions, 'what

is the notion of a logic?' (which is implied in the first paragraph above), and 'how can we

determine whether or not the logic-in-the-framework correctly represents the logic-as-

first-presented?' Except to observe that schematic consequence offers an interesting

notion of consequence *via* abstracted lexical meanings, we will not be concerned with

Aczel's presentation of *schematic consequence*. It would nevertheless be interesting to

know what *schematic consequence* has to say about the notion I develop below, of

Leibnizian language. I am also interested in the notion of a universal syntax – though I

believe I contribute more to the problem than to the solution. My suspicion is that

1

Leibnizian language could not be characterized within a universal syntax of the sort Aczel suggests via schematic consequence. But I cannot presently confirm or disconfirm this suspicion. If it is true, then Leibnizian language is a limit to which Aczel's schematic syntax cannot reach. In turn, schematic syntax is disqualified as a candidate for a truly *universal* syntax. But this disqualification would only be technical, since I have only a partial notion of what a Leibnizian language would be. In any case, the notion of Leibnizian language is closer to natural language than to formal language, since as will be discussed below [at pp. 48 & *ff.*], a Leibnizian language must be a *lingua characterica*, as are all natural languages, whereas the language of schematic consequence must be a *calculus rationator*, as are all formal symbolic languages. So Aczel's schematic syntax is still universal for all of our usual work-a-day formal languages. (By 'work-a-day' formal languages, I mean programming languages, the mathematics of astrophysics, geophysics, mathematical linguistics, economics, and so on.)[1]

Rather than universal syntax, this essay concerns diagonal argument and its relationship to (Leibnizian) *lingua characterica*. In particular, I argue that 'Leibnizian' language cannot be modeled via the formal resources of 'syntactic' languages. (By 'syntactic' language, I mean classical formal language, including first-order logic, second-order logic, and in general, all classical languages which work via manipulation of symbols.) What lets Leibnizian language out of the universe of syntactic languages is that Leibnizian language has as many predicate extensions as there are subsets of the set of sentences (where a sentence is a finite string of symbols), or of the subsets of the

---

[1] Other possible theories of universal syntax might be found in *gaggle theory* (*see* [DUN1] and [DUN2]), and for another possibility, *see* J.R.Brown's 'Proofs and pictures' [BRO].

natural numbers. Thus there are not enough sentences in a syntactic language to pick out every subset of the naturals. Of course, since second-order logic quantifies over its predicates, of which there may be an uncountable infinity, it might capture Leibnizian language *via syntactic resources*. That is, Leibnizian language might turn out to be no more expressive than second order logic.

The validity of the diagonal method stands independently of its instances in *calculus rationator* of symbolic logic. Accordingly, diagonal argument imposes constraints on symbolic logic, rather than the reverse. Instead of being a kind of proof for mathematical logic alone, we shall see that diagonal argument has applications well beyond mathematical logic to the theory of natural language, the semantic paradoxes. As further confirmation, we can observe that within the general discipline of mathematical logic, diagonal argument provides philosophical-sounding conclusions for set theory, metamathematics and computability theory. For Fregean set theory, the principle of set comprehension fails, since the Russell set (RS) cannot be consistently either included or omitted from itself. That is, we can reason both that $RS \in RS$ and that $RS \notin RS$. In computability theory, it is demonstrable that no universal Turing machine computes all functions – that is, the halting problem is unsolvable. In metamathematics, Gödel's incompleteness theorems show that proof and truth are not the same things, and due to elaboration by Tarski [TAR2][2], the truth-predicate is not even definable in classical first-order logic. None of these philosophical-sounding conclusions are provable in a formal language, and neither is the language of schematic consequence. Indeed, many of these

---

[2] *see also*, [MEN, p.151] and [KLE1, p.501]. Kleene: 'Truth definitions for formal systems were originally investigated by Tarski (1932, 1933). He established that, if an (effective) formal system including the usual number theory is consistent, it must be impossible to express the [truth] predicate... For [if the truth predicate were expressable] then the reasoning of the Epimenides paradox could be carried out in the system.'

conclusions are not even expressible or definable in formal languages (in consequence, for example, the undefinablility of the truth predicate). Yet diagonal argument does bring us to these philosophical-sounding conclusions. Accordingly, diagonal argument is not an exclusively formal kind of argument, i.e., it is not exclusively a *calculus rationator* kind of argument. Thus, as is argued in the rest of this essay, diagonal argument is – in the general case – a method of proof for the *lingua characterica* kind of language. (As the *calculus rationator* kind of language is subsumed within the *lingua characterica* kind, the general case of diagonal argument includes the special cases of the *calculus rationator* languages.)

RUSSELL'S QUESTION

The general context of our investigation into diagonal argument will be that of

Russell's question (RQ).

(RQ)   Why are some diagonal arguments good, whereas others are bad?[3]

Much turns on what is meant by 'good' and 'bad' in (RQ), since there are many

interesting kinds of diagonal argument. For instance, Keith Simmons proposes that the

'good' diagonal arguments are reasoned upon 'well-determined' sets, whereas the 'bad'

are reasoned upon (false suppositions or) 'non-well-determined' sets [SIM, pp.27-37;

especially p.29]. But Simmons' distinction does not exhaustively classify all diagonal

argument. Of course, some of the bad diagonal arguments can be immediately

recognized as such, for they (purport to) prove contradictions. (Of course, if a

contradiction is actually proved – or seems to have been proved – then we must either

take the proof as an absurdity for a *reductio* argument, *or*, we must either change our

logical system or employ a paraconsistent system.) Since contradictions are easy to

identify by their syntactic form, we can effectively locate and dismiss them. Thusly at

least some of the 'bad' diagonal arguments are identified. But there are other diagonal

arguments, the conclusions of which are (or seem) utterly absurd, yet without having a

contradictory form. These are the 'Russell absurdities' (as I will call them). Every

contradiction is an absurdity, but not every absurdity is a contradiction. Some non-

contradictory absurdities are very hard to attribute to any error, except that (quasi-) error

of violating our intuitions about absurdity in the first place. But intuitions inevitably

change, and with time, a number of the past absurdities become re-evaluated as theorems.

---

[3] I have been unable to locate (RQ) in Russell's writings. But Keith Simmons refers to Russell's question –
'a question first asked by Russell' – in his Universality and the Liar [SIM, p.20].

5

(In the early history of mathematical logic, these are the theorems of Cantor's powerset, Russell's set, and Gödel's incompleteness proofs.) So if we want to decide whether or not a given diagonal argument is 'good', we might simply wait. What the teacher's teacher cannot conceive, the teacher may not be able to fault, and the student may embrace. This change in the minds of mathematical logicians is interesting in its own right, from both historical and philosophical perspectives. But it cannot qualify as a means of distinguishing the 'good' non-contradictory absurdities from the 'bad' ones, whereas that is precisely what (RQ) exhorts us to find. (Hereafter, we refer to the non-contradictory-but-absurd conclusions and their diagonal proofs as, 'RA', for *R*ussell *a*bsurdities.)

> (RA) A 'Russell absurdity' is an intuitively absurd thesis which is argued validly, and which does not have the form of a contradiction.

There is credibility in the claim that there are no hard-to-decide diagonal arguments left, leastways, none for first-order logic. So if there remains an interesting case of an (RA), it must ostensibly be an argument in some language other than those of first-order logic. But given the great semantic generality of first-order logic, there is not much left which cannot be expressed formally. Although there are significantly more expressive formal systems – such as those of second-order logic[4] – first order logic is

---

[4] An example of the greater expressive capacity of second-order logic is easily seen in the difference between first-order and second-order mathematical induction. In a first order system, induction is captured by a schema which applies only to the predicates given in the system, i.e., X is used as a substitutional variable. 'X' in the first-order induction schema,

$$\{ X(0) \wedge (\forall m)[X(m) \rightarrow X(m+1)] \} \rightarrow (\forall n)X(n),$$

'stands in' for the predicate-names available in the first-order language. (We understand m and n to be natural numbers.) 'X' does not range (i.e., *is not quantified*) over all of the *properties* of the natural numbers, but only the available predicate-*names*. By contrast, second-order induction involves quantification over all of the properties of natural numbers. The second-order axiom of induction can be expressed in this way:

already expressive enough to represent much of natural language (*see* [JAC, *chapters* 2.1,

2.2, 5 & 6] *and* [JOH]). Thus natural language is the most semantically universal, first-

order logic is nearly as universal as English. So first-order logic has a *very* high degree

of universality, and first-order logic has enjoyed some impressive (though partial)

successes as a theory of natural language. Noam Chomsky, for instance, founded the still

flourishing field of mathematical linguistics (*see*, for example, [CHO1] and [CHO2]),

whereby natural language is represented in a formal (usually first-order) language.

Simplified, this is the notion that natural language is really just a special case of formal

language – that natural language *reduces to* or *is completely expressable by* formal

language.[5]

It is a key thesis of this essay that some diagonal argument shows that there is a

somewhat coherent notion of natural language which does not completely reduce to an

equivalent expression in first-order logic, and that (when clarified) this notion may afford

a greater degree of semantic universality than can be achieved via first-order systems.

Those (RA)s which are not decidable in first-order logic might be decidable in a non-first

order language – (possibly) such as second-order logic – of higher semantic generality.

Although this notion of language (which I call 'Leibnizian') is incomplete, it suggests an

---

$(\forall X)[\ \{ \ X(0) \wedge (\forall m)[X(m) \rightarrow X(m+1)] \ \} \rightarrow (\forall n)X(n) \ ]$.

Thus the first-order axiom of induction implies an infinity of axioms, *viz.*:

$\{ \ P_0(0) \wedge (\forall m)[P_0(m) \rightarrow P_0(m+1)] \ \} \rightarrow (\forall n)P_0(n),$
$\{ \ P_1(0) \wedge (\forall m)[P_1(m) \rightarrow P_1(m+1)] \ \} \rightarrow (\forall n)P_1(n),$

and so on, such that every predicate-name is some P_i, for $i \in$ NN. Necessarily, there are extensions of natural numbers (i.e., subsets of NN) which none of the P_i name. (This follows trivially from the fact that there are uncountably many subsets of NN, but only denumerably many predicate-names.) The second-order axiom entitles us to assert induction via every subset of NN, whether or not the subset (i.e., *property extension*) has yet been assigned a name.

[5] Below, at pp.46-54, we revisit this notion as the *mathesis universalis* view of (natural) language.

uncomplicated way of deciding (at least some) of the Russell absurdities. These decisions are obtained by simply re-expressing the (RA)s in terms of Leibnizian language. For example, consider a famous (RA), that of the Russell set (hereafter, 'RS'). (This (RA) is no longer understood as an absurdity, though it was at the time of its discovery.) Ordinarily, the absurdity of the RS is understood as a proof that (RS $\in$ RS) $\leftrightarrow$ (RS $\notin$ RS), which is a contradiction, and in this form, it does not qualify as an (RA) – since no (RA) is a contradiction. However, the RS can also be understood as showing that because of the contradiction of the RS, the principle of set comprehension fails, or else Frege's set theory is inconsistent. According to this principle, every intuitively coherent definition of a set – such as 'RS is the set of all non-self-membered sets' – consistently defines the set's membership. (In the case of RS, it is often said that the definition of RS is not coherent, since, according to Russell, the definition of RS violates the famous 'vicious circle principle'.) But the RS does lead to the following (RA) – or at least it did for Frege, who was disinclined to accept the vicious circle principle as a solution for the problem of the RS.)

> (FRA) Either set comprehension is incorrect or else Frege's set theory is inconsistent.

(For 'FRA', read '*F*rege's *R*ussell *a*bsurdity'.) (FRA) is solved via Leibnizian language by allowing the membership of RS to be ambiguous, in the sense that RS has many possible memberships. We distinguish these memberships by adding primes to the symbol 'RS'. Thus, there is one membership, RS', for which RS $\in$ RS'. In another membership, RS'', we have RS $\notin$ RS''. Thus, for Leibnizian language, the contradiction of RS resolves to,

8

$$(RS \in RS') \leftrightarrow (RS \notin RS''),$$

which does not have the form of a contradiction. In this Leibnizian interpretation of RS, there is no contradiction to force us to consider and resolve (FRA). Thus, Leibnizian language 'solves' (FRA).

Of course, to 'solve' a logical problem by translating it into another logical language has the *prima facie* appearance of an *ad hoc* form of argument. But it is not always an unsound mode of argument. Consider for example, Gödel's second incompleteness theorem of the unprovability of consistency. Strictly speaking, the second incompleteness theorem is not a proof that consistency is unprovable, but rather that *either* the Principia Mathematica is inconsistent, *or* the Principia cannot prove its own consistency. In *this* form, Gödel's second incompleteness theorem implies no decision as to whether or not consistency *actually* is provable in the Principia. Thus, to conclude that consistency is unprovable, we must suppose that the Principia is consistent. Since the Principia is consistent if and only if consistency is unprovable (in the Principia), the supposition that the Principia is consistent is logically equivalent to the desired *conclusion*, that consistency is unprovable. And so it is *ad hoc* to reason in this way that consistency *actually is* unprovable in the Principia. But this conclusion is not without support, for there is independent cause to hold that the Principia is consistent. It is an empirical fact that, as a formal theory of the natural arithmetic, the soundness of the Principia has overwhelming confirmation and no significant disconfirmation. The Principia's success as a theory of the natural arithmetic is at least as well confirmed as the most successful theories of physics – Newtonian, relativity and quantum. Were the Principia actually inconsistent, then we might have found some inconsistencies by now.

9

Another independent reason to hold that the first-order subsystem of the Principia is consistent is found in Gentzen's proof via transfinite induction [GEN]. This is a rigorous proof of the Principia's consistency, but it is not proved in the language of the Principia, nor in any first order system of logic. Gödel's second incompleteness theorem concerns only the provability of consistency of the Principia, in the language of the Principia. That is, Gödel shows only that (if it is consistent) then the Principia cannot prove its *own* consistency. Gödel's theorem does not block a consistency proof which is argued within a language other that the Principia. Gödel's second incompleteness theorem shows that there is no purely *logical* reason to hold that consistency is *actually* unprovable via proof in the Principia, but with non-logical support for the consistency of the Principia – via empirical confirmation and transfinite induction – it is possible to reasonably conclude that the Principia actually does not prove its own consistency. So our reasons for accepting the unprovability of consistency are just as much dependent on empirical confirmation and transfinite induction as they are on Gödel's theorem.

The Leibnizian view of language is similarly dependent on a strictly non-logical fact: Diagonal argument does not always have to be evaluated as if it were a first-order kind of formal argument. In particular, Keith Simmons' theory of diagonal argument does not cover all possible kinds – it is possible to conceive diagonal argument as occurring within a language of higher semantic generality than Simmons' theory affords. But the Leibnizian notion of natural language which will be developed below offers no refutation of any other notion (or model) of natural language, and particularly not those notions offered by first-order logic. Thus the Leibnizian notion of natural language might *compete* with the formal notion, but neither can refute the other. Nevertheless, the

Leibnizian notion is presently far too weak and incomplete to offer any *effective* means of modelling natural language. It is also hard to see how to strengthen the Leibnizian notion. Notwithstanding its weaknesses, this thesis is not a straightforward one, and besides the technical discussion, will include material from the history and philosophy of diagonal argument. Indeed, the vast bulk of this essay is devoted more to 'setting the scene' than to 'proving the point'. I would have preferred to 'prove' than to 'set', but at this stage of research, there is greater need of setting than proving.

The next part of this essay is devoted to a discussion of some of the history and philosophy of diagonal argument. The part immediately following explains Keith Simmons' theory of diagonal argument [SIM, pp.20-37], the comments of some other logicians on diagonal argument, and then a short philosophical critique of Simmons' theory. The motivating question of these parts was first expressed by Russell, where he asks (RQ).

SOME HISTORY AND PHILOSOPHY OF DIAGONAL ARGUMENT AND RUSSELL'S QUESTION

Diagonal argument was first presented by Georg Cantor in [CAN1] (*see also* [KLE2, pp.180-183]), but is nevertheless suggested in earlier logical puzzles, all the way back to Epimenides' liar paradox. This pre-Socratic self-referential paradox can be profitably explained (if not solved) by assuming that they are diagonally argued. Pre-Socratic philosophers debated these paradoxes thoroughly, but of course, their understanding of 'self-referential' paradox is distant from the powerful analysis offered by the voluminous contemporary philosophical and formal literature. This literature, exemplified by work such as Douglas Hofstadter's philosophical Gödel Escher Bach [HOF] and Raymond Smullyan's formal Diagonalization and Self-Reference [SMU2], is accordingly the more appropriate language for the following analysis of diagonal argument. However, we will briefly consider some of the so-called 'semantic' paradoxes, and express them as if they were diagonally proved. Inasmuch as Epimenides' liar has the same form as the 'semantic' liar, our analysis of the 'semantic' liar serves also as an analysis of Epimenides' liar.

Let us consider a diagonal analysis of Epimenides' liar. For reference, the non-diagonal expression of Epimenides' liar is as follows. (For 'EL', read 'Epimenides' liar.)

(EL$_0$)   All Cretans are liars.

Epimenides himself was a Cretan, and thus, by asserting (EL$_0$), Epimenides calls himself a liar. However, (EL$_0$) does not immediately imply a contradiction, since if any Cretan were to tell the truth just once, then (EL$_0$) is simply false (i.e., without also implying that (EL$_0$) is true). But (EL$_0$) can be strengthened to,

12

(EL₁)   Epimenides is a liar,

since Epimenides is a Cretan. And (EL₁) can be strengthened to,

(EL₂)   Everything Epimenides says is a lie,

since otherwise, if Epimenides sometimes tells the truth, then Epimenides' liar is not

paradoxical. Now as (EL₂) is something which Epimenides says, we have,

(EL₃)   The assertion of (EL₃) is a lie.

Finally, we employ the supposition that if a person's claim is a lie, then the claim is false.

(This supposition is not in all cases true, since it is credible to define a lie as the assertion

of a claim which the speaker believes is false – even though the claim might actually be

true. For example, suppose Smed comes indoors from a sunny sky and attempts to lie by

saying 'its raining'. But after Smed has come inside, and before he attempts to lie, the

weather may change to a thunderstorm.   Then Smed's lie is actually true.)   This

supposition that a lie is a falsehood leads us to the following formulation of Epimenides'

liar.

(EL₄)   (EL₄) is false.

One more change to Epimenides' liar brings us to the ordinary semantic paradox of the

liar sentence, (EL₅).

(EL₅)   This sentence is false.

(EL₅) is abstracted a significant distance from *Epimenides'* liar (i.e., from EL₀),

but it is not abstracted *too* far, for (EL₅) has exactly the logical structure which makes

Epimenides' liar a paradox. Notwithstanding the close relationship between (EL₀) and

13

(EL_5), we reason below according to (EL_4), since this is the expression most amenable to treatment as a *diagonal* paradox. For reference below, we allow (EL_4) to be expressed as s^*.

s^* s^* is false.

(This expression of s^* is not intended as a definition, but rather as a preview of the definition to follow below. s^* will be defined as a *fixed point* on the falsity predicate, p_F.)

We now consider some formal definitions and theorems which are needed for a diagonal expression of the paradox of s^*. First, we let the set SENT have all and only the declarative sentences of English as members. Note that there is exactly a countable infinity of these sentences, since every sentence is a finite string of the letters of the alphabet (including spaces, punctuation and brackets). (Proof is omitted.) So SENT is countable. Because SENT is countable, its members can be arranged into a list, as follows.

$s_0, s_1, s_2,$ and so on.

The sentences – i.e., the members of SENT – are referred to collectively as 'the s_i'. Note that as every sentence of English is some s_m, for a natural, m, it follows that there is some n such that $s^* = s_n$. (That is, the fixed point on p_F must be a sentence, and therefore it must be one of the s_i.)

We define the set PRED of the *sentential* predicates of English. By a 'sentential' predicate, I mean one which applies to the sentences of English. For example, the members of PRED include, '... has five words', '... is self-referential', '... is a palindrome', '... is false', and so on. Some predicates which are not members of PRED

14

include, '... is blue', '... is east of Lethbridge', and the like. (We could have allowed PRED to include *all* of the predicates of English without weakening the following arguments, but it increases clarity to exclude them from PRED at the outset.) We conceive the members of PRED extensionally, such that a predicate is a subset of SENT. Thus PRED $\subseteq$ $\wp$(SENT). Although the s_i are exhaustively listed above, not all the predicates can be listed, for there are more subsets of SENT than there are natural numbers. (This fact follows immediately by Cantor's powerset theorem, which I do not prove here.) Because $\wp$(SENT) is uncountably large, it cannot be hoped that its members could be exhaustively listed, as the s_i can be. There are not enough symbols of the form p_n (for n, some natural number) to name every member of $\wp$(SENT). But we can name *some* of the members of $\wp$(SENT), and that some are left un-named does not block or invalidate the diagonal analysis below. We let (some of) the members of $\wp$(SENT) to be listed as below.

p_0, p_1, p_2, and so on.

The predicates of this list are referred to collectively as 'the p_i'.

We assume for *reductio* that every one of the p_i can be uniquely paired with a single extension (that is, with a single unique subset of SENT). According to this assumption, there must be exactly as many predicate-extensions as there are p_i – to wit, exactly denumerably many. It will be shown that the *reductio* assumption is false, that even though the p_i are countable, they cannot each be uniquely paired with one extension. In particular, it will be shown that the predicate, '... is false' (p_F), cannot be paired with one extension unless it is paired with another, and then with an uncountable infinity

15

more. From this, it follows that either p_F has no extension, or that it has uncountably many.

It is important to stress that the *reductio* proof is not employed to prove that PRED is uncountable, even though we already know that PRED *is* uncountable. Rather, what the *reductio* argument is intended to prove is that the number of extensions which are named by the denumerably-many p_i is greater than the number of p_i in the first place. In other words, there is no one-to-one function from the p_i to the extensions named thereby. This logical manouvre – of stopping short of an otherwise well-warranted conclusion (that the extensions of the p_i are uncountable) – is not novel, having precedent in the reasoning of intutionist logicians. Intuitionists will grant that Cantor's diagonal proof of the uncountability of the real numbers proves that there is no one-to-one *onto* function from the naturals to the reals, but they will not grant the further conclusion that there are more reals than naturals. In contradistinction from the intuitionists, this essay does not go to challenge the conclusion that sets such as those of the reals, and of PRED, are uncountable. Rather, the uncountability of these sets is beside the point. What *is* material to this essay is that if predicates such as p_F have *one* extension, then *the set of the extensions of the denumerably-many p_i is uncountable*. Let us henceforth name the set of the extensions of the p_i as 'EPI', for '<u>e</u>xtensions of the <u>p</u>$_i$'.

In continuance of our diagonal analysis of the paradox of s^*, we now assert a fixed point theorem, via which s^* is defined.

> (FPT1) For every predicate p, there is a sentence s such that s = 's is p'. Equivalently, for every p, there is an s such that '$s \leftrightarrow p(s)$' is provable.
>
> (FPT2) Formally, the fixed point theorem is '$\forall x \exists y(s_y \leftrightarrow p_x(s_y))$'.

16

If we now instantiate p_x as p_F, we obtain,

(FPT3) $\exists y(s_y \leftrightarrow p_F(s_y))$.

And now we instantiate s_y as s_a (for a, some natural number), and define s^* as s_a. I stress that s_y is not (existentially) instantiated as s^*; nor is it *assumed* that s_a is s^*. Rather s^* is *defined* to be s_a, for whichever sentence s_a happens to be. Now the fixed point theorem on p_F is,

(FPT4) $s^* \leftrightarrow p_F(s^*)$.

(I do not attempt to prove the fixed point theorem, formally – although (FPT4) could be formally proved.) Notwithstanding the fact that (FPT4) is expressed in a formal way, we understand it as a natural language expression: 's* holds true if and only if s* is false'. In natural language, sentences can be directly presented without any need of formal definitions. The resources of natural language allow us to construct many sentences like s*, where self-attribution (or self-reference) can be easily accomplished – unlike formal self-attribution, which normally requires complex devices of one kind or another, such as Gödel-numbering.

Now with all the above preparatory discussion completed, we construct the following array, '(EDA)' (for *Epimenides diagonal array*), upon which the following reasoning is based.

	s_0	s_1	s_2	s_3	. . .	s^*	. . .
p_0	0	1	0	1			
p_1	1	1	0	1			
p_2	1	1	1	0			
p_3	0	0	1	0			
.							
.							

(EDA)

$p_\mathrm{F} \mid$ 1   1   0   0 ... α ...

. $\mid$

. $\mid$

We interpret (EDA) as follows. Across the top, we list the s_i. Down the side, we list the p_i (which constitute a subset of PRED). In the body of (EDA), there is a two-dimensional array of the symbols '0' and '1'. The rows of the array specify the membership of the p_i. Where there is a '1', the co-ordinate member of the s_i is a member of the co-ordinate member of the p_i. For example, suppose that the predicate name p_a has the row, 101011101111... Then the extension of p_a has s_0, s_2, s_4, s_5, s_6, s_8, s_9, s_{10}, and s_{11} as members. (p_a may or may not have as members some of the sentences after s_{11}.) Above, at the intersection of p_3 and s_2, there is a '1' – we interpret this to mean, 's_2 is a p_3', or '$p_3(s_2)$ is true'. Were there a '0' at this intersection, it would be interpreted to mean 's_2 is not a p_3', or '$p_3(s_2)$ is false'. Also, we understand the values of the array classically, such that every intersection has either '0' or '1', not neither and not both. If we reason concerning an intersection that it is 'not 1', then it is '0', and vice-versa. And furthermore, we treat truth classically, such that 'it is false that ϕ is false' means 'ϕ is true', and vice-versa.

We now proceed to the paradox. At the intersection of p_T and s^*, there is the symbol, 'α'. According to the interpretation of (EDA), α must be either '0' or '1'. Suppose α is '0'. Then $p_\mathrm{F}(s^*)$ is false. By (FPT4), it follows that s^* is false. But $p_\mathrm{F}(s^*)$ means 's^* is false'. So as $p_\mathrm{F}(s^*)$ is false, it follows that it is false that s^* is false. To wit, s^* is true – but this contradicts the hypothesis that α is 0. Suppose instead that α is '1'. Then $p_\mathrm{F}(s^*)$ is true. By (FPT4), it follows that s^* is true. But as $p_\mathrm{F}(s^*)$ means 's^* is

18

false', it follows that s^* is false. Accordingly, α should be '0', but this contradicts our hypothesis that α is '1'.

Thus, $\alpha = 0$ if and only if $\alpha = 1$. Such is the paradox. From this point, there are several avenues of sound reasoning. In intutionist logic, our assumptions of classical truth are weakened so as to block the derivation of contradictions. Alternatively, it might be argued that '$p_F(s^*)$' is undecidable, and thereby the intersection of p_F and s^* is not required to have either value, '0' or '1'. It might also be argued that p_F is indefinable as a subset of EPI, thus excluding p_F from (both) the p_i (and PRED). Similarly, s^* could be identified as an 'impredicative' (viciously self-referential) sentence, and for this reason s^* may be excluded from the s_i. All of these treatments are motivated by fact of the consistency of the formal first-order arithmetic, for no inconsistent theory can provide a sound model of arithmetical truth. (However, a paraconsistent system *can* provide a model of arithmetical truth which is – or might be – negation-inconsistent, but which is not absolutely inconsistent.) My treatment of (EDA) does not have this virtue – that is, I do not propose a means of preserving consistency. Rather than present a novel means of preserving consistency with respect to the values ('0' or '1'), I mean the analysis of (EDA) to indicate a novel *diagnosis* of the paradox. That is, I hope to say something novel about the *cause* of (natural language) diagonal paradox. It is not even my intention to say anything about the cause of *formal* first-order diagonal paradox.

To diagnose a paradox, it must exist, and it must be paradoxical – otherwise, either there exists no paradox to diagnose or the paradox is unproblematic. For the sake of argument, we accordingly grant soundness to the argument which concludes with the contradiction that $\alpha = 0$ if and only if $\alpha = 1$. It is unclear how this contradiction is to be

represented on the array of (EDA), but let us first consider two ways of representing the value(s) of α which are not options as diagnoses of the present paradox. *First*, we cannot represent α as being simultaneously '0' and '1'. Otherwise, we must say that s^* both is and is not a member of the (single) extension of p_F, but then the extension of p_F is not a set (because its membership is ill-defined), and *a fortiori*, this extension is not a well-defined subset of SENT. But if p_F does not name a subset of SENT, then p_F is no predicate of natural language (leastways, not if the natural-language sentential predicates are conceived extensionally). I resist the conclusion that natural language has no extensional falsity predicate, since the speakers of (virtually) all of humanity's natural languages quite ordinarily use a falsity predicate, and moreover, that falsity predicates have extensions, i.e., that natural language falsity extends to some set or other of false sentences. (Nevertheless, the natural language falsity predicate may not be *universal* – in the sense that this predicate may not have the same meaning for all uses and languages. Yet, natural language does have at least one falsity predicate which is stable for most applications within natural language.) So if we were to diagnose the paradox as a consequence of natural language's *actually* having no extension for the (ordinary[6]) falsity predicate, then we have failed to address the problem itself of *ordinary natural-language falsity* (where natural language is conceived extensionally). *Second*, I resist representation of α as "neither '0' nor '1' ". Were we nevertheless to do so, then the membership of the extension of p_F is again ill-defined. We would have to grant that s^* both is not in the extension of p_F and that it is not the case that s^* is not in the extension

---

[6] Of course, quasi-ordinary falsity predicates *are* possible in various formal languages. But even to assert that the natural-language falsity predicates are *actually* quasi-ordinary falsity predicates, it must be presumed – for it certainly could not be *proved* – that the ordinary natural-language falsity predicates are *completely* modeled via formal notions of quasi-ordinary falsity. If this assertion is nevertheless granted, then we have failed to diagnose the problem *itself* of ordinary natural-language falsity.

of p_F (i.e., that s^* *is* in the extension of p_F). And it follows again that the extension of p_F is not a well-defined set, and not a subset of SENT. And once again, the *ordinary* natural language notion of falsity has no well-defined extension. And thus we again will fail to address the problem *itself* of extensionally-conceived natural language falsity.

In both of the above two (non-) options of diagnosing the paradox of ordinary natural language falsity, it is (or could be) argued that the contradiction of natural falsity entitles argument by *reductio ad absurdum* that natural falsity is incoherent in the first place, and thus that it is no disservice to natural language to replace natural falsity by formal quasi-falsity, for anything is better than attributing truth to a contradiction. But, as is well-known, argument by *reductio* can go to negate *any* assumption upon which the derivation of the contradiction (i.e., the *absurdity*) depends. In the reasoning above of the ('first' and 'second') options for treatment of the falsity paradox, the *reductio* is employed to negate the assumption that natural language falsity predicates have well-defined sets of (false) sentences as extensions. But there are other assumptions which the *reductio* can negate. We could soundly argue a *reductio* against the supposition that PRED is denumerable – though this is not our purpose. (In any case, a *reductio* such as this does nothing to solve the problem of how to determine extensions for the natural language falsity predicates. This is because natural fixed points, such as s^* or the liar sentence, will remain even if (even *though*) PRED is not denumerable. Another supposition for possible negation by the *reductio* is that the sentential predicates can all be coherently assigned a well-defined extension. If we argue that p_F has no (well-defined) extension, then we are brought to treat natural falsity as formal quasi-falsity, and so, to model natural language via a formal notion of natural language. If instead, we negate the

21

supposition that the sentential predicates can be treated extensionally, then we must advert to some non-extensional notion of natural language. It would not be enough to say that the non-extensional notion of natural language is *intensional*, since intensions are meaningless without their extensions.)

Here, we employ the *reductio* to negate the supposition that there is a one-to-one onto function from the p_i to EPI (i.e., the set of the extensions of the p_i), and thus that EPI is uncountably large. (This treatment has already been suggested above.) Consequently, at least some of the p_i must have multiple extensions. Otherwise, some extensions (some subsets of SENT) must lack a p_i to express them, and then it could not be coherently maintained that natural language has the resources to express every extension. Accordingly, if natural language is to match every extension (every subset of (SENT), then some of the p_i must express multiple extensions. It can be shown that p_F, in particular, extends to uncountably many subsets of SENT. Because of this, we can treat p_F as an ambiguity, having multiple members of EPI for its extensions.. First (from the conclusion that EPI is uncountable), p_F has at least two extensions – one in which $\alpha = 0$, and another in which $\alpha = 1$. However, there are at least infinitely many fixed points on p_F. Trivially, there is an infinite series of the sentences,

$$s^*, (s^* \vee s^*), (s^* \vee s^* \vee s^*), (s^* \vee s^* \vee s^* \vee s^*), \text{ and so on.}$$

Obviously, since we can reason that s^* both is and is not a member of the extension of p_F, we can reason that $(s^* \vee s^*)$ *is* in the extension of p_F, and that $(s^* \vee s^* \vee s^*)$ is *not*. For convenience, let us express the immediately above series of fixed points on p_F by the series,

22

$$s*^1, s*^2, s*^3, s*^4, \ldots, s*^x, \ldots,$$

such that $s*^x$ is a disjunction of x-many $s*$. So for example, $s*^7$ is,

$$(s* \vee s* \vee s* \vee s* \vee s* \vee s* \vee s*).$$

We now make explicit the presence of the $s*^i$ on (EDA), below.

	s_0	s_1	...	$s*^1$	...	$s*^2$	...	$s*^3$	...
p_0	0	1							
p_1	1	1							
p_2	1	1							
p_3	0	0							
.									
.									
p_F	1	1	...	α^1	...	α^2	...	α^3	...
.									
.									

(EDA)

With respect to the three first $s*^i$, each of the following series of '0's and '1's is a legitimate extension for p_F.

```
1 1 ... 0 ... 0 ... 0 ...
1 1 ... 1 ... 0 ... 0 ...
1 1 ... 0 ... 1 ... 0 ...
1 1 ... 0 ... 0 ... 1 ...
1 1 ... 1 ... 1 ... 0 ...
1 1 ... 0 ... 1 ... 1 ...
1 1 ... 1 ... 0 ... 1 ...
1 1 ... 1 ... 1 ... 1 ...
```

Taking into account that the $s*^i$ are accordingly denumerably many, there are uncountably many series of '0's and '1's, each of which specifies a subset of SENT, and thus also, a possible extension for one of the p_i. (This follows from the well-known fact that the set of infinite strings of '0's and '1's is uncountable.)

23

Once again, this treatment of the paradox of p_F is not intended as a solution, but rather as a diagnosis. But it may not yet be clear what this diagnosis has to say about the paradox – we return to this matter below.

For another example of diagonal paradox, let us focus on Cantor's paradox. The universal set, U, is straightforwardly defined as the set of all the self-identical objects of the universe (or formally, $U = \{$ all objects, S, such that $S = S \}$). Of course, just like empirical objects such as apples, starlight and roadmaps, abstract objects are also self-identical. The constellations, the theory of evolution, the westerly (270°) compass-heading, the Charter of Rights and Freedoms, are all self-identical objects. Of course, sets are also self-identical abstract objects, and accordingly, they are members of U. Thus U includes all of its own subsets as members. (We will write '#(S)' for the cardinality of S; that is, for the number of members of S. Also, the *powerset* of S is the set of subsets of S, and we write this, '$\wp(S)$'.) Accordingly, the cardinality of the powerset of U is no greater than that of U itself (i.e., #($\wp(U)$) $\leq$ #(U), as everything in any set is already a member of U). But this cannot be, for Cantor's 1892 diagonally-proved powerset theorem [CAN2] (see also [KLE2, pp.180-183] *and* [MEN, p.183]) proves that for all sets, S, #(S) < #($\wp(S)$). (This theorem holds even when #(S) is a transfinite cardinality.) So, with $U = S$, it follows that #(U) < #($\wp(U)$), by Cantor's theorem. But by the universality of U, there can be no more members in *any other set* than there are in U. In particular, this means that #(U) $\geq$ #($\wp(U)$), with $\wp(U)$ as some 'other set'. (Note that '#(U) $\geq$ #($\wp(U)$)' is equivalent to '~{#(U) < #($\wp(U)$)}'.) But this is now a contradiction. Thus, to evade the contradiction, one must give up on either the existence of U or on the soundness of the powerset theorem. For a time, the notion of U

24

would not be let go. But the powerset theorem also could not be let go – for no fault could be found in it. So long as neither the existence of U nor the powerset theorem could be given up, one remained mired in paradox. (Cantor's paradox is sometimes reasoned according to the assumption that there is a largest transfinite cardinality, rather than according to the assumption that there is a universal set. These two initial formulations of Cantor's paradox are equivalent.)

Similarly, the liar paradox (i.e., Epimenides' paradox) is sometimes argued to show something that native speakers usually resist – that the natural language truth predicate is not extensional – that the set of all and only the true natural language sentences is ill-defined. That is, since *either* (a) the liar sentence and its negation are both in the set of the true natural language sentences – rendering this set inconsistent – *or* (b) neither the liar sentence nor its negation is a member of the set of the true natural language sentences – in which case the liar sentence is true, and thus the set of *all* (and only) the true natural language sentences does not contain at least one true sentence (namely, the liar sentence). In case (a), the set of all and only the true natural language sentences includes at least one falsehood (i.e., either the liar sentence or its negation), and so the natural language truth predicate does not extend to a set of (all and) *only* the true natural language sentences. In case (b), the natural language truth predicate does not extend to at least one natural language truth: the liar sentence. Yet it is hard to see what is at fault with the liar paradox – its *form*, at least, does not appear to be invalid. However, that diagonal argument has counter-intuitive instances is well-known. Besides Cantor's powerset theorem, two other unexpected diagonal arguments are those of Russell's 1905 proof of the inconsistency of the set-theoretic fragment of Frege's

25

Begriffsschrift and Gödel's 1931 incompleteness proofs. Gödel's 1931 diagonal theorems are responsible for proving a variety of counter-intuitive philosophical-sounding conclusions, including the failure of the logicist project and the inconflatability of proof with truth.[7] At the time of their discovery, these conclusions were unexpected, but neither of them has ever been convicted of significant error. Notwithstanding their philosophical flavour, these results are formally provable. Accordingly, the philosophical consequences of these theorems are well-developed, rigorously obtained and they strongly resist refutation. One might even say that they are *immune* to refutation, since if they were refuted, then inconsistency would immediately follow. However, inconsistency does not in all cases block further logical analysis. There are rigorous systems of paraconsistent logic in which *negation* inconsistency does not imply *absolute* inconsistency – whereby inconsistency is non-trivially managed.

The discussion above goes only to highlight the problem of deciding the (RA)s. The difficulty of making these decisions is well evidenced in the history of diagonal argument. Although mathematical logicians have accepted the correctness of Cantor's proof for over a century, it struck Cantor's contemporaries (including Cantor himself) as a mystery. Even now, talented and clear-minded (beginning) students resist Cantor's proof, since the first intuitions of most students report that there can only be one infinite cardinality.[8] The intuitions of (beginning) students are similarly offended by

---

[7] That the first incompleteness theorem has these two consequences is well-known. An accessible discussion of some of these will be found in [KNE, pp.712-742]. Of course, Hofstadter's [HOF] also makes an accessible discussion. For a more formal treatment of Gödel's theorems, see Boolos [BOO2], Kleene [KLE1], or Mendelson [MEN].

[8] This observation is based on no scientific evidence and no argument whatsoever. Rather, personal experience gained while tutoring beginning students of logic, and during general discussions on the logic of diagonal argument, the overwhelming majority have resisted – sometimes vociferously – the idea that some infinities are bigger than others (whether or not these infinites are completed or extant). Even professional

indefinability of the truth predicate, a diagonal result obtained by Tarski [TAR2].[9]

However, the intuitions of present-day logicians are difficult to offend by diagonal arguments such as those noted above. Indeed, some kinds of diagonal argument have undergone extensive refinements and generalizations. In particular, Gödel's incompleteness theorems are the subject of a voluminous technical literature – for two representative examples, see Smullyan's [SMU1] and Boolos *in* [JEF, *chapters* 26, 27 and 30]. Gödel's theorems have also been skilfully explained for the general public. Two well-known such explanations are due to Raymond Smullyan [SMU3] and Douglas Hofstadter [HOF]. It is also notable that an accessible article about incompleteness has been published in Scientific American [DAW2]. In the academic field of mathematical linguistics, first-order logic is studied as a means of modeling natural language, and thus the incompleteness theorems play a role in all such models (*see* [PART, part 'D', pp.317-429]). These examples of diagonal argument are now so well developed and integrated that one could be easily forgiven for believing that diagonal argument holds no more surprises for logic. One could even hold the view that diagonal argument has been *mastered* by its theory in mathematical logic, and thus that if diagonal argument has been unintuitive in the past, this is only because the logicians of the past did not have the benefit of contemporary research.

---

scientists (two, an astrophysicist and neuroscientist) have been unwilling to admit the validity of the diagonal method.

[9] There is no attempt here to draw conclusions about the correctness (or incorrectness) of anybody's logical intuitions. The claim is rather that the intuitions of students are usually, at first, violated by logical theorems such as Gödel's incompleteness proofs. Of course, the student's intuitions are expanded and refined the further he advances in his logical training – but this does not diminish the fact that (most of) the beginning students *start out* by doubting the validity of diagonal argument. The content of beginning students' intuitions is not relevant – just the existence of differences between the student's and the professional's logical intuitions. It is not claimed that a student's intuitions *show* any conclusions of a logical kind – only the (loosely) *meta*-logical fact that the student's intuitions and the professional's intuitions usually differ.

In this view, it may seem that the mysterious instances of diagonal argument (i.e., those diagonal arguments which conclude with the (RA)s) are just those instances which are not expressed within the formal languages of mathematical logic. Argumentative support for this view is found in the fact that these formal languages are highly general, both expressively and semantically. In consequence, very little is left which cannot be *said*, and very little else to speak *of*. The expressive generality of a formal language, say L, lies in L's adequacy to express the first-order formal arithmetic. (Hereafter, we say that a formal language is *strongly expressive* if it is sufficiently expressive to axiomatize the formal arithmetic.) But the axiomatization of the formal arithmetic requires only modest formal resources (such as a truth-functional syntax, some axioms – such as the Peano axioms – and a few rules of logical consequence), and thus Gödel's incompleteness theorems can be proved in *every* strongly expressive first-order language. As is well known, the incompleteness theorems employ highly expressive devices, sometimes called *Gödel-functions*, by which some of the metalanguage of the formal arithmetic is expressed in the object-language. These devices also require only the modest logical resources of the formal arithmetic.

Despite its modest resources, the formal arithmetic actually affords very generous degree of power, and is more than ample for the purposes of most mathematical logicians. L's *semantic generality* is given in two ways. First, by L's having all consistent models in its domain, be they finite, denumerable or uncountable.[10] *Second*, by the Church-Turing-Gödel thesis (hereafter simply *Church's thesis*) [CHU] any intuitively computable function is formally computable within L. Thus, provided that Church's thesis is true, it follows roughly that the functions of natural intuition reduce to

---

[10] That is, by *Skolemizing* one domain into another. (See [BOO2], pp.147 & *ff.*)

28

functions which are formally computable in L. Accordingly, there is very little left for other systems of logic to analyze, which L cannot. It must nevertheless be granted that within formal language, diagonal argument *is* unproblematic, or at least that it is unproblematic *so far*. But it would be a mistake to suppose furthermore that problematic (*i.e.*, mysterious) instances of diagonal argument – the (RA)s – can *always* be made unproblematic by revising the argument for treatment within a formal language.

But this is no reason to think that (RQ) is closed in the general case – neither for diagonal argument in formal language nor for diagonal argument in natural language. Although there are several strategies for answering (RQ), these strategies tend to offer only partial answers. Keith Simmons' theory of diagonal argument achieves one such answer to (RQ), and it is the best and most general answer that I have yet encountered. For the present, however, I want to discuss some other treatments of (RQ).

HERZBERGER AND BOOLOS ON SEMANTIC PARADOX

As confirmation of the fallibility of logicians' intuitions about absurdities, we consider articles by Hans Herzberger [HER1] and George Boolos [BOO3]. *First*, in his 1980 article, 'New paradoxes for old', Herzberger argues that the semantic paradoxes are in a state of 'eternal recurrence'.

> There is familiar pattern whereby resolutions for old paradoxes engender paradox anew. ... [O]ne would like to hope there might possibly be a "way out". But the lessons of history offer no encouragement in this respect. And the ideas I want to pursue here only show us how to quicken the cycle of recurrence. [HER1, p.109].

Herzberger's general conclusion is that no matter what devices one has for the pacification of paradox (that is, devices for deciding the (RA)s), there must always be the possibility of a new paradox, the provability of which is – as yet – unconstrained by any existing device. It is important to note that Herzberger's argument invokes considerable formal resources, but is not intended to lead to any formal results. Rather than being about paradox-in-formal-language, [HER1] is about paradox-in-natural-language. (Were there authentic paradoxes in a (non-trivial) formal language then it must be inconsistent, which conclusion is *either* to be flatly denied – since the classical systems of logic are assuredly *not* inconsistent – *or* we must advert to one or another paraconsistent system.) So if Herzberger is correct, then natural-language paradox is inescapable, and this is a near-absurdity to anyone who expects the eventual development of a (maximally complete and consistent) formal theory of natural language, for any such theory must admit one or more of Herzberger's 'recurring' paradoxes, and *a fortiori*, be inconsistent. Again, this inconsistency need not block rigorous logical analysis. Herzberger's

30

paradoxes could be employed to motivate paraconsistent logic. But for the time being, we keep our focus on consistent (principally first-order) systems of logic. In consequence, it follows that Herzberger's arguments confirm a central thesis of this (my) essay, *viz.*, that attempts to model natural language *via* formal language must be either incomplete or trivial. If otherwise, formal models are adequate to completely and non-trivially express natural language, then it must follow (via a 'genetic' argument) that the natural laws of the workings of the brain (i.e., the laws of biology, chemistry, neurophysics, and so on) are formally inconsistent (i.e., all formal theories thereof are inconsistent). Of course, the actual expressions of natural language (i.e., those bits of natural language which have been *actually expressed* by some human – or perhaps, even some *machine* – at some past time) are inconsistent, since anyone can lie, mis-speak himself, change his mind, and so on, but these cases are beside the point. They are problems of the language-user, whereas our interest is in the problems imposed by the language itself.

Having said this, it is important not to overstate the problem Herzberger presents for formal theories of natural language. Although there is always a formally undecidable natural language predicate, Herzberger's analysis shows that each new such predicate is more and more specific and limited with each new paradox-pacifying device. Thus with each successive iteration of these devices, the paradox becomes less and less problematic, and more and more like a 'mere technicality'.

Second, in Boolos' very short 1990 'On "seeing" the truth of the Gödel sentence' [BOO3], he observes that,

> [t]o concede that we can see the truth of the Gödel sentence for PA, in which only a fragment ... of actual mathematical reasoning can be carried

31

out, is not to concede that we can see the truth of Gödel sentences for more powerful theories such as ZF set theory, in which almost the whole of mathematics can be represented. [BOO3, p.389]

And furthermore,

I suggest that we do not know that we are not in the same situation vis-à-vis ZF that Frege was in with respect to naïve set theory ... before receiving ... the famous letter from Russell... It is, I believe, a mistake to think that we can *see* that mathematics as a whole is consistent. ... Are we really so certain that there isn't some million-page derivation of "0=1" that will be discovered some two hundred years from now? [BOO3, p.390] [*emphasis mine*]

Unlike Herzberger's, Boolos' article does not attempt an analysis or explanation of the 'eternal recurrence' of authentic paradox – and in consequence, it is perhaps somewhat the weaker of the two articles. However, Boolos' comments pertain to the availability of authentic paradox *for symbolic logic* (rather than the formal theories of natural language), which Herzberger's arguments do not. This suggests that the problem of deciding the (RA)s remains a (potential) problem for *symbolic logic*. For the present, however, there is no valid reason to suspect that ZF actually is inconsistent, and so we work under the view that it is consistent.

HODGES ON THE PHENOMENON OF AMATEUR REFUTATIONS OF
DIAGONAL ARGUMENT

Wilfred Hodges discusses the fact that many 'amateurs' (as he calls them) have made, and likely will continue to make, unsuccessful and unpublishable attempts to refute diagonal argument. Nevertheless, it seems that many of these amateurs invested a great deal of careful thought, and a few have managed to produce answers to (RQ), the errors of which are not obvious. It should be noted that most of these amateur logicians propose to answer (RQ) by arguing that *all* diagonal argument is bad. Accordingly, these amateurs' answers can be immediately faulted, since it is (virtually) universally accepted that at least some diagonal argument is 'good'. In these cases, the error may not be obvious, but the fact that there is an error is not in question.

The foregoing observations are given in Hodges' paper, 'An editor recalls some hopeless papers' [HOD]. In this paper, Hodges hoped only to discover the motivations of the amateurs' criticisms of diagonal argument, and not to develop an answer to (RQ). The question Hodges wants to answer is, "why so many people devote so much energy to refuting this harmless little [diagonal] argument – what had it done to make them so angry with it? [HOD, p.1]" Hodges does not propose to correct these 'hopeless papers'. Rather, the purpose is to present some of the reasoning behind the amateurs' (putative) refutations of diagonal argument, and thereby to help other editors when they are confronted with an amateur (putative) refutation of diagonal argument. Hodges concludes that some important points of basic logic are taught (by professional logicians) either "very badly, or not at all [HOD, p.1]." At greater length, Hodges says,

> a small number of the criticisms [of diagonal argument] are fair comment
> on misleading expositions [of logic text books]. A much larger number of

the criticisms are fair comment on some serious and fundamental gaps in the logic that we teach. Even at a very elementary level – I'm tempted to say *especially* at a very elementary level – there are still many points of controversy and many things that we regularly get wrong. [HOD, p.19]

The 'fundamental gaps' of logic instruction are not all gaps in the accuracy of the instruction – for in the formal case, diagonal argument is usually very technical, and would not have been taught in logic courses of a 'very elementary level', *at all*. Moreover, the errors made by the amateur logicians are principally errors which would vitiate any logical argument and so even a beginning student of logic could be expected to have avoided them. For example, some of the authors of the attempted refutations fail to realize the basic fact that "to attack an argument, you must find something wrong with it [HOD, p.5]." Other attempts to discredit diagonal argument include misguided attacks on justifications for suppositions or on the rules of logical consequence – justifications and rules which no well-trained mathematical logician would fault. All of these kinds of attempted refutations include obvious errors, and therefore do not bear further consideration.

Yet there are two papers (neither of which is explicitly considered by Hodges[11]) which attempt apparently stronger refutations. The first of these two papers (that of Leon Gumanski [GUM2]) shows that the author has considerable technical skill, but also includes a number of clear errors. It is not always obvious where Gumanski makes his errors, but the conclusions of his arguments are definitely mistaken. The errors of another paper (that of myself, [KIN]) are harder to see. [KIN] does not criticize diagonal argument in general, but only in the case of Gödel's incompleteness theorems, and then

---

[11] In fact, Hodges' paper does not include any references to the papers he discusses, in order to not violate copyrights of their authors.

only by implication. [KIN] nevertheless includes at least three critical errors. Both papers were peer-reviewed and published.

GUMANSKI'S PUTATIVE REFUTATION OF DIAGONAL ARGUMENT

Gumanski's paper was apparently translated to English, possibly by Gumanski himself, from another language. Unfortunately, the translation is not quite fluent, and since Gumanski's argument is technically presented, it is sometimes difficult to understand the argument. It is appropriate, therefore, to grant Gumanski a degree of charity. Gumanski begins with the conclusion of his 1983, 'On decidability of the first order functional calculus' [GUM1], wherein he claims to have given a decision procedure for first order logic. [GUM1] was published, but only in abstract. It is therefore impossible to identify the errors of [GUM1], which seem to be the principal source of the errors of [GUM2]. Gumanski states that the result of his 1983 paper contradicts "the well-known Church's theorem [CHU], and so the problem arises what an error has been committed in the proof of the theorem [GUM2, p.45]" Already there should be alarm, since if first order logic were decidable – in such a way that Church's theorem is contradicted – then Gödel's incompleteness theorems can hardly be sound. If, *contra* Church, first-order logic were decidable, then the provability of the Gödel sentence, G, must also be decidable. If G is decided to be provable, then it is false, and then there is a proof of a false sentence, and the inconsistency of first-order logic follows. If G is decided to be refutable, then its negation, $\neg G$, is thereby decided to be provable. (In the following, the predicate, 'PR' is the provability predicate, and the symbol '$\ulcorner ... \urcorner$' is Gödel-quotation.) From the provability of $\neg G$, it follows that $\text{PR}(\ulcorner \neg G \urcorner)$ is provable, by *derivability condition one*, stated below on page 46. But it also follows from the provability of $\neg G$ that $\neg\neg\text{PR}(\ulcorner G \urcorner)$, by Gödel's first incompleteness theorem: $G \leftrightarrow \neg\text{PR}(\ulcorner G \urcorner)$. Supposing that the logical rule of double-negation-elimination is available

36

(as we presently do), it follows that PR($\lceil G \rceil$) is provable. But then both PR($\lceil \neg G \rceil$) and

PR($\lceil G \rceil$) are provable. It is not exactly a formal contradiction for both PR($\lceil \neg G \rceil$) and

PR($\lceil G \rceil$) to be provable, but it is an important inconsistency – for the set of sentences

which satisfy the PR predicate must be inconsistent, and thusly, the provable sentences of

first-order logic imply a formal contradiction. So if first-order logic is decidable, then

first-order logic is inconsistent. The quasi-contradiction PR($\lceil G \rceil$) $\land$ PR($\lceil \neg G \rceil$) can lead

to inferences which do have the form of a contradiction. We write '<u>PR</u>' for the set of

sentences which satisfy the PR predicate. Then,

$$\text{PR}(\lceil G \rceil) \leftrightarrow (\lceil G \rceil \in \underline{\text{PR}}), \text{ and,}$$

$$\text{PR}(\lceil \neg G \rceil) \leftrightarrow (\lceil \neg G \rceil \in \underline{\text{PR}}).$$

For non-paraconsistent systems, we now have

$$\lceil G \land \neg G \rceil \in \underline{\text{PR}}.$$

If we already have proofs for both PR($\lceil G \rceil$) and PR($\lceil \neg G \rceil$), then we also have a proof for

$G \land \neg G$, which formula clearly has the form of a contradiction. Another way of making

the quasi-contradiction of PR($\lceil G \rceil$) $\land$ PR($\lceil \neg G \rceil$), a contradiction of the form, $A \land \neg A$,

employs an unprovable formula, which we have already seen as (KE) and (PPE): *viz.*, for

all formulas, A, 'PR($\lceil A \rceil$) $\rightarrow A$'. Supposing that (KE) is provable, PR($\lceil G \rceil$) implies G,

and PR($\lceil \neg G \rceil$) implies $\neg G$. By conjoining these two (supposedly) provable formulas, G

and $\neg G$, we can prove the explicit contradiction, $G \land \neg G$. However, as (KE) is

unprovable, (KE) cannot *actually* lead to a proof of $G \land \neg G$. But there is still another

step that strengthens the attempt to obtain an explicit contradiction from PR($\lceil G \rceil$) $\land$

37

PR($\ulcorner \neg G \urcorner$). Although (KE) is not provable, its metalinguistic counterpart *is* a theorem of the metalanguage, namely,

$$\text{if } \mathbin{|\!\!-} \text{PR}(\ulcorner A \urcorner), \text{ then } \mathbin{|\!\!-} A, \hspace{3cm} \text{(MKE)}$$

(for all formulas, A). Thus, if we already have a proof of PR($\ulcorner G \urcorner$), *then* we can prove G. (The same holds for PR($\ulcorner \neg G \urcorner$); if it is provable then so is $\neg G$. Thusly, if we already have proofs for PR($\ulcorner G \urcorner$) and PR($\ulcorner \neg G \urcorner$), then $G \wedge \neg G$ is provable. However, as (MKE) is metalinguistic, so is the conclusion it leads to:

$$\begin{aligned} &\textit{if } \mathbin{|\!\!-} \text{PR}(\ulcorner G \urcorner) \textit{ and } \mathbin{|\!\!-} \text{PR}(\ulcorner \neg G \urcorner), \\ &\textit{then } \mathbin{|\!\!-} (G \wedge \neg G). \end{aligned}$$

Furthermore, if Gödel's proof of incompleteness is unsound, then other central theorems of logical metatheory are also implicated in error. The unprovability of consistency, the inconflatability of proof with truth, the unsolvability of the halting problem and much of recursion theory could also be expected to fall. Because of this, Gumanski's argument is already hard to take seriously – too many logicians would have had to make the same mistakes over and again, starting (at the latest) with Gödel's 1931 [GOD1], and running continuously right up to the present. Yet Gumanski does apparently intend that these central theorems of formal logic should fall. Indeed, he wishes to implicate *all* diagonal argument with vitiating error. Gumanski states outright,

> The aim of the present paper is to demonstrate that despite its distinguished credentials the [diagonal] method is unreliable and *all* the purported proofs in which it is employed – though not necessarily their theses – ought to be doomed to oblivion. [GUM2, p.45] [Emphasis added.]

One might already be tempted to put Gumanski's papers away, and never return to them. Yet, on the assumption that there actually is a problem with the diagonal method, then

38

we must expect it to be a *subtle* problem. If the problem were sufficiently subtle, then it is (at least weakly) plausible that it has escaped the notice of mathematical logicians for the thirteen or so decades since Cantor first introduced his diagonal method, however unlikely this seems. So we grant Gumanski another dose of charity and consider his arguments a little longer.

Gumanski's thesis pivots on the matter of whether or not the counter-diagonal exists. In indirect diagonal argument, the only assumption Gumanski allows for negation by *reductio* is the supposition that the counter-diagonal does exist. (Gumanski discusses his reasons for imposing this restriction, but they are contrary to the usual account of indirect argument and quite uncompelling.) Thus, for Gumanski, indirect diagonal argument can show only that the counter-diagonal does *not* exist, or else the argument must be unsound. Direct diagonal arguments are also criticized for not including proof that the counter-diagonal exists. In this case, Gumanski seems to be satisfied with argument to the effect that the existence of a definition of a counter-diagonal is not sufficient to demonstrate its existence. It is implied that there is no reason besides definitional *fiat* to think a counter-diagonal exists. Here, Gumanski seems again to argue *contra* a result of formal logic which is both long-standing and universally accepted. The fixed point theorems prove precisely that there *are* effectively specifiable formal constructions which work as counter-diagonals. But Gumanski neither mentions nor discusses fixed point theorems, and this appears to be a very serious omission. Gumanski's errors can be clearly and compellingly demonstrated. We consider Gumanski's claims as they relate to the diagonal ARRAY (EDA) [*see* p.19 above]. The diagonal of (EDA) is given by the ARRAY's values at the intersections of p_0 and s_0, p_1

39

and s_1, p_2 and s_2, and so on. The value of the diagonal is thus, '0110...'. Accordingly, the value of the counter-diagonal of (EDA) is '1001...'. Now, if Gumanski is correct, and (EDA) has *no* counter-diagonal, then he must maintain that the counter-diagonal string does not *exist*. But Gumanski's inference here is very hard to take seriously, for it is intuitively obvious that our method of specifying the counter-diagonal of (EDA) is certain to yield an infinite string of '0's and '1's. *A fortiori*, (EDA)s counter-diagonal *does* exist. Gumanski might argue that although the value of our so-called counter-diagonal exists, it is not the genuine counter-diagonal of (EDA). But this claim is also very hard to accept – for what else could the genuine counter-diagonal of (EDA) be? Furthermore, even if we have failed to correctly determine (EDA)s counter-diagonal, then the correct counter-diagonal of (EDA) *must* be some other infinite and effectively specifiable string of '0's and '1's.

Gumanski makes another serious error where he says,

> we must consider the existence of [counter-] diagonal sequences as highly uncertain and their definitions as inadmissible, at least as long as it is not demonstrated that the defined diagonal sequence exists. Consequently, if such a demonstration has not been given, we must recognize any proof (even a direct one) that employs the diagonal method to be fallacious. As a matter of fact, the demonstration is lacking in proofs which are not based on axiomatic set theory. [GUM2, p.51]

The first and second sentences of this quote are not well supported by Gumanski's foregoing arguments. The third sentence is plainly mistaken. Gödel's 1931 is a paradigm example of diagonal argument, and it is not based on set theory, and since 1931, there have been a great many non-set-theoretic re-formulations of Gödel's incompleteness proofs. There have also been a great many extensions of incompleteness to formal systems which are not characteristically set-theoretic.

Gumanski next announces "a new shape" for the "problem of antinomies". (Read for '(NSA)', *new shape of antinomy.*)

> (NSA) An antinomy is not a mysterious logical puzzle, nor is it a threat to the very foundation of logic. It is simply a quite innocuous proof in which the derived absurdity demonstrates that ... *the defined object* [i.e., the counter-diagonal] *does not exist and the applied definition is inadmissible.* [GUM2, p.53] [Emphasis mine.]

The keystone of Gumanski's argument is the emphasized sentence above – all of his conclusions depend on it. While there is no obvious error with the above claim that there are no (true) antinomies for formal logic – that is, no 'show stoppers' – Gumanski's argument for this claim is not strong. Finally, Gumanski advances the following two conclusions. (For '(GC1)' and '(GC2)', read *Gumanski's conclusion one,* and *Gumanski's conclusion two.*)

> (GC1) [W]e must consider all proofs constructed according to the diagonal method and based on axiomatic set theory as not convincing, as mere samples of logical deduction grounded on arbitrary assumptions. The method does not make a reliable instrument for scientific investigations. [GUM2, p.55]

> (GC2) A. Church in his proof that the functional calculus of first order is undecidable has made use of the diagonal method. And that is his mistake. The calculus has its own objective properties independent of our will. It is decidable as I managed to demonstrate in my paper [GUM1]. No arbitrary decisions, no assumptions, no axioms can help it. [GUM2, pp.55-56]

Although (GC1) is expressed only for diagonal argument which is 'based on axiomatic set theory', we can strengthen it to include all classical first-order languages in which the formal arithmetic is expressable. But (GC1) is faulty whether or not it is strengthened, for (the emphasized part of) (NSA) is false. The fixed point theorems do precisely what Gumanski denies – they prove the *existence* of the fixed points concerned. Indeed, Haskell Curry has shown how to construct a fixed point *combinator.*

41

The first sentence of (GC2) is unobjectionable, but the second sentence is too bold by far, and is poorly supported in any case, since it depends on Gumanski's false claim that fixed points do not exist. The third sentence may also be granted, but it is hard to see the connection it bears to the rest of the paper; this sentence appears more as a metaphysical thesis than a logical one. The fourth sentence may be true, but seems much more likely to be false. If 'the calculus' (i.e., first-order logic) were decidable, this would undermine a huge part of formal logic, namely, those parts which are consequent to incompleteness. The fifth sentence could also be granted, but seems to need further explanation, and in any case, this sentence appears to have more of a rhetorical purpose than a logical one.

As a final comment, we may say that even if there are subtle causes which render diagonal argument unsound, Gumanski has been unsuccessful in his effort to reveal them. His reasoning depends uncritically on several mistakes, and until those (apparent) mistakes are addressed, Gumanski's conclusions must be evaluated as – to use Gumanski's own words – "highly uncertain" and "not convincing".

KING ON A PROBLEM CONCERNING DIAGONAL ARGUMENT

My 2001 [KIN] can be construed as a challenge to the correctness of diagonal argument, although this was not my intent at the time it was written. (I have never conceived this paper as an attack on diagonal argument.) Like Gumanski's paper, mine includes (at least two) vitiating errors, and there can be no serious debate about their existence. Unlike Gumanski's paper, mine is very short and the argument is reasonably clear, or leastways, reasonably clear to logicians. In spite of the brevity and clarity of [KIN], its errors are still hard to see -- or at least, they were hard for me to see. (I believe that the errors were also hard for a number of professional logicians to see, though only for a short time. My errors were explained to me principally by Jonathan Seldin and Bryson Brown, who also brought me to a more complete understanding of the complex interplay of metalogical concepts and theorems. I am also indebted to Dov Gabbay, Charles Morgan and Solomon Feferman for helpful comments. Jeff Pelletier was also very helpful, sending me the papers by Hodges, Gumanski and Zvonimir Sikic [SIK].)

[KIN] is divided into three parts. We will consider only a simplification of part one, which is fairly straightforward. In part two, we consider an argument that the formula, 'PR($\lceil A \rceil$) $\rightarrow$ A', is provable (where 'PR' is a provability predicate, and A is an arbitrarily chosen well-formed formula). Part three will not be restated here, as it is only an object-linguistic proof of a formula which expresses Gödel's first incompleteness theorem, in the formal language of Gödel's 1931 paper. The following (putative) proof is modified in two ways from its presentation in [KIN]. *First*, derivability predicates (or 'turnstyles' -- '|--') are inserted before each line of (KP) in order to stress that this (putative) proof is meta-linguistic, and thus that it does not necessarily give an object-

43

linguistic proof of either G or P. (Indeed, there is *actually no* object-linguistic proof of either G or P.) *Second*, it is made explicit that (KP) depends on the provability of the formula, $\{PR(\ulcorner F\urcorner) \to F\}$ (which is stated shortly below as '(KE)'), and accordingly, the conclusion of (KP) is expressed as a conditional: '*if* |-- $\{PR(\ulcorner F\urcorner) \to F\}$, *then* |-- P' (where F and P are any two arbitrarily chosen formulas). (For 'KP', read 'King's paradox'.) Part one:

(KP) Let L be a sound classical language in which the first incompleteness theorem is provable. Let $PS(x,y)$ express 'x is (the Gödel number of) a proof schema for the formula (Gödel-coded by) y'. Let PR be a provability predicate for L. (We define '$PR(x)$' as '$\exists y\{PS(y,x)\}$'). Let G be a Gödel sentence for L, and let the formula P be arbitrary.

(0)	|-- $\{PR(\ulcorner G\urcorner) \to G\}$	[Assumption]
(1)	|-- $\{G \leftrightarrow {\sim}PR(\ulcorner G\urcorner)\}$	[The 1st incompleteness theorem]
(2)	|-- $\{{\sim\sim}PR(\ulcorner G\urcorner) \vee G\}$	[A consequence of (1)]
(3)	|-- ${\sim\sim}PR(\ulcorner G\urcorner)$	[Assumption for $\vee$-elimination on (2)]
(4)	|-- G	[(0) and (3)]
(5)	|-- G	[Assumption for $\vee$-elimination on (2)]
(6)	|-- G	[$\vee$-elimination: (2) to (5)]
(7)	|-- $PR(\ulcorner G\urcorner)$	[By 'if |-- G then |-- $PR(\ulcorner G\urcorner)$']12
(8)	|-- ${\sim}PR(\ulcorner G\urcorner)$	[(1) and (6) by *modus ponens*]

(where the asterisk * appears beside line (4))

(7) and (8), when conjoined, form a contradiction. Now by *ex falso quodlibet*, it follows that |-- P. Whereas the assumptions at (3) and (5) have been discharged, that of (0) has not. Thus the conclusion is, '*if* |-- $\{PR(\ulcorner G\urcorner) \to G\}$, *then* |-- P'.

It is already well-known that, in consequence of Löb's theorem, if |-- $\{PR(\ulcorner G\urcorner) \to G\}$, then the consistency of L must fail. This is the first error of my paper.

---

12 This meta-linguistic formula ('if |-- G then |-- $PR(\ulcorner G\urcorner)$') is well-known as one of the *derivability conditions*. If anything can count as an axiomatization of provability, then so can the derivability conditions. But there is also a formal proof of this derivability condition, outlined as follows. Let $n = \ulcorner G\urcorner$; let m be the Gödel number for the concatenation of lines (1) to (6). Then m is (the Gödel number of) a proof schema for G. Thus $PS(m,n)$ holds, and as PS is computable, $PS(m,n)$ is provable. *A fortiori*, PR(n) is also provable, and as $n = \ulcorner G\urcorner$, $PR(\ulcorner G\urcorner)$ is provable.

The second error occurs in part two, where I attempt to prove by mathematical induction that for all formulas, F,

$$\text{(KE)} \quad \vdash \{PR(\ulcorner F \urcorner) \to F\}.$$

(We shall henceforth call this formula, '$\vdash (PR(\ulcorner F \urcorner) \to F$', *King's error*, or 'KE'.) I will not reproduce part two here. It suffices to note that my employment of mathematical induction is inessential and invalid, since the conclusion drawn does not actually depend on the correctness of the induction case.

If (KE) were provable, then it would follow that Gödel's incompleteness theorems harbour an inconsistency. Since the incompleteness theorems depend on a diagonal argument, their inconsistency could be taken as an absurdity for *reductio* against the soundness of Gödel's diagonal argument. My first error is one of failing to acknowledge that there is already significant literature proving that (4) does not follow from (3) – leastways, not in any consistent formal system. The part-two attempt to prove (KE) is much the same error as that of the first part. However, strictly speaking, the second error is not the attempt itself to find proof for (KE), but the commission of a mistake in that attempt.

(KP) AND THE LOGIC OF PROVABILITY: HENKIN'S PROBLEM,
LÖB'S THEOREM AND THE DERIVABILITY CONDITIONS

We now delve deeper into the literature on Henkin's problem, Löb's theorem and

the derivability conditions, which proves the unprovability of (KE). The implication of |-

- PR($\lceil F \rceil$) from |-- F is well known as the first 'derivability condition' ('DC1', shortly

below). Informally, the reasoning behind this consequence runs as follows.

For any formula, F, if we have |-- F, then F is provable. Thus there is a proof

schema, H, such that PS($\lceil H \rceil, \lceil F \rceil$) holds true. As PS is recursive, if follows that |--

PS($\lceil H \rceil, \lceil F \rceil$). By the above definition of PR, we now have |-- PR($\lceil F \rceil$). With 'F' = 'G',

we have |-- G, and thus it follows that |-- PR($\lceil G \rceil$). (The reader may note that line (6) of

(KP) depends on the supposition of line (4) that |-- (PR($\lceil G \rceil$) $\rightarrow$ G), which is false.

Accordingly, we should accept neither line (6) nor line (7) of (KP). We will consider this

error immediately below, but it nevertheless remains that (for all formulas, F) if |-- F then

|-- PR($\lceil F \rceil$), and thus also, if |-- G then |-- PR($\lceil G \rceil$).)

(DC1) is a metalinguistic formula. As might be expected, the validity of (DC1)

implies validity for its object-linguistic counterpart -- *prima facie*, '|-- $\{F \rightarrow$ PR($\lceil F \rceil$)$\}$'

follows by conditional proof from 'if |-- F then |-- PR($\lceil F \rceil$)'. The converse of (DC1) is 'if

|-- PR($\lceil F \rceil$) then |-- F', which is a metalinguistic counterpart of (KE). (Let us call this

formula, 'MKE', for 'metalinguistic KE'.) Both (DC1) and (MKE) are valid formas of

inference. However, unlike the object-linguistic counterpart of (DC1), (KE) is not valid.

This fact was first proved by Loeb, in his famous *Loeb's theorem* [LOB]. This theorem

came as a reply to a problem presented by Leon Henkin. Henkin's problem is (HP),

following.

46

(HP) Where H is a formula of a strongly expressive system, T, such that $\lceil H \rceil = \lceil PR(\lceil H \rceil) \rceil$, is H provable in T, or is H independent of T? (*see* [HEN, p.160, *problem 3*])

The formula, H, is a (so-called) 'Henkin sentence'. Stated more straightforwardly than (HP), Henkin's problem is,

(HP*) is it provable or not that $H \leftrightarrow PR(\lceil H \rceil)$?

Since it concerns the provability of a biconditional, (HP*) can be separated into the following two questions.

(HP*1) Is it provable or not that $H \rightarrow PR(\lceil H \rceil)$?

(HP*2) Is it provable or not that $PR(\lceil H \rceil) \rightarrow H$?

As has already been discussed, all formulas of the form '$F \rightarrow PR(\lceil F \rceil)$', are provable, and thus $H \rightarrow PR(\lceil H \rceil)$ is also provable. (This is due to the satisfiability of Löb's *condition III* and *condition IV* [LOB, p.116], which are now understood to be expressed by (DC1).) It is (HP*2) which Löb's theorem answers. Löb's theorem is (LT) below.

(LT) For every formula, F,
    if $\vdash (PR(\lceil F \rceil) \rightarrow F)$, then $\vdash F$. (*see* [LOB])

It is an immediate consequence of (LT) that when F is a Henkin sentence, H (i.e., $\lceil F \rceil = \lceil H \rceil = \lceil PR\{\lceil H \rceil\} \rceil$), we have,

$\vdash \{PR(\lceil H \rceil) \rightarrow H\}$, (†)

and thus,

$\vdash H$.

It follows immediately that (HP*1) and (HP*2) are both answered in the affirmative, but this is so *only* when H is a Henkin sentence. Otherwise, H might be a contradiction and

47

then by (†) a contradiction is immediately proved. Thus even though $PR(\lceil H \rceil) \to H$ is provable whenever H is a Henkin sentence, it is not provable in the general case.

For the sake of expositional completeness, it is necessary to state the other derivability conditions. These conditions appear to been first introduced by Löb, who counted five of them, but these have since been reduced to three. (Even in these three, there appears to be a redundancy, for (DC3) is a *prima facie* consequence of (DC1).) The three are,

for all formulas, A and B,

(DC1)   If $\vert\!\!-\!\!- A$, then $\vert\!\!-\!\!- PR(\lceil A \rceil)$.
(DC2) $\vert\!\!-\!\!- \{PR(\lceil A \to B \rceil) \to (PR(\lceil A \rceil) \to PR(\lceil B \rceil))\}$.
(DC3) $\vert\!\!-\!\!- \{PR(\lceil A \rceil) \to PR(\lceil PR\{\lceil A \rceil\} \rceil)\}$.

At the end of his short paper on (HP), Löb explains his (LT) as if it were a theorem of natural language. This explanation is strikingly similar to Curry's paradox, and it would be interesting to know whether Löb and Curry discovered this paradox independently of each other. (Curry's paradox plays a key role in the latter parts of this essay.) Löb writes,

> The method used in the previous proof leads to a new derivation of paradoxes in natural language. For let A be any sentence, and let B be the sentence,
>
> "If this sentence is true, then so is A."
>
> Now we easily see that, if B is true, then so is A. That is, B is true. Hence A is true. We have thus shown that every sentence is true. [LOB, p.117]

Still, an apprentice logician could be forgiven for seeking a proof of (KE) – or so I believe – especially if he had not yet encountered the literature which proves the unprovability of (KE). Indeed, there is proof of a metalinguistic formula which strongly

48

(but erroneously) suggests that (KE) is provable, namely, (PPE) below. It is known that for all formulas, F,

$$\text{if } |\!- PR(\ulcorner F \urcorner), \text{ then } |\!- F. \qquad\qquad \text{(PPE)}$$

(Read for '(PPE)' 'provability predicate elimination'. Note that (PPE) is (MKE).) If we now suppose $|\!- PR(\ulcorner F \urcorner)$ then by (PPE), it follows that $|\!- F$. Thus, if we already have $|\!- PR(\ulcorner F \urcorner)$, then by (PPE) we have $|\!- F$. It thus appears that if we already have $|\!- PR(\ulcorner F \urcorner)$, then (PPE) implies (KE), but this appearance misleads.

In order to use (PPE) to prove (KE), one must first have $|\!- PR(\ulcorner F \urcorner)$, *for arbitrary F*. It is not enough to *assume* the provability of $PR(\ulcorner F \urcorner)$; one must have *actual proof*. This fact is manifest in the observation that whenever F is not provable, there is – in the *first* place – no n such that $PS(n, \ulcorner F \urcorner)$ holds, and *a fortiori*, $PR(\ulcorner F \urcorner)$ does not hold either. So, when F is not provable, the assumption $PR(\ulcorner F \urcorner)$, *is false*, notwithstanding any assumption that it is true. In other words, $PR(\ulcorner F \urcorner)$ *immediately implies* $\sim PR(\ulcorner F \urcorner)$ whenever F is not *actually* provable. Accordingly, (KE) is no consequence of (PPE).[13]

The proof of (KE) from (PPE) is *good-looking* (at least for some beginning students of logic), but it is not *good*. (That is, the proof of (KE) from (PPE) is not sound.) That (KE) does not follow from (PPE) seriously undermines *ordinary* intuition, but this undermining does not – *cannot* – soundly imply the provability of (KE).[14] The result (that (KE) does not follow from (PPE)) is better understood as an unexpected fact. Yet if one were only an amateur logician, it would be easy to satisfy him or her that (KE)

---

[13] A person might think that I am here arguing that (KE) – '$|\!- PR(\ulcorner F \urcorner) \rightarrow F$' – is provable in the object language. But this is not so, as re-consideration of sentences nine through fourteen of this paragraph will confirm.

[14] It might be thought that this comment is tempting only 'if we slide too easily from the meta-linguistic demonstration that there is a provability predicate to the assumption that a parallel argument can be laid out in the object language'. This thought is correct, since a *beginning student* might be -- *actually is* (as I believe) – prone to committing the error of this slide.

does follow from (PPE) (even though (KE) does not *actually* follow from (PPE)), and thus that Gödel's incompleteness theorems harbour inconsistency (even though there is no *actual* inconsistency in Gödel's incompleteness theorems). This goes to show that there is mystery in Gödel's diagonal argument, which could easily draw a student away from forming a correct understanding of the proof-theoretic issues surrounding (PPE) and (KE).

It is also worthy of mention that the argument of [KIN], which we now see is merely 'good-looking', could be converted into a 'good' argument of the unprovability of consistency. (The possibility of this conversion was brought to my attention and given by Jonathan Seldin.) The fact that this conversion is possible is further evidence that although [KIN] includes vitiating errors, it is not *garbage*. (Hodges uses this word – 'garbage' – to describe the attempted refutations of the diagonal argument which he considers in his paper.)

TWO NOTIONS OF LANGUAGE: LINGUA CHARACTERICA AND
CALCULUS RATIONATOR

Via two notions (*lingua characterica* and *calculus rationator*) which predate

Gottlob Frege's Begriffsschrift [FRE1] and the ensuing development of mathematical

logic, we observe that natural language is a kind of mathematical logic only if natural

language can be completely subsumed within a *calculus rationator*. I argue that natural

language is not a *calculus rationator*, but rather a *lingua characterica*. In contemporary

terms, this is the view that natural language is *language-as-universal-medium-of-all-*

discourse (i.e., the view of *lingua characterica*), and I oppose the view that natural

language is *language-as-logic* (i.e., a *calculus rationator*).

A paper by Jean van Heijenoort engages this distinction in the context of Frege's

thoughts on the matter [vHE2]. As van Heijenoort explains it, the distinction between

lingua and *calculus* concerns the semantic universality of formal logic, such that *lingua*

characterica is the more universal, and *calculus rationator* the less. But the distinction is

not exclusive. In Frege's view of logic, the propositional calculus is *calculus* but not

lingua, whereas quantification theory (i.e., that of Frege's Begriffsschrift) is not merely

calculus, but also *lingua*. It is a central point of van Heijenoort's paper that Frege did not

conceive quantification theory in the same way as it is now understood, say, within the

discipline of symbolic logic, where quantification theory can be conceived altogether

divorced from semantic issues. At least in its early days, Frege intended the

Begriffsschrift to be embedded in *lingua*. (And when his successors went *mathesis*

universalis (i.e., a semantically universal *calculus rationator*), Frege went vociferously

anti-psychologistic.) Via its being embedded in the *lingua characterica*, Frege conceived

51

quantification theory as semantically universal. Frege seems moreover to have assumed

that quantification theory is complete, in the sense that Frege – and later also Russell –

read the <u>Begriffsschrift's</u> judgement stroke (i.e., the turnstyle: '|--') to express truth rather

than provability, as it is now read. Van Heijenoort comments further,

> The universality of logic [i.e., of quantification theory] expresses itself in an important feature of Frege's system. In that system the quantifiers binding individual variable range over all objects. ... Boole [i.e., Boole's propositional calculus] has his universe class, and De Morgan his universe of discourse, ... [b]ut these have hardly any ontological import. They can be changed at will. ... For Frege it cannot be a question of changing universes. One could not even say that he restricts himself to *one* universe. His universe is *the* universe. ... Frege's universe consists of all that there is, and is fixed. [vHE2, p.325]

Van Heijenoort observes that owing to the universality which Frege conceived for the

<u>Begriffsschrift</u>, "nothing can be, or has to be, said outside of the system" [vHE2, p.326].

Although he recognizes that any formal system needs meta-systemic notions and rules

(i.e., notions such as *consistency, completeness*, the *independence* of axioms), Frege

never discusses any of these notions.

Moreover, the <u>Begriffsschrift</u>-as-*lingua* has to be learned in the same way as

natural language is learned, as van Heijenoort explains, "by suggestions and clues".

Also,

> Frege repeatedly states, when introducing his system, that he is giving 'hints' to the reader, that the reader has to meet him halfway and should not begrudge him a share of 'good will'. The problem is to bring the reader to 'catch on'; he has to get into the language. [vHE2, p.326]

Ultimately, Frege's conception of the <u>Begriffsschrift</u>-as-semantically-universal-*lingua*

failed. The initial cause of this failure is the Russell set (hereafter, 'RS'). The definition

of RS is both intuitively well-formed, and more important for Frege, it is an *admissible*

definition in the language of the <u>Begriffsschrift</u>. But, as we well know, RS's membership

52

is inconsistent (i.e., since it can be proved in the Begriffsschrift both that RS $\in$ RS and

RS $\notin$ RS). It seems that Frege immediately recognized the scope of the problem which

RS visits on the Begriffsschrift, for he despaired of the success of his philosophical thesis

(i.e., the thesis *contra* Kant's doctrine of the synthetic *a priori* that mathematics can be

completely represented in an analytic language, that is, in a logical language). This thesis

of Frege's is now known as the 'logicist' thesis, although Frege himself never used this

term. In other ways, the Begriffsschrift was still a great success, for example, in virtue of

Frege's creation of a generalized quantifier theory, the function-argument distinction and

the sense-reference distinction. (The last of these was actually first presented in Frege's

Sinn und Bedeutung [FRE3], but was nevertheless implicit in the Begriffsschrift.) So

other logicians did not give up on Frege's insights, and made considerable efforts to

maximize the successes of so-called 'Fregean' languages. After all, even though the

consistency of the Begriffsschrift failed, it does not follow that Kant's synthetic *a priori*

is thereby re-confirmed. In their Principia Mathematica [WHI], Russell and Whitehead

adapted the pivotal insights of the Begriffsschrift to a new symbolic system for formal

language. The two most important contributions of the Principia were a much-simplified

and more intuitive notational system, and type theory. Type theory pacified the problem

of the RS by rendering so-called 'impredicative' definitions (*see* [KNE, pp.513-521])

inadmissible – if the RS cannot be defined, then it has no membership relation, and no

inconsistent membership and hence no contradiction (within the Principia) could be

proved from the RS.

But Russell sets are not the only kind of formal notions that weaken Fregean

languages. The famous Gödel sentence, which is similar to the RS in that they are both

fixed points, implies that not every mathematical truth corresponds with a mathematical

proof. This result, due to Gödel [GOD1], is widely regarded as proof that Frege's

logicism is an unrecoverable failure, for it demonstrates not only that the Principia is

incomplete, but also that *every* Fregean language is incomplete. It follows immediately

that the formal arithmetic is incomplete*able* as a Fregean language. However, research

into Fregean language has not only continued, but *flourished*.

Had Frege succeeded in his efforts, he would have achieved a mode of language

which Jaakko Hintikka identifies as *mathesis universalis* [HIN, pp.58-59]. In words we

have already used, a *mathesis universalis* could be understood as a semantically universal

calculus rationator. Were Frege to have achieved *mathesis universalis* for the

Begriffsschrift, natural language could be subsumed within formal quantifier theory,

without relevant loss of expressability, and with the added capacity to mechanically

determine mathematical and logical truth and falsity. Of course, *mathesis universalis* still

has not been achieved, and is moreover expected to be unachiev*able* via Fregean formal

systems (in consequence of Gödel's incompleteness theorems).

In section two of his paper, Hintikka meets and further develops the *lingua-
calculus* distinction of van Heijenoort's paper.

> [Frege] is ... representative of the type of view labelled by van Heijenoort
> 'logic as language,' [i.e., *lingua characterica*] which perhaps rather ought
> to be called in more general terms the view of language as the universal
> medium of all discourse. According to this view ... we cannot escape the
> basic semantical relationships that connect our language with reality.
> Since they are presupposed in anything we say, they cannot be
> meaningfully talked about. [HIN, p.58]

That is, we cannot do the semantic theory of a language *from within* that language

without begging the question of what should be the content of that semantic theory. And

54

this is why Frege eschewed formal semantics. On languages *calculus rationator*, Hintikka writes,

> on the view labelled by van Heijenoort 'logic as calculus' [i.e., *calculus rationator*], we can meaningfully and nontrivially discuss the links between our own language and reality. ... In other words, we can to some extent at least think of our language and its logic as if it were a calculus, not in the sense that it is a meaningless formal game as a calculus is, but in the sense that it can be reinterpreted like a calculus. [HIN, pp.58-59]

Of course, if *calculus* is subsumed within *lingua*, which it must be if our *lingua* is to be semantically universal, then the development of a semantics for the *calculus* inherits the problems of developing a semantics for the *lingua*. Accordingly we must assume that *lingua* can be reinterpreted as *calculus* – for otherwise, we can do no semantic theory at all. As Hintikka puts it, "the development of all serious truth-conditional semantics (model theory) obviously presupposes adopting the conception of language as calculus. [HIN, p.59]" Consequently, "we can practice systematical semantics only if we can meaningfully discuss these relationships as we cannot do on the view of language as the universal medium." In the *calculus rationator* view, the meta-systemic questions which Frege eschewed are straightforwardly definable. Hintikka again:

> [L]ogical validity is, when it is explicitly defined, a concept which requires a tacit reinterpretation of the representative relationships between language and reality: a [formal] sentence [of a *calculus*] is valid if and only if it is true on every possible reinterpretation of its nonlogical ... concepts. Hence a completeness proof (a proof that each logically valid sentence of a given language is provable) presupposes the idea of language as calculus. [HIN, p.59][15]

Yet one is cautioned to avoid using truth-conditional semantics to explain or model the semantics of a *lingua characterica*.

---

[15] Hintikka's notion of validity here is the formal notion, not the notion of a *necessary* or *tautological* truth. Hintikka's notion depends only on our understanding of the *calculus'* logical constants.

55

[A] mistake – 'the first dogma of the philosophy of language' – often is in the offing when an attempt is made to elicit [universal] semantical morals from the rules of logical inference which naturally are formulated in syntactical (formal) terms and which establish, when correctly applied, logical truth rather than truth *simpliciter*. [HIN, p.61]

It could be argued that as long as logical truth and truth *simpliciter* are indistinguishable, there is effectively no distinction between *calculus* and *lingua*.

Simmons' comment, that the semantic paradoxes are "generated by diagonal arguments in exactly the same way that the set-theoretic paradoxes are generated" [SIM, p.37]. This comment depends on the view that the arguments which produce the semantic paradoxes are all subsumable within a *calculus rationator*. Notwithstanding Simmons' comment, diagonal argument *in general* cannot be easily shown to be characteristically a kind of set-theoretic or first-order argument. The origin of diagonal argument – with Georg Cantor – resembles more closely a kind of mathematical argument than logical or set-theoretic. Moreover, it is diagonal argument which has consequences for formal systems of logic (and formal set theory) rather than the other way around – diagonal argument is not a feature *of* formal logic so much as it is a (sometimes metalinguistic) mode of argument *about* formal logic. This is seen clearly in consideration of, say, the incompleteness theorems. These theorems are established (in part) by diagonal argument, and since the incompleteness theorems also impose a *limitation* on formal proof (*viz.*, that proof and truth are inconflatable), it follows that diagonal argument shows that consistent formal systems must be weaker (in the *first* place) than we might have hoped. But if all diagonal argument – and in particular, all *bad* diagonal argument – were argument in first-order logic, then why not, on account of its invalidity, simply banish *all* diagonal argument from first-order logic? Similarly, were

56

diagonal argument of a kind that is consequent of the axioms and rules of inference for a formal system, it should not be possible to use diagonal argument to derive antinomy *in the first place* – for these axioms and rules are known to yield consistent closure of the axioms of first order logic. Moreover, we cannot banish diagonal argument by *fiat*, so long as diagonal argument is well-founded in the rules and resources of proof in first-order logic, which it is.

In Simmons' defence, it must be noted that he did generalize upon the *dimensions* of a diagonal argument, but this generalization still affords only a partial account of diagonal argument, since it is already implied in formal set-theory and first-order logic. Thus generalization with respect to the number of dimensions of a diagonal argument still does not bring us to the conclusion that diagonal argument is subsumed within a *calculus rationator*. This constraint upon Simmons' theory is imposed not by an inherent limitation on diagonal argument *in general* (or at least, there is no independent support for the existence of an inherent limitation), but is rather imposed by the conception of diagonal argument *as* formal logic. Simmons' formal definition of n-dimensional (n finite) diagonal argument does not step beyond the contexts of formal set-theory or logic. Thus the generalization which Simmons provides can only be partial.[16]

Moreover, the fact that the form of a *lingua characterica* diagonal argument mirrors that of its formal cousins is no proof that the *lingua characterica* argument is best or most accurately analyzed as if it were a formal language (i.e., a *calculus rationator*). Formal proof does not constrain diagonal argument; rather, diagonal argument imposes constraints on formal proof. In what follows 'diagonal argument' shall mean the

---

[16] We are on the edge of inconsistency here, but not the inconsistency of first-order logic – only that of a *lingua characterica*, of natural language.

formally-unconstrained (*lingua*) kind of diagonal argument. Since the following diagonal

arguments are in the *lingua characterica*, we are free to consider explanations and

conclusions which would render a *calculus rationator* inconsistent.[17]

---

[17] A person might think that 'without the rules of a calculus to tell us when we've got a real consequence
and when we're jumping to conclusions, we can explore, but cannot be all that confident about our results.'
This thought apparently proceeds from the supposition that the rules of a sound *calculus rationator*
argument (i.e., a sound first-order proof) are, or are monotonically preserved, within the rules of argument
in *lingua characterica*. But to make this supposition appears to beg the question, 'is *lingua characterica*
best, most completely, or most accurately analyzed as if it were a formal language (i.e., as a *calculus
rationator*)?' However, if this question *is* begged, and the formal methods (i.e., those of a *calculus*) are
impotent to provide an analysis of the question, then we might still expect an adequate analysis in a *lingua*.

SIMMONS' THEORY OF DIAGONAL ARGUMENT,
AND RUSSELL'S QUESTION AGAIN

In his book, Universality and the Liar [SIM], Keith Simmons offers a general

theory of diagonal argument. Simmons' theory is contained in pages 20 to 44, and the

following material is a paraphrase and simplification of Simmons' theory. What will not

be well-represented in the following exegesis is that Simmons' theory of diagonal

argument is formal – in virtue of its being written in a formal first-order language.

Because Simmons' theory is formal, it is capable of strong results. For example,

Simmons generalizes as to the number of 'dimensions' of a diagonal argument. The

following summary of Simmons' theory includes only such material as is needed to argue

the rest of the essay.

We begin with the following quote from Simmons,

> There are arguments found in various areas of mathematical logic that are
> taken to form a family: the family of *diagonal* arguments. Much of
> recursion theory may be described as a theory of diagonalization; diagonal
> arguments establish basic results of set theory; and they play a central role
> in the proofs of the limitative theorems of Gödel and Tarski. Diagonal
> arguments also give rise to set-theoretic and semantic paradoxes. What do
> these arguments have in common: What makes an argument a diagonal
> argument? And to ask a question first raised by Russell, why do some
> diagonal arguments establish theorems, while others generate paradoxes?
> [SIM, p.20]

In other words, Russell's question – (RQ) – concerns the problem of distinguishing the

'good' diagonal arguments – those which eventually establish theorems – from the

merely 'good-looking' diagonal arguments – those which establish antinomies.

Presently, our usual means of deciding the (RA)s is to observe whether or not a

given diagonal argument proves a falsehood or contradiction. This is part of Simmons'

methodology for recognizing 'bad' diagonal arguments. If a diagonal argument does

prove a falsehood then it must be unsound, for no sound proof proves a falsehood. The only thing we could do with a diagonally-proved falsehood is use the falsehood as an absurdity for a *reductio* argument. Nevertheless, as Simmons himself observes, there is nothing wrong with the reasoning (nor the validity) of a contradiction-proving diagonal argument [SIM, p.29] – for otherwise the argument could hardly be contradiction-*proving*. Rather, it is the *soundness* of one or another *supposition* of the paradoxical diagonal argument which is at fault. For greater generality, it should be said that paradoxical diagonal arguments have *absurd* conclusions, since although every contradiction is absurd, not every absurdity is a contradiction. (Simmons does not note this important fact.) Also, the non-contradictory absurdities are not in all cases object-linguistic, i.e., they are not all *proved* formally. (Of course, some non-contradictory absurdities are formally proved – such as Gödel's proofs that the set of the formally provable sentences is not the same as the set of the true sentences. Such absurdities are generally re-interpreted as having always been non-absurd, as being unexpected facts.) For example, it is a non-contradictory but absurd thesis that the semantic theory of a *lingua* can be done from within that *lingua* itself, independently (i.e., without invoking) the *actual* semantics of that *lingua*, say English. Simmons' strategy of identifying the bad diagonal arguments depends on our already having an independent means – presumably an algorithm of some sort – of identifying absurdities. But we have no such algorithm, for if we did, then (RQ) would be a closed question. Moreover, we should recall that the conclusions of several early diagonal arguments – e.g., those of Cantor, Russell and Gödel – were at first thought to be absurd, but are now counted among the key theorems of formal logic. Since logicians' sense of absurdity has misdirected them

60

in the past, we should not expect contemporary logicians' sense of absurdity to be immune to such misdirection.

Let us now delve into the technical points of Simmons' theory of diagonal argument. As has already been noted, a diagonal argument can have any finite number of dimensions, but we will be concerned only with the two-dimensional cases. The main components of a diagonal argument are (according to Simmons) the SIDE, the TOP, the VALUES, the ARRAY, the DIAGONAL, the *value of the diagonal* (hereafter, D.VAL), the *countervalue of the diagonal* (hereafter, C.VAL) and the ROWs of the ARRAY. Simmons defines each of these components in a first-order language of set theory, but we will specify them according to the ARRAYs pictured below.

Consider (A1).

	t0	t1	t2	t3	...
s0 |	m	w	m	w	
s1 |	w	w	m	w	
s2 |	w	w	w	m	
s3 |	m	m	w	m	
. |					
. |					

(A1)

The SIDE is the set, {s0, s1, s2, s3, ...}; the TOP is {t0, t1, t2, t3, ...}; and VALUES is {m, w}. The ARRAY of (A1) is the field of the symbols 'm' and 'w'. It will be helpful to define two sub-parts of the ARRAY, namely, the ROWs and the CELLs. A CELL of an ARRAY is an ordered pair, $\langle si, tj \rangle$, such that $si \in$ SIDE and $tj \in$ TOP. Thus the set of all the CELLs of (A1) is the set of all $\langle si, tj \rangle$ such that $i,j \in$ NN. (We read for 'NN' the set of the natural numbers.) For the value ('m' or 'w') of the CELL $\langle si, tj \rangle$, we write, $\langle\langle i, j \rangle\rangle$. Thus $\langle\langle 2,0 \rangle\rangle$ = w and $\langle\langle 0,2 \rangle\rangle$ = m. (Simmons does not define the CELLs, but he might have done so, and perhaps as set out above.) For the kth ROW (hereafter, ROW$_k$) of (A1), we

61

write, «*k*,0» «*k*,1» «*k*,2» «*k*,3» ... The ARRAY of (A1) is now just the set of all its ROWs: {ROW₀, ROW₁, ROW₂, ROW₃, ...}.

The next two definitions are at the heart of a diagonal argument: D.VAL and C.VAL. D.VAL is the value of the 'leading' diagonal of (A1), which starts in the upper left and continues to the lower right. In general, the D.VAL of an ARRAY is «0,0» «1,1» «2,2» «3,3» ... Thus the D.VAL of (A1) is mwwm... The C.VAL of (A1) is the reverse of D.VAL, or wmmw... (For convenience, we might also write D.VAL as «d,0» «d,1» «d,2» ... and C.VAL as «c,0» «c,1» «c,2» ...) In general, we define the C.VAL as follows.

(DCV) For all natural numbers, *n*,

$$\text{«c,}n\text{»} = \begin{cases} m & \text{if «d,}n\text{»} = w, \text{ (that is, if «}n,n\text{» = w),} \\ & \text{and,} \\ w & \text{otherwise.} \end{cases}$$

(For 'DCV' read, 'definition of C.VAL'.) Next, Simmons specifies what it means for D.VAL or C.VAL to occur as a ROW of an ARRAY. D.VAL occurs in (A1) as a ROW if and only if there is a natural number, *n* such that ROW$_n$ = D.VAL. C.VAL occurs as a ROW of (A1) if and only if there is a natural number, *n* such that ROW$_n$ = C.VAL.

We are now prepared to prove what Simmons calls the *basic diagonal theorem*. The proof below is not formal, and in any case is only outlined. Nevertheless, the reasoning of the argument below is valid, and is formally proved by Simmons [SIM, pp.25-26, *Theorem 2.1*].

> Suppose that C.VAL occurs as a ROW of (A1). Then there is a natural number, *n*, such that ROW$_n$ = C.VAL. (Thus «c,*n*» = «*n*,*n*».) If «*n*,*n*» = m, then by (DCV), «c,*n*» = w. Similarly, if «*n*,*n*» = w, then «c,*n*» = m.

But if $\langle\langle n,n\rangle\rangle = m$, then as $\langle\langle n,n\rangle\rangle = \langle\langle c,n\rangle\rangle$, we have $m = \langle\langle n,n\rangle\rangle = \langle\langle c,n\rangle\rangle = w$. And thus, $m = w$, which is false. If otherwise $\langle\langle n,n\rangle\rangle = w$, then we have $w = \langle\langle n,n\rangle\rangle = \langle\langle c,n\rangle\rangle = m$, and $w = m$, which is the same contradiction. Thus there is no n such that $\mathrm{ROW}_n = \mathrm{C.VAL}$.

As already been mentioned above , Simmons distinguishes 'good' diagonal arguments from 'bad' diagonal arguments according to whether they prove theorems or establish paradoxes (respectively). Of course, the basic diagonal theorem is a 'good' diagonal argument, but could be either of two kinds of 'good' diagonal argument: the 'direct' and the 'indirect'. (Jonathan Seldin has commented that every indirect diagonal argument can be equivalently expressed as a direct diagonal argument, but this need not invalidate Simmons' analysis.)

According to Simmons, in a direct diagonal argument, the SIDE, the TOP and the ARRAY are all 'well-determined' sets, and D.VAL and C.VAL are well-determined in the sense that every 'digit' – m or w – of these values can be effectively determined. The conclusion of a direct diagonal argument is simply that the C.VAL does not occur as a ROW.

In an indirect diagonal argument, one (or possibly more) of SIDE, TOP and ARRAY are *assumed* to be well-determined, for *reductio ad absurdum*. In (A1), SIDE is specified as a denumerable set of the s_i, such that each s_j is paired uniquely with ROW_j. Accordingly, we might say that s_j is the 'name' of ROW_j. To treat (A1) as though it were an indirect diagonal argument, we might assume (for *reductio*) that even though SIDE is denumerable, its members collectively name each and every infinite sequence of 'm's and 'w's. We assume, that is, that every possible value for a ROW of the ARRAY is

63

named by some member of SIDE. But SIDE is not well-determined, for the C.VAL is not named by any member of SIDE. Thus there is always a possible ROW value without a name.

The 'bad' diagonal arguments are neither direct nor indirect. In a 'bad' diagonal argument, the components are all assumed to 'exist' (be well-determined), but one or more of them is actually not well-determined. Besides finding the presence or absence of unsound assumptions, Simmons presents no other means of identifying a 'bad' diagonal argument. He states also that "there is nothing wrong with the *reasoning* of a bad diagonal argument – rather, what is at fault is the assumption that all of the diagonal components exist" [SIM, p.29]. Moreover, it is a simple matter to convert a 'bad' diagonal argument into a 'good' one. A 'bad' diagonal argument can be converted into a direct or an indirect diagonal argument. To make it direct, replace the non-well-determined component with one that is well-determined. To make it indirect, convert the unsound assumption into an assumption for later negation by *reductio ad absurdum*.

Following the material on 'good' and 'bad' diagonal arguments, Simmons presents and analyzes many examples for discussion. We will consider a few of them in turn. Gödel's first incompleteness theorem[18] is analyzed as a direct diagonal argument. It is assumed that every predicate (what Gödel calls a 'class sign') defines the set of numbers of which the predicate is true. In particular, the unprovability predicate, S, is assumed to be decidable (i.e. S is assumed to be provable of exactly those naturals of which it is true). (In general, a decidable predicate, $S(x)$, is one such that for all natural numbers, x, either $S(x)$ |-/- $\perp$ or ~$S(x)$ |-/- $\perp$.) S, along with the numbers of which it is

---

[18] Strictly speaking, Simmons does not argue via Gödel's incompleteness *theorems*, but rather via the informal outline given at the beginning of Gödel's 1931 paper [GOD1].

true (or provable), functions very much as a countervalue. Under the constraint of *omega*-consistency, it follows that neither the Gödel sentence, '[S;m]', nor its negation, '~[S;m]' is provable. Thus S is not formally decidable in the object-language, although the same proof (i.e., Gödel's incompleteness proof) supports the meta-linguistic decision that the Gödel sentence is true.

In Simmons' theory the set of all non-self-membered sets (the Russell set) is a countervalue for a bad diagonal argument. The unsound assumption is that the Russell set exists (i.e., that it is well-determined), and the conclusion by *reductio* is that this set does not exist. Simmons also considers a few more classic examples of diagonal argument. These include the unsolvability of the halting problem (which Simmons says is proved by a good diagonal argument); Georg Cantor's paradox of the universal set (a bad diagonal argument which is convertible to a good indirect diagonal argument); Cantor's powerset theorem (a good indirect diagonal argument); and the formal indefinability of the (informal) notion of *total algorithmic function* (a direct diagonal argument).

Next Simmons argues an extension of his analysis of formal diagonal arguments to the informal cases of the semantic (i.e., natural language *lingua characterica*) paradoxes. It is via this extension that Simmons makes the (somewhat trifling) errors that I wish to highlight. Simmons' running example for these cases is the heterological paradox. (It is assumed that the reader is familiar with the basic form of this paradox. Here the SIDE is the set of the extensions of the English predicates.) These extensions are sets of individuals. TOP is the set of the predicate-signs of English; SIDE lists the domain, and the VALUES of the ARRAY – either, say, '0' or '1' – express set

membership.) Without significant discussion of the matter, Simmons proceeds to analyze, and eventually to solve, the heterological paradox via formal resources [SIM, pp.35-37]. In particular, Simmons invokes Tarski's indefinability theorem to explain heterologicality. However, it appears to me that the soundness of Simmons' assumption (that natural language heterologicality is accurately modeled via formal resources) is suspect. I hold this suspicion notwithstanding the fact that (as Simmons cites) Tarski weakly supports Simmons' assumption.

Simmons considers the following as the possible treatments of the heterological paradox. *First*, we might implement a Tarskian hierarchy of languages. We thereby replace the general predicate, 'heterological' with a denumerable hierarchy of heterological predicates: 'heterological$_0$', 'heterological$_1$', 'heterological$_2$', and so on, such that neither 'heterological$_n$' nor its negation is a member of the set of the heterological$_m$ predicate-signs (for all natural numbers, n and m, such that $m \leq n$). (That is, the predicate-*sign*, 'heterological$_n$' is not a member of the extension of the *predicate* heterological$_m$.) *Second*, we could deny that the heterological predicate has an extension, since if it did have extension, then that extension violates the law of the excluded middle. *Third*, it could be argued that although there is a concept of heterologicality *for* and *in* English, no English predicate *expresses* this concept.

The second diagnosis is the closest to that which I advance in this essay. Commonly, from a violation of the law of non-contradiction one reasons via 'truth-gaps' or 'truth-gluts'. Thus when asked whether 'heterological' is heterological, we might answer, 'yes and no', or 'neither yes nor no'. Our account leans toward the notion of a truth-glut. However, it is not a glut of *truth-values* that we will employ, but rather a glut

66

of *predicate-extensions*. This kind of glut is easy to explain by distinguishing it from another possible notion of *predicate-extension glut*. For example, we might have a predicate-extension glut for a given predicate, P, by having (for some individual, $\lceil Q \rceil$) both $\lceil Q \rceil \in P$ and $\lceil \sim Q \rceil \in P$. The problem with this kind of glut is that it results in inconsistent predicate-extensions, via which contradiction may be easily derived. The alternative notion of a predicate-extension glut is to have *multiple extensions*. We might hold that the extensions of P are both $\{\lceil Q_0 \rceil, \lceil Q_1 \rceil, \lceil Q_2 \rceil, ... \lceil Q_n \rceil, ...\}$ and $\{\lceil Q_0 \rceil, \lceil Q_1 \rceil, \lceil Q_2 \rceil, ... \lceil \sim Q_n \rceil, ...\}$. But for the time being, we put this notion of an extension glut aside, and return brefly to Simmons' first and third diagnoses.

Simmons' first and third diagnoses have the effect of imposing (or discovering) expressive limitations on natural language. In the first case, the general predicate 'heterological' is eliminated in favour of a hierarchy of heterologicality predicates. (Of course, Tarski hierarchies are not unintuitive for mathematical logicians, but for present purposes, the more relevant opinion is that of ordinary people who know no formal logic.) The third diagnosis is even more unintuitive because it proposes that 'heterological' is an inexpressable predicate, whereas this predicate *is* expressed by speakers of English, even though it cannot be consistently decidable. So the third diagnosis goes to deny a plain fact – that the speakers of natural language quite routinely use terms such as 'heterological', and that they *successfully convey* a meaning by the use of these terms. However, to be fair to the formal treatment of heterological predicates, it must be noted that the heterological paradox is not paradoxical without the supposition that there is one heterological predicate which applies to *all* possible cases of the predicate's use. But there is no compelling reason to think that this supposition is true,

for neither the formal heterological predicate, nor the natural. That is, it might be the case – in the first place – that heterology cannot be correctly modeled as if it were a single predicate which applies (or does not apply) universally to all predicates and/or words. There is confirmation of this view – that there is no universal heterological predicate – in Herzberger's [HER1] where he considers many heterological predicates which have successively more and more restricted cases of application.

It could be argued that the first and third diagnoses are effective explanations of, and suggest resolutions for, natural language predicates such as 'heterological'. But such an argument appears to depend on circular reasoning, since the principal reason for accepting this argument is just that it disarms the paradox of whether 'heterological' is heterological. There is no independent reason to believe that the only way to disarm the heterological paradox is along the lines of Simmons' first and third diagnoses, that is, along the lines that there is no universally applicable natural language heterological predicate. It is no solution of the paradoxes of natural language to assert that natural language is reducible to first-order logic, whether or not such assertions are supported by the supposition or discovery that natural language heterology predicates are not universally applicable. It cannot be the case that formal analysis alone proves that natural language is, or is subsumed within, the *calculus rationator* of first-order logic. It is apparent that mathematical logic is a strong and highly general means of modelling natural language, but even if such models enjoy stellar success, it does not follow that natural language *is* first-order logic. To put the matter differently, we have no proof that the best *possible* explanation of natural language paradox is the one provided by the *calculus* of first-order logic, even though it certainly provides the best *existing* account of

natural language. Even if parts of natural language can be successfully modelled separately from other parts, that is, by multiple mutually distinct formal models, the *full* (i.e., *complete*) modelling of natural language, by a *single* formal model may be too problematic to hope for. At the present, there is no such full model, although there is also no cause to believe that a full formal model of natural language is altogether impossible.

Unfortunately, Simmons does not engage these considerations, and passes by the above challenge, that natural language is best understood as a system of first-order logic. Simmons' clearest support for the view that natural language is first-order logic is in the following single sentence. "These semantic paradoxes [i.e., the paradoxes of natural language] are generated by diagonal arguments in exactly the same way that the set-theoretic paradoxes are generated. [SIM, p.37]" We can grant Simmons this claim without granting the unstated implication that the semantic paradoxes *are* set-theoretic. Nevertheless, it remains that natural language diagonal paradox is substantively 'parallel' with (i.e., *similar to*) the diagonal paradoxes of first-order logic. Accordingly, the resolutions of these kinds of paradox may also be substantively similar – it is possible (at least) that these kinds of paradox and their resolutions are *identical*.

I have three criticisms of Simmons' theory of diagonal argument: (a), (b) and (c). (a) Simmons considers only those diagonal arguments which are formulable in (ZF) set theory or first-order logic. Thus, for Simmons to analyze natural language diagonal arguments via his theory of diagonal argument, he must assume that natural language diagonal arguments can be soundly and completely represented via set-theoretic resources. As has already been noted, this assumption may be sound, and at present, this

assumption provides the best existing theory of natural language and its paradoxes. Accordingly, Simmons' assumption is very well supported, for it is *instrumental* to the best existing theory of natural language. Moreover, Simmons' assumption does not commit us to the view that natural language *actually is*, or *actually reduces to*, the *calculus* of first-order logic. These *calculus rationator* models of natural language can serve as *experimental confirmation* of the thesis that natural language actually is *calculus rationator*, but not *proof*. Of course, such proof is not needed to justify the pursuit of *calculus rationator* theories of natural language, for all that Simmons must be committed to is the representability of natural language via a set-theoretic *calculus*. However the fact that there is (presently) no proof that natural language is a *calculus* leaves the question (infinitesimally) open, 'is natural language *lingua* actually formal *calculus*?' Before this question can be answered, there must be developed some successful non-*calculus* account of natural language, or proof (of some semantically universal non-*calculus* kind) that no such account can succeed. And it is unlikely that with existing analytic tools any account of natural language – whether via *calculus* or not – could be proved to be the same thing as natural language. *A fortiori*, it is unlikely that any professional logician would attempt to prove or disprove the claim that natural language actually is the same thing as, or is adequately and fully modelled by, this-or-that *calculus* or non-*calculus* account of natural language.

(b) As has already been noted, Simmons does not distinguish between absurdity and contradiction. But this is an important distinction, as *all* of the (RA)s are absurdities but not contradictions. Those diagonal arguments which conclude with an (RA) are highly interesting, whereas those diagonal arguments which conclude with contradictions

70

are plainly unsound (i.e., if they are not treated by a dialethic system of paraconsistent logic) and thus (except for the job of figuring out how the proofs of these contradictions have gone wrong), they are also uninteresting. It is necessary for any *general* theory of diagonal argument to explain *all* diagonal argument, specifically including diagonal argument which concludes with an (RA).

(c) For Simmons, one of the marks of a 'bad' diagonal argument is that it assumes the existence of a non-well-determined set (TOP, SIDE or ARRAY), but not toward a *reductio* argument. However, Simmons does not adequately explain the notion of 'well-determinedness'. It is implied in several places that a non-well-determined set does not *exist*. Otherwise, Simmons only suggests that we adopt the notion of well-determinedness from, say, ZF set theory. It is hard to resist the impression that Simmons wants to hold that all non-well-determined sets are altogether trivial and uninteresting. Whereas it makes sense to say that no non-well-determined sets exist *in ZF*, it does not immediately follow that no non-well-determined sets exist *at all*. Even if it actually is the case that no *existing* set is non-well-determined, there are other rigorous notions of sethood, such as 'non-well-founded' [ACZ2] and 'fuzzy' sethood (*see* [DID], [KLI] and [ROS]). Not every fuzzy or non-well-founded set is well-determined. Thus well-determinedness is not a requisite of all interesting or rigorously considerable sets. It would be helpful here to have the necessary and sufficient conditions of well-determinedness, but Simmons does not state these conditions. Presumably, a set is well-determined if and only if its membership is well-determined. In turn, a set's membership is well-determined if and only if for every object, a, of the domain, a is either determinately *in* or *out* (but not neither and not both) of the set. When a set is said to

71

'exist', this properly means that the set's membership is well-determined. (As an aside, it should be noted that when we say that this-or-that set *exists*, we use this word in a semantically restricted sense, such that the applicability of the word – 'exists' – is not universal. A well-determined set exists, but not everything which exists is a set. (For instance, cats and dogs exist.) And not everything which does not exist is a non-well-determined set. (For example, unicorns do not exist, but are not sets of any kind.)

But fuzzy sets do not satisfy even this presumed notion of well-determinedness. An object, *w*, might be '75% in' the fuzzy set, *Z*. In this case, *w* is '25% out' of *Z*. Thus *w* is both in and out of *Z* (though to different degrees), contrary to the above presumed notion of well-determinedness. In consequence of this, the fuzzy notion of sethood appears to be independent of the notion of well-determinedness, for otherwise, fuzzy sethood is a trivially inconsistent notion. Let alone the fact that not every fuzzy set is *the same thing as*, nor even the same *kind* of thing as, some one or many well-determined sets, it is not even the case that fuzzy sethood can be consistently described or modeled via the notion of well-determinedness, for otherwise it must again be the case that fuzzy sethood is an altogether incoherent notion. So there are (there *exist*) rigorous non-trivial notions of sethood which are not compatible with the notion of well-determinedness, such as the notions of non-well-founded and fuzzy sethood.

WELL-DETERMINEDNESS AND WELL-DEFINEDNESS

Simmons adds to the puzzle of the meaning of 'well-determined' when he considers the paradox of Jules Richard [RIC]. I will not reproduce Richard's paradox here, as it is likely to be well-known by the readers, and in any event, the details of Richard's paradox are not relevant at this point. What we will focus on is Richard's own diagnosis of his paradox – that the set (which Richard labels, 'E') of those numbers which are defined by an expression of French is *non-well-defined*. (For Simmons, if a set is non-well-defined, then this is sufficient to conclude that it is non-well-determined.) According to Simmons, the problem with Richard's set of definitions, E, is that 'we lack a clear notion of *definability*' [SIM, p.28], but this is a somewhat abstracted diagnosis of the paradox. A closer diagnosis is to say that *under the supposition that E is enumerable, E is inconsistent.*

To show that E's enumerability implies a contradiction, we reason as follows. First, we set the TOP to be the digit-places of the numbers defined by the members of E. (TOP will be the set of the natural numbers.) SIDE is the set E, and the ROWs of the ARRAY are the numbers defined. Now the French expression, G (the counter-diagonal value), is a member of E, but the number defined by G cannot occur at any ROW of the ARRAY. (I omit proof of this claim.) Otherwise, if the number defined by G is stipulated to occur at some ROW, then the value (number) of that ROW is inconsistent – either the number paired with definition G is not the number defined by G, or else the number defined by G is not the counter-diagonal of the ARRAY.

One might be tempted to 'solve' this paradox by stipulating that G is not a member of E. This solution would be endorsed by Simmons – it is given by an indirect

diagonal argument which assumes for *reductio* that G is a member of E. However, the more economical solution appears to be to hold that the set of the numbers defined by the members of E is non-enumerable. According to this solution, none of the ROWs of the ARRAY has the number defined by G. Nevertheless, G is an expression of French, and it can certainly be enumerated among the other members of E. But since the number defined by G can never occur as a ROW, it follows that the set of the numbers defined by the members of E is non-enumerable, for there must always be an 'extra' number, for we are always able to effectively define a counter-diagonal number. (It is noteworthy that the numeral-naming resources of French are significantly limited. However, as a natural language, French is *expansible* from within the French language in such a way that a French speaker will understand the expanded numeral-naming resources.)

Unless it is argued that enumerability is a necessary condition on well-determinedness, we may not conclude that the set of numbers defined by the members of E is non-well-determined. However, the ZF-notion of well-determinedness is not the only interesting one, and as has already been noted, the non-well-determinedness of E does not imply that E is altogether non-existent. All that must be granted is that the set of the numbers defined by the members of E is non-enumerable, and since there are many well-known non-enumerable sets (such as the set of the real numbers, or of the powerset of the natural numbers), it is no vitiation if we discover another. It remains as a vexing puzzle that although E is enumerable, and supposing that each member of E defines a single determinable number, the set of the defined numbers is non-enumerable. Ordinarily, the non-enumerability of a set amounts to its having uncountably many members, but I think that in the case of the set of numbers defined by E we should resist

74

this additional conclusion. In the first place, no expression could legitimately be a *definition* of a (single) number if it defined either *no* number or *more than one* number. Thus, under the supposition that every member of E defines a single number, the number of numbers defined by the members of E can be no greater than the number of members of E. In the second place, constructivist logicians will not grant that the non-enumerability of a set implies that it *actually* has uncountably many members. Constructivist logicians will grant that there is no one-to-one onto function from the naturals to the set of numbers defined by E, but they will not grant that the non-existence of this function implies that there are more numbers defined by E than there are members of E.

It appears to me that the *real* problem with E is not that it is non-well-determined. For set-wise determinedness is a matter of set membership, and more specifically of the membership relation ('∈'). But the membership of E is not in question until we reason upon it that a finite expression of English (for Richard, *French*) *actually fails* to define a number. When we do reason that G defines no number, we do not seem to be reasoning upon the set E, but upon a different set, say, E'. G clearly defines a counter-diagonal number (which Richard labels, 'N') for E, notwithstanding the fact that N could not be a member of, and simultaneously a counter-diagonal for, E. But when N is *added* to E, we get a different set of numbers, E'. G defines N and N is a counter diagonal for E, and so N is not a member of E. We may argue that since G *does* define a number (namely, N), G must be a member of all of the finite expressions which define a number. But when we add N to E, we obtain a new set of numbers, E' (such that E' = E∪{N}). There is still a finite expression (say, G') which defines a counter-diagonal for E', say N'. Now, N' is

75

not a member of E', even though G' clearly defines N'. If we now add N' to E' we obtain yet another set of numbers (say E''), for which there is yet another finite expression (say G'') which defines yet another counter diagonal number (say N''). We could continue with this reasoning *ad infinitum*, always failing to include the counter-diagonal number (N^*) in the set of numbers (E^*) in the terms of which N^* is defined. It appears that what makes Richard's paradox a paradox is just that in this (above) way, E fails to pick out all the numbers which its members define.

It is therefore clear that no ARRAY can include all the numbers definable by a finite expression, but this is not quite the same as saying that the set of the defined members is non-well-determined – for any set, E, of the finite expressions which define a number, there is a set N of the defined numbers, and the membership of N is easy to specify. (That is, since the number defined by any given member of E can be specified digit-by-digit, that number can be exactly specified.)

The discussion above suggests a further comment concerning sets such as E and N, and expressions such as G. Note that G is self-referential, in the sense that the number defined by G is specified in terms of each member of N. Under the supposition that G ∈ E, the number defined by G is one of the members of N. Now, we let r be the row number at which the number defined by G (let this number be g) appears. Then the value of rth row is simultaneously the value of the counter-diagonal of N. So the rth digit of g is defined in terms of the rth digit of the number at row r. Thus, as (by hypothesis) $g = r$, it follows that the rth digit of g is defined in terms of the rth digit of g. In this way, G is self-referential.

Observe now that, according to the expression G, the kth digit of g is different from the kth digit of the (number at the) k^{th} row of the ARRAY of N. So it is impossible for the rth digit of g to be the same as the rth digit of the rth row of the ARRAY given by (the enumeration of) N. Accordingly, the (number at the) r^{th} row of the ARRAY cannot simultaneously be g. Thus, the expressive capacity for self-reference carries (or seems to carry) enough expressive power to allow us to reach beyond the limits of sets like N – that is, to reach beyond the membership of sets like N. No matter what the membership of N may be, we may yet identify numbers which are not members of N. In particular, we may always reach beyond the membership of *well-determined* sets, which *must* have fixed memberships. This notion of 'reaching beyond' the membership of well-determined set has an intuitive counterpart in an informal problem about the size of the universe. Supposing that the universe is bounded (as are well-determined sets), we may yet coherently ask, 'what lies beyond these boundaries?' The answer may be 'nothing', but then we might simply propel something past those boundaries – say, a misanthropic executive of some insurance agency – and then the answer to the question could be, 'a just place for an insurance executive to be'. (This appears to suppose that beyond the boundaries, there is *empty space* for the executive to fill, but even if there is no empty space beyond the universe's boundaries, and there is *absolutely* nothing – no space, no existence and no possibility – then this would seem to be an even better place for the executive to be.) But the problem remains, for we have not freed ourselves from the executive; by propelling him into his just place, we have only pushed the boundaries of the universe outward, and made it a little bigger than it was before – for now there is *something* beyond what were the boundaries of the universe.

77

CURRY SENTENCES, FIXED POINTS AND UNDECIDABLITY

It is Gödel sentences which are the better known examples of fixed points, and so one might wonder why I have chosen to reason according to Curry sentences instead. As neither Gödel sentences nor provability predicates are essential to the arguments I want to make, it is actually advantageous to exclude them from our consideration. The problem with Gödel sentences is that they are *so* well-known that, in using them, a writer can easily become distracted or engulfed by the voluminous literature surrounding Gödel's 1931 results. This literature runs the range from philosophical explorations, to technical outlines, to strictly formal proofs and extensions. The incompleteness theorems also have implications for mathematics, computer science, artificial intelligence, computational linguistics, and generally any formal theory in which the formal arithmetic is a sub-theory – that is, virtually *all* formal theories.

The following treatment of Curry sentences and predicates is adopted from the treatment given to fixed points by Boolos and Jeffrey in their [BOO2, pp.170-172]. (For Curry's own treatment of this paradox via the λ-calculus, see his short 1942 [CUR2].)

Curry sentences, (CS), are formulas, C, such that,

(CS) $C = {}^{\backprime}C \rightarrow Q{}^{\prime}$,

for an arbitrarily chosen sentence, Q. It is a simple matter to derive a contradiction from (CS).[19]

---

[19] Curry's original version of his paradox runs as follows.

1.	$C \rightarrow C$	[Postulated tautology]
2.	$C \rightarrow (C \rightarrow Q)$	[1, Rule of equality]
3.	$(C \rightarrow (C \rightarrow Q)) \rightarrow (C \rightarrow Q)$	[Postulated tautology]
4.	$C \rightarrow Q$	[2, 3, *modus ponens*]
5.	C	[4, Rule of equality]
6.	Q	[4,5, *modus ponens*]

78

(CP)   Let 'Q' = '$\sim C$'. Thus by (CS), 'C' = '$C \to \sim C$'. (*) Suppose $\vdash$ C. Then $\vdash (C \to \sim C)$. By *modus ponens*, $\vdash \sim C$. So if $\vdash C$ then $\vdash \sim C$ – and thus $\vdash (C \to \sim C)$, by conditional proof. So $\vdash C$. (The supposition at (*) is discharged.) And then $\vdash \sim C$, whence contradiction.

(Read for (CP), 'Curry's paradox'.)

Now we show that if L is consistent, then C is undecidable.

Suppose $\vdash C$. Then by (CP), $\vdash \sim C$, whence contradiction. So if L is consistent, then $\nvdash C$.

Suppose $\vdash \sim C$. Then by (CP), $\vdash \sim\sim C$, whence contradiction. So if L is consistent, then $\nvdash \sim C$.

Therefore, if L is consistent, then $\nvdash C$ and $\nvdash \sim C$. *A fortiori*, either L is inconsistent or C is an undecidable formula.

Below, much is made of the fact that Curry sentences are diagonal expressions, in the sense that Curry sentences are (or are expressable as) fixed points. By arithmetizing L, we gain Gödel quotes $\lceil ... \rceil$ and disquotes $\lfloor ... \rfloor$. We define a fixed point on the predicate,

(1) $\lfloor x \rfloor \to Q$,

as a natural number, n, such that,

(2a) $n = \lceil \lfloor n \rfloor \to Q \rceil$.

Let there now be a predicate, $C(x)$, which might be called a 'Curry predicate', such that $\lceil C(x) \rceil = \lceil \lfloor x \rfloor \to Q \rceil$. Therefore,

(2b) $\lceil C(n) \rceil = n$,

and,

(2c) $n = \lceil C(n) \to Q \rceil$.

It is assumed that there is a fixed point theorem for Curry predicates, and that via this theorem, the numeral of n can be effectively computed.

79

Note that Curry predicates appear to be distinct from the ordinary predicates. Ordinary predicates can be understood as functions which take objects (of the domain of L) to truth letters (i.e., to 't' or 'f'). By contrast, Curry predicates appear more as functions from formulas to formulas. (As a side comment, it seems likely that this difference in kinds of predicates is connected to the interesting fact that in Church's λ-calculus and Curry's combinatoric logic, Curry predicates can be expressed without the use of Gödel-quotes.)

Curry predicates fall within a more general kind of predicate, which we will call *substitution predicates*. A substitution predicate is any L-formula, R, which is open at the place of a sub-formula of R, say S. For example, let R be the formula, '$A \lor (B \land C)$'. If we open R at the place of the sub-formula, B, then we get a substitution predicate, R': '$A \lor (\lfloor x \rfloor \land C)$'. (We take x to range over the natural numbers – i.e., those natural numbers which Gödel-code L-formulas.)

Let n be the Gödel number of the L-formula T. Then $R'(n) =$ '$A \lor (T \land C)$'. Provided that L has the resources to arithmetize its own syntax, any substitution predicate may be expressed in the object language. In the above example, $R'(n)$ is an object-linguistic predicate which is logically equivalent to the formulas '$A \lor (\lfloor n \rfloor \land C)$' and '$A \lor (T \land C)$'. That is,

$$R'(n) \quad \leftrightarrow (A \lor (\lfloor n \rfloor \land C))$$
$$\leftrightarrow (A \lor (T \land C)).$$

Because substitution predicates are expressable in the object language, the distinction between them and ordinary predicates need not be troublesome. As long as L has Gödel

quotation, all of the substitution predicates are expressed by some of the ordinary predicates.

We now proceed to construct an ARRAY in which Curry sentences will be analyzed.

```
(A2)        f0  f1  f2  f3  . . .
      s0 |  1   0   1   1  . . .
      s1 |  0   0   1   0  . . .
      s2 |  1   1   0   1  . . .
      s3 |  0   1   0   1  . . .
       . |  .   .   .   .
       . |  .   .   .   .
```

SIDE = *the set of all substitution predicates in L*
= { s0, s1, s2, s3, ... }

TOP = *the set of all formulas of L*
= { f0, f1, f2, f3, ...}

CELLS = { ⟨0,0⟩, ⟨0,1⟩, ⟨1,0⟩, ⟨0,2⟩, ⟨1,1⟩, ...}
= *the set of all ordered pairs of natural numbers*

VALUES = {0, 1}

We interpret the values of (A2) as follows. For all naturals i and j, «i,j» = 1 if and only if $si(\lceil fj \rceil)$ is true, and otherwise, «i,j» = 0. We also require that the values of (A2) obey the law of the excluded middle. Accordingly, the rows of the values of (A2) express the *extensions* of the members of SIDE. Since the values of (A2) (i.e., at the cells of (A2)) also express truth, the array expresses truth as if it were extensional. Thus (A2) can be understood as a test of the notion that truth is extensional.

It is worthwhile to note that it is already well-known that truth-in-(the object language of)-*L* is not extensional, leastways, not wherever the syntax of *L* can be arithmetized. (Truth-in-L *is* extensional in the metalanguage – provided that 'true' is not also an object-linguistic predicate.) It amounts to much the same to say that truth-in-*L* is

81

not extensional if L has the resources necessary for Gödel-qotation. It is furthermore well-known that no truth predicate for L is decidable in L, in consequence of Tarski's theorem of the undefinability of truth [TAR1]. And the undecidability of truth predicates is tantamount to their not being extensional. (In the present example, we expressly omit truth predicates and provability predicates from SIDE.) Yet (A2) goes to show the non-extensionality of some members of SIDE in a different – and hopefully novel – way.

For convenience, we define the rows of (A2) as follows.

For all natural numbers, i,

$$\text{ROW}i = \text{‘}«i,0» «i,1» «i,2» «i,3» \dots\text{’}.$$

We are now sufficiently prepared to form diagonal arguments in terms of (A2). The simplest such diagonal argument on (A2) is exemplified by Georg Cantor's famous proof of the uncountability of the real numbers. For this argument, we need specifications of the value of the 'leading' diagonal (i.e., the diagonal from the top left cell of (A2) toward the bottom right), and its countervalue. D.VAL is the value of the leading diagonal, and C.VAL is the counter value.

$$
\begin{aligned}
\text{D.VAL} \quad &= \text{‘}«0,0» «1,1» «2,2» «3,3» \dots\text{’} \\
&= \text{‘}1\ 0\ 0\ 1 \dots\text{’}
\end{aligned}
$$

$$
\begin{aligned}
\text{C.VAL} \quad &= \text{‘}0\text{'}\ at\ digit\text{-}place\ x\ if\ «x,x» = \text{‘}1\text{’},\ and\ \text{‘}1\text{’}\ otherwise \\
&= \text{‘}0\ 1\ 1\ 0 \dots\text{’}
\end{aligned}
$$

(We write for the substitution predicates for D.VAL and C.VAL, 'sd' and 'sc' respectively.) We have already considered diagonal arguments of the form of (A2) and so we state only an outline of the present diagonal argument. We assume for *reductio* that SIDE includes all of the substitution predicates in L. But this cannot be so, since if it were, there must be some natural, n, such that $\text{ROW}n = \text{C.VAL}$. If $\text{ROW}n = \text{C.VAL}$,

82

then «n,n» = 0 if and only if «n,n» = 1. But this is a contradiction, and therefore, there is no such n, and *a fortiori*, sc is not a member of SIDE.

To set the point aside, the above diagonal argument can be taken to show that there are uncountably many possible extensions for the substitution predicates of L. However, SIDE is countable, since every substitution predicate is a finite string of the basic symbols of L. These strings can then be ordered serially, first according to the length of a string, and then alphabetically within each length. (For convenience, we omit consideration of all those strings which do not express a substitution predicate.) In particular, sc is a finite string of the basic symbols of L, and thus it is expressed at some (finite) place in the serial ordering of the substitution predicates. Since every substitution predicate occurs at some finite place from the beginning of the list, the set of substitution predicates (i.e., SIDE) is countable. But it remains that the extension of sc cannot consistently appear as a ROW.

In systems of mathematical logic, problems of this kind – i.e., those of counter-diagonals such as sc – are already solved. In the solution via mathematical logic, sc is said to be undecidable or undefinable. The extensions of an undefinable predicate then pose no difficulty, because they are not accessible to mathematical logic, that is, because the language does not have enough predicate signs to name every possible predicate extension. In a loose sense, then, mathematical logic solves the problem of undefinable predicates via the observation that not every extension is picked out by some substitution predicate – that is, by observing that not every possible extension is definable, nor even *nameable*, by some substitution predicate. Yet mathematical logic's treatment of undefinable predicates preserves consistency while still providing a high degree of

expressability. And this is strong confirmation of the correctness of mathematical logic's treatment.

LEIBNIZIAN LANGUAGE

The Leibnizian notion of language, developed below, is not drawn from Leibniz'

logical work, but rather from his Discourse on Metaphysics (hereafter simply 'the

Discourse'). The logic of the Discourse is furthermore not easy to uncover, for even if

Leibniz did have a clear conception of his metaphysical logic, he certainly did not have

the tools of contemporary formal logic. Without these tools, Leibniz can do little more

than suggest the logical system of the Discourse. (The main tools which Leibniz lacks

are those of set theory and the diagonal method of proof. The bulk of the other

contemporary tools were developed in the context of Frege's Begriffsschrift, including

quantification theory, recursion theory and proof theory. But it is only the diagonal

method which is relevant to the following discussion.) Because Leibniz did not know of

diagonal proof, it might appear that he could hardly have had a clear notion of the

Leibnizian notion of logic. But it is not at all unlikely that Leibniz could conceive of

Leibnizian language. Leibniz was a highly talented deductive thinker – he created the

infinitesimal calculus independently of Newton, and it is Leibniz' symbolism which is

still used, rather than Newton's. Leibniz also conceived of actual infinities at a time

when most thinkers could conceive only potential infinities. (Contemporary

constructivist logicians still deny the existence of actual infinities, and there is good cause

for this denial. But the constructivist denial does not preclude the utility – in non-

constructivist contexts – of supposing the existence of actual infinities.) Another

important logical contribution attributed to Leibniz is the notion of a possible world.

Besides the difficulty of extracting a notion of language from the Discourse, there

are some others. *First*, the logic of the Discourse is ontological and it is particularly an

85

ontology of God and creation. Accordingly, the logic of the Discourse is not subject-neutral, whereas contemporary logic *is* subject-neutral. Indeed, the subject-neutrality of first order logic is one of its strengths – a subject neutral logic can be employed to analyze (virtually) any semantic domain. *Prima facie*, however, the logic of the Discourse applies only to the metaphysics of God, or theology. *Second*, Leibniz employs God as an explanatory device, whereas some quite sophisticated philosophical analysis denies the existence of God. Russell went so far as to attempt to remove God from the Discourse, according to his supposition that Leibniz included God only so as to not contradict 'the prevailing opinions of his time'. But Russell's interpretation is deeply flawed, since whether or not God exists, and whether or not Leibniz believed that God exists, it remains that God is an instrumental component of the Discourse. Leibniz defines God very simply as *the absolutely perfect being*. The definition of God has an axiomatic role in the Discourse; it never changes, and virtually all the rest of Leibniz' metaphysics hangs on it. *Third*, the predicate/object distinction of contemporary logic is backwards from that of Leibniz. In contemporary logic, (with 'ϕ' being a predicate, and 'b' an object) '$\phi(b)$' means approximately, 'b is a ϕ'. In the logic of the Discourse, the expression is more helpfully written, '$b(\phi)$', meaning approximately 'ϕ is a (predicate of) b'. Just as in contemporary logic, Leibniz conceives ϕ as a predicate and b as an object. Thus in contemporary logic, we take two objects, x and y, to be identical if, for all predicates, Z, $Z(x)$ is true if and only if $Z(y)$ is true. This is an expression of Leibniz' law. However, for his Discourse, Leibniz' law is properly, 'two predicates W and Z, are identical if and only if for all objects, x, $x(W)$ is true if and only if $x(Z)$ is true. The problem of this predicate/object reversal is not very serious, as it requires only a simple

86

re-conception of the distinction. (It is unclear at this point whether or not the two predicate/object distinctions are equivalent, such that anything accomplished via Leibniz' distinction can also be accomplished via the contemporary distinction. It would be interesting to explore this question, but it is not strongly relevant to this essay.)

On the other hand, Leibniz employs some notions which approximate to key contemporary notions. Principally, these are the notions of decidability, analyticity and necessity. Furthermore, Leibniz employs consistency in two ways, both evocative of the role of consistency for first order logic. *First*, the predicates true of a substance must be consistent with each other – for no inconsistent substances can exist. (In first order logic, no inconsistent set of formulas has a model; that is, no inconsistent models exist.) This notion could be called intra-consistency. *Second*, all existing intra-consistent objects must be consistent with each other. This could be called inter-consistency. The need of intra-consistency is clear, as it corresponds tightly to the ancient principle of non-contradiction. The need of inter-consistency (which Leibniz calls 'compossibility') is somewhat less clear. Inter-consistency is given by Leibniz' argument that an intra-consistent object exists if and only if the object is compossible with the other objects of a maximally perfect domain of objects.

The logic of the Discourse includes logics of both human rationality and divine knowledge, but it does not strongly connect humanity with divinity. It is obvious that Leibniz violates the adage, 'dare not to ponder the mind of God', as he manages to say quite a lot about divine knowledge, but nevertheless, Leibniz fails to render divine knowledge intelligible to human rationality – as Leibniz himself indicates at several places. The disconnect between human rationality and divine knowledge is a disconnect

87

between kinds of understanding. For an explanation of this disconnect, Leibniz invokes a matter of infinity. Whereas God can conceive actual infinity 'in one intuition', humans cannot, and can work only with potential infinities. The incomprehensibility of divine knowledge by humans is perhaps the most critical problem facing the Discourse, since it greatly weakens its theodical impact. If human rationality and divine knowledge are too completely disconnected, then God becomes so different from humanity that He bears little or no relevance to humans, human values, successes and foibles. However, if we are prepared to accept the incomprehensibility of divine knowledge, then we need not propose that one system of logic characterizes both human rationality and divine knowledge.

Indeed, there is high utility in the hypothesis that human logic and divine logic are two distinct (even disjoint) kinds of logic. We shall call this hypothesis the 'two languages hypothesis', or '(2LH)' for short. We will not be hereafter concerned to develop the logic of human rationality, except according to its distinctions from the divine logic. It is the divine logic which we will develop into the 'Leibnizian' notion of logic and language.

The atomic components of the Discourse are the objects, which Leibniz calls the 'substances'. A substance is uniquely determined by the collection of those predicates which are true of the substance. Thus the substances are not material; they are not even empirical. (For Leibniz, the substances are representers of the actual constituents of reality. So the substances are not material, but are more like minds, though these minds are in general simpler and more limited than a human's mind. Indeed, Leibniz is explicit

88

about this in a letter[20], where he says that there is no contradiction in supposing that there are two 'conflicting' substances, but only 'disharmony'.) For the collection of predicates true of a substance, Leibniz uses the term 'notion'. We will not use Leibniz' term since we have already used it in a different sense. Instead, we continue to use the term 'extension' for what Leibniz calls a 'notion'. Put in more familiar terms, a substance is a *set* of predicates. It is unclear whether or not Leibniz would agree that a substance is a set, but the Discourse strongly suggests this view. We shall call this view the 'set-substance analogy'.

One reason for Leibniz to have denied the set-substance analogy is that he conceived the intra-*in*consistent substances to be incoherent and thus non-existent, but contrariwise, it is not the case that for every inter-inconsistent pair of set-substances, X and Y, either X or Y is incoherent. That is, it does not follow from the fact that X and Y are inter-inconsistent that (at least) one of them is intra-inconsistent. Where X and Y are inter-inconsistent, Leibniz does say that one of them – or rather, one of their corresponding substances – does not exist, but here Leibniz conceives 'existence' as 'the *actual* existence of the substance, X or Y, *in the world*. Of course, the intra-inconsistent sets are also non-existent, but here there are two kinds of non-existence. If X is intra-inconsistent, then it does not exist *in the world*. But then X is also non-existent as an abstraction. That is, X does not qualify as a set-substance because its membership is inconsistent, and thus X cannot be even a candidate for set-substance-hood, nor even for sethood. When it is said that some set-substance, X, *exists*, we take this to mean that even if X is inter-inconsistent with some other set, say Y – such that X does not exist *in the world* – X might still exist as an intra-consistent set.

---

[20] At present, I do not have a citation for this letter.

We now show that there are non-denumerably many substances in God's language. *First*, we assume that there are denumerably many predicates, and that they can be put into a list, as below. (For '(LP)' read 'the set of the listed predicates', and for 'p_n', read 'the nth predicate of the list'.)

(LP) $\{ p_0, p_1, p_2, \dots \}$

In the set-substance analogy, which may or may not be strongly supported by the Discourse, we allow *every* set of predicates to be a substance. Of course, not all such substances actually exist. For example, the set-substance analogy permits a substance with the predicates, '... is a (kind of) horse', '... cannot be harnessed', '... perceives and shuns evil', '... is big and white and strong', '... is magical', and '... has a single long horn growing from the forehead'. This is the substance of a unicorn, and of course, unicorns do not *actually* exist. But the set-substance analogy does not compel us to hold that there are actually-extant unicorn-substance-sets. Rather, all the set-substance analogy requires is that there exists an *abstraction* of the unicorn substance. Similarly, we say that the substance-set of the number 2, with predicates such as, '... is an even prime', '... is between 1 and 3', and so on, actually exists. The substance of 2 is an abstraction, but not merely so, for it *actually* exists as a term of the *actually* extant language of English. By contrast, the substance with the predicate '... is the second even prime' is a *mere* abstraction, and exists only as such. In this way, we can speak non-trivially of substances with impossible predicates such as '... is a square circle', '... is the last natural number', '... is green and red all over', and so on. Thus the set of all the abstractly extant substances is just the set of all subsets of (LP). In the above assumption, every predicate is uniquely subscripted by a natural number. Accordingly, every set of predicates is

expressable by a subset of (NN). (This follows trivially from the fact that there is a total one-to-one onto function from (LP) to (NN).) Thus the substance-extension,

$$\{ p_4, p_8, p_{92}, p_{3012} \}$$

is expressed by the subset of (NN),

$$\{ 4, 8, 92, 3012 \}.$$

Conversely, every subset of (NN) expresses a substance extension. It therefore follows that there are as many substance extensions as there are subsets of (NN). The set of all sub-sets of (NN) is the *powerset* of (NN), written, ' $\wp$(NN)'. (For the *cardinality* of a set, A – i.e., the number of members of A – we write '#(A)'.) Now by Cantor's powerset theorem (which is a diagonally-proved result), it follows that #(NN) < #($\wp$(NN)). By definition, (NN) is denumerable, and thus $\wp$(NN) has more than denumerably many members. There are more members of $\wp$(NN) than there are members of (NN) – to wit, $\wp$(NN) is uncountable.

In the context of the set-substance analogy, we now have it that there are non-denumerably many substances, or at least, that there are *potentially* non-denumerably many substances. (The constraints imposed by intra- and inter-consistency might restrict the number of *really* possible substances to a merely denumerable quantity.) And it immediately follows that the deific logic has non-denumerably many *truths*. The insufficiency of human language to express deific language is a consequence of the fact that human language can have only denumerably many sentences. If there is only one truth per sentence, then human language can potentially express only denumerably many truths, and can actually express only a finite number of truths.

91

The language which Leibniz attributes to God, I have dubbed 'Leibnizian'. Nevertheless, Leibniz is able to use his theory of God's language to solve some of the most persistent problems about God, such as those of free choice, culpability for sin, the existence of evil, and the existence of God. While it is not possible to know exactly how Leibniz would explain God's language, it is clear that this language is supposed to be adequate to 'do' Leibniz' metaphysics. Via human language and within the limits of human understanding, we cannot 'do' metaphysics, and must accept the compromise of 'doing' mere physics. It is certain that Leibniz would not have explained God's language as I propose below – for I will use contemporary logical resources for my proposal, and these resources were not available to Leibniz. In particular, Leibniz did not know the formal treatments of diagonal argument, fixed points or impredicative (self-referential) paradox. And of course, Leibniz did not know about Curry's paradox, in the terms of which I present the notion of Leibnizian language. Accordingly, 'Leibnizian language' is a mainly honorary appellation.

A 'LEIBNIZIAN' TREATMENT OF CURRY'S PARADOX

Let sm be a Curry predicate; thus sm is also a substitution predicate. We write the extension of sm as follows.

$$«m,0» «m,1» «m,2» ...$$

As sm is a Curry predicate, there is a natural, n, such that sn is a Curry sentence. Thus $«m,n» = 0$ if and only if $«m,n» = 1$. (Proof is omitted.) We are just as well supported in saying that the extension of sm is,

$$«m,0» «m,1» «m,2» ... «m,(n-1)» 0 «m,(n+1)» ... \tag{5a}$$

as that it is,

$$«m,0» «m,1» «m,2» ... «m,(n-1)» 1 «m,(n+1)» ... \tag{5b}$$

For purposes of the following investigation, let us accordingly say that sm has at least *two* extensions, (5a) and (5b). It is apparent that in order for sm to *actually* have the two above extensions, we must violate the law of the excluded middle with respect to the requirement that each of the si may have exactly one extension. Allowing sm to have two extensions is counterproductive if each of the si must have *exactly* one extension. Indeed, if our present reasoning were expressed in a formal language such as L, no predicate has multiple extensions (except for the undecidable predicates, which either have no extension or an only partial extension). But we do not follow the one-extension-per-predicate rule here. So we violate the law of the excluded middle with respect to predicate extensions, but we do not violate this law with respect to the cells of the array. That is, we still hold that (for all naturals, i and j) $«i,j»$ is 0 or 1, but not neither and not both.

It is well to note that even if we assert the excluded middle with respect to all the components of (A2), there will remain a contradiction which must be pacified in some way. If we propose to solve the treatment of sm's extension by allowing sm to have multiple extensions, which is a counterintuitive notion, we are still not straying too far from the solutions by Russell, Gödel and others, since they also forward counterintuitive claims (or rather, these claims were once *thought* to be counterintuitive). On the other hand, our treatment of sm's extension cannot accurately be characterized as a solution – since it is not at once obvious how to rigorously and consistently apply multiple extensions to single predicates. (So far as I know, there is no literature in which treats 'Leibnizian' language as a language of ambiguity, but of course there is a great deal of literature on ambiguity and ambiguity logic.[21] I am hopeful that that at least some of this literature has helpful or compelling suggestions of how to 'fill out' the Leibnizian notion of language which I only begin to develop here.) Nor will a rigorous and consistent solution be attempted. Rather, we use the notion of assigning multiple extensions as a diagnostic tool. The conclusion we will attempt to reach is that although there are only countably many predicates, the fact that they cannot all be listed together with their extensions is indicative of the uncountability of the extensions of the substitution predicates. Some substitution predicates, then, are diagonalized out of the (denumerable) set of substitution predicates, in the terms of which the diagonal predicate is defined. So even though the expressions of a language are countable, the extensions of the counter-diagonal expressions cannot be listed one-to-one with the expressions of the substitution

---

[21] Owing to time constraints, I have not yet built a bibliography of this material. However, any further development of the themes of this (my) essay must include a substantive consideration of the literature on ambiguity.

sentences in general. Thus a countable language may (if self-arithmetized) have an irreducably uncountable set of predicate extensions.

Let us illustrate the matter with a new array, (A3).

```
(A3)      f0 f1 f2 f3 ...
      s0| 1  0  1  1 ...
      s1| 0  0  1  0 ...
      s2| 1  1  0  1 ...
      s3| 0  1  0  1 ...
      . |  .  .  .  .
      . |  .  .  .  .
```

As the set of all substitution predicates of L (i.e., SIDE of (A3)) is denumerable, there is a natural number, p, such that sp is a Curry predicate. The set of all the well-formed-formulas of L (i.e., TOP of (A3)) is also denumerable. Included in TOP are infinitely many Curry sentences. We list some of the Curry sentences below.

$C(q)$ such that $q = \lceil C(q) \rceil$
$C(q')$ such that $q' = \lceil C(q) \wedge C(q') \rceil$
$C(q'')$ such that $q'' = \lceil C(q) \wedge C(q') \wedge C(q'') \rceil$

and so on, for all formulas $C(q^x)$, and for x, a string of x-many primes (with $x \in NN$). Clearly, there are denumerably many Curry sentences. We add this detail to (A3), yielding (A4).

```
(A4)         f0        f1      .. C(q)=fa .. C(q')=fb .. C(q'')=fc ..
       s0 |  1         0    ..   1       ..   1      ..   0           ..
       s1 |  0         0    ..   1       ..   0      ..   0           ..
       .  |  .         .    ..   .       ..   .      ..   .           ..
       .  |  .         .    ..   .       ..   .      ..   .           ..
       sp |  «p,0»    «p,1» ..  «p,a»    ..  «p,b»   ..  «p,c»        ..
       .  |  .         .    ..   .       ..   .      ..   .           ..
       .  |  .         .    ..   .       ..   .      ..   .           ..
```

As sp is a Curry predicate and as each of the values «p,y» (for fy one of $C(q)$, $C(q')$, and so on) is a Curry sentence, it follows that the values «p,a», «p,b», «p,c», and so on, are all inconsistent. That is, for every «p,y» (such that fy is a Curry sentence) it may be proved

95

that «*p,y*»=0 if and only if «*p,y*»=1. Let us now simplify (A4), so that TOP lists only the

Curry sentences, cs0, cs1, cs2, and so on.

```
(A5)          cs0     cs1     cs2     cs3    ..
        s0|    1       0       1       1     ..
        s1|    1       0       1       1     ..
        . |    .       .       .       .     ..
        . |    .       .       .       .     ..
        sp|  [p,0]   [p,1]   [p,2]   [p,3]   ..
        . |    .       .       .       .     ..
        . |    .       .       .       .     ..
```

There are now uncountably many extensions for *sp* in (A5). The set of all extensions for

sp is the set of all infinite strings of 0s and 1s. (The proof of this – that the set of all

infinite strings of 0s and 1s is uncountable – is already easy to see and is not reproduced

here.) Finally, as *sp* has uncountably many extensions, the set of all extensions of SIDE

is uncountable.

LEIBNIZIAN LANGUAGE, FORMALLY

The main concepts of Leibnizian language are defined as follows.

$P =$ The set of all the predicates. (It is assumed that P is denumerable, though it might be (as I think, *probably is*) uncountable. It is also assumed that the logical properties of P-as-denumerable hold monotonically valid within the logical properties of P-as-uncountable. It is a consequence of these two assumptions that the arguments below are not made unsound by re-interpreting them for P-as-uncountable.)

$p_0, p_1, p_2, \ldots$ The predicates themselves. These are also sometimes referred to collectively as 'the p_i'. The fact that the p_i are listable in the way above follows from the assumption that P is denumerable. 'p_x' is used as the *name* of the xth predicate.

$S =$ The Leibnizian 'universe' – the set of all the (abstractly extant) substances. As every subset of P is the extension of a substance, $S = \wp(P) = \{ \text{ all sets, } Q, \text{ such that } Q \subseteq P \ \}$.

$s_0, s_1, s_2, \ldots$ The substances themselves, sometimes referred to collectively as 'the s_j'. 's_x' is used as the *name* of the xth substance.

$\lceil s_x(p_y) \rceil =$ The semantic valuation of the formula, '$s_x(p_y)$' – one of the truth-letters, 'T' and 'F'.

(Here again is a place where the literature would illuminate my essay, but which literature I have not (yet) researched. I refer to Alexius Meinong's existence/subsistence distinction. According to this distinction, a 'golden mountain' does not *exist*, but does *subsist*, and as a subsistent entity, we may talk meaningfully about the mountain and the gold of which it is composed.)

One version of substance-comprehension, 'SC' – a close cousin of *set-comprehension* – and two versions of the principle of the excluded middle ('EM1' and 'EM2') for Leibnizian language:

> (SC)   Corresponding to every intuitively comprehensible substance-extension, *there is an actual (abstract) substance*. For example, the substance of 'truth' is comprehensible as the set of all true instances of the schema, '$s_x(p_y)$'.

(SC) is treated as if it were true, notwithstanding any possible arguments *contra*. But the truth of (SC) is not provided by *mere* fiat – for we have a good meta-theoretical reason to hold truth for (SC). We must preserve consistency with Leibniz' ideas, or fail to describe a genuinely *Leibnizian* notion of language. For suppose, *contra* (SC), that some intuition of a substance fails to consistently define a substance-extension. For example, the membership of the Russell set is inconsistent or non-existent. If there is a 'Russell substance', then it too has an inconsistent extension. Of course, there are trivially-inconsistent substance-extensions, such as those of a 'square circle' or a 'married bachelor'. Clearly, there are no empirically-extant square circles or married bachelors, and thus, there are no such substances either. But it is a different matter to conclude that

98

an inconsistent substance is non-existent *altogether*. No inconsistent substance has *empirical* existence, but it does exist as an *abstraction*. This is not an unfamiliar notion of existence. Hereafter, 'α exists' means 'α exists *as an abstraction*'. We note furthermore that the set-substance-analogy requires that every subset of the set of predicates *is* a substance-as-abstraction, even when its membership is inconsistent. (An inconsistent membership of a substance, E, occurs whenever there is a predicate, ϕ, such that we can reason either,

(a) both $\phi \in E$ and $\sim\phi \in E$, or,
(b) both $\phi \in E$ and $\phi \notin E$.

Note that the inconsistency of (a) is *among predicates*, whereas that of (b) is *among substance-extensions*.)

The law of the excluded middle does not in general hold true for Leibnizian language, unless (SC) is false. Two notions of the excluded middle, (EMa) and (EMb), which correspond with the two notions, (a) and (b), of substance-inconsistency, above.

(EMa) For all substances s_x and predicates p_y,
either $p_y \in s_x$ or $\sim p_y \in s_x$.

(EMb) For all substances s_x and predicates p_y,
either $p_y \in s_x$ or $p_y \notin s_x$.

$$\lceil s_x(p_y) \rceil = \quad \begin{cases} T & \text{if } p_y \in s_x, \\ & \\ F & \text{otherwise.} \end{cases}$$

At this point we must make a distinction among substance-names which is peculiar to Leibnizian language. Leibniz' metaphysical logic is treated throughout the Discourse as working differently for God than for humans. In particular, God knows the membership of all substances *for every predicate*. (We say that, in this sense, God knows

99

every substance 'to infinity'.) Humans, on the other hand, can generally know the membership of a substance only finitely – that is, for a finite number of predicates. Of course, humans can know some substances to infinity – for a few trivial examples, there are the substances of all the even-numbered predicates, the empty substance and the 'universal' substance, which has every predicate. Hereafter, we represent a substance, say s_m, which is known only finitely, up to say, the nth predicate, by, '$s_m{}^n$'. Similarly, a substance (the mth) which is known to infinity is represented by, '$s_m{}^\infty$'. The distinction between the finitely- and infinitely-known substances enables us to express a notion of predicate-extension conflation. We define the conflation of two substances by a substance name as follows.

> For every finitely-known substance, s, there are m' and m'' such that sd
>
> conflates sr with $s_{m''}^{q''}$ if and only if,

    (1) $q > n$; and,

    (2)   either $p_q \in s_{m'}^{q'}$ and $p_q \notin s_{m''}^{q''}$,

             or $p_q \notin s_{m'}^{q'}$ and $p_q \in s_{m''}^{q''}$; and,

    (3) $s_{m'}^{q'}$ and $s_{m''}^{q''}$ are identical with s_m^n up to and including predicate p_n.

We can now identify an undecidable substance, namely, the 'truth-substance', or s_T. It is a tricky task to have a predicate be an *expression* of a given truth while at the same time *being* a predicate true of s_T, as the Leibnizian subject/predicate distinction requires. Once this task has been accomplished, the formal definition of s_T will be ungainly, whereas a simpler (but not exactly Leibnizian) notion of s_T will suffice.

100

Accordingly, we use the following simplified notion of the *truth-substance*. The simplified notion of s_T requires that a truth-expressing predicate be written with subscripts for both a substance and a predicate. Thus the truth-expressing predicates will be written as, '$p_{x,y}$', where x is the subscript of a substance and y is the subscript of a predicate. $p_{x,y}$ is taken to express the formula, $s_x(p_y)$, or the sentence 'p_y is a predicate of s_x'. Note that every truth-predicate, $p_{x,y}$, is identical to some other singly-indexed predicate, p_z. To match a truth predicate with its singly-indexed equivalent, we employ a function, say f, which takes ordered pairs of natural numbers to natural numbers. That is, f is a total one-to-one function from NN×NN into NN. One function like f is suggested by the Gödel-coding function, *viz.*:

$$p_z = p_{x,y} \text{ if and only if } z = 2^x \times 3^y.$$

According to the simplified notion of the 'substance of truth', s_T,

$$s_T = \{ p_{x,y} \mid p_y \in s_x \}.$$

Similarly, the 'substance of falsehood' is s_F,

$$s_F = \{ p_{x,y} \mid p_y \notin s_x \}. \tag{1}$$

Let us now set $x = m$, and note that the formula, '$p_y \notin s_m$' has exactly one free (subscript-) variable, y. ('$p_y \notin s_m$' is a kind of substitution predicate.) We now define a *fixed point* on the formula, '$p_y \notin s_m$', as a subscript, n, such that $p_n = p_{m,n}$. (Proof that there actually *is* an n such that $p_n = p_{m,n}$ is omitted, in the assumption that there is an instance of the fixed point theorem for the formula, '$p_y \notin s_m$'.) Now consider the formula, '$p_y \notin s_F$', and a fixed point on it, q, such that $p_q = p_{F,q}$.

101

We now reason as follows. Suppose $p_q \in s_F$. Then as $p_q = p_{F,q}$, $p_{F,q} \in s_F$. Then by (1), $p_q \notin s_F$, whence contradiction. Now suppose $p_q \notin s_F$. Then by (1), $p_{F,q} \in s_F$. But as $p_q = p_{F,q}$, it follows that $p_q \in s_F$, whence contradiction. So $p_q \in s_F$ if and only if $p_q \notin s_F$.

In mathematical logic, when confronted by such a contradiction, one reasons according to some paradox-pacifying device, some of which we have already considered. (These include, for example, Russell's type theory and Tarski's predicate hierarchies.) It might be observed that the definition of s_g is impredicative, in such a way that Russell's vicious circle principle is violated, and s_g is accordingly not admissible in Russell's (ramified) type theory.

In its extension to Leibnizian language, Russell's principle forces us to say either that s_g is not a substance (even though it is a subset of P), or else that no subset of P is the extension of the intuitive definition of s_g. But there is a third option. According to the hypothesis above, we may understand the p_i and the s_j to be conflationary, understood under the restraint imposed by the finitude of all our calculations. (That is, let us not suppose that all the memberships of all the s_i are completely determined, simply because a name or definition is applied, for the definitions are also conflationary.) Some extensions are definable digit-by-digit to infinity, e.g., 1111111... or 1010101010... Other extensions are simply not yet computed (such that *distant* digits – say, the $(2^{1000000})^{\text{th}}$ digit of π), or are not computable within the next, say, two trillion years. So as s_g conflates an uncountable infinity of substances, we may suppose (for the time-being, at least) that the extension of s_g is not the countervalue of the extension of $\sim s_g$. (This appears as a double negative, and a trivial notion for that reason.) But I do not propose to

treat the extensions of s_g and $\sim s_g$ classically. I deny the law of double negation (as do also intuitionist logicians) with respect to the extensions of the substances. It is not always the case that $p_x \in s_y$ if and only if $p_x \notin \sim s_y$. Of course, this also goes to deny the excluded middle for Leibnizian language.

We can suppose that s_g conflates two extensions $s_{g'}$ and $s_{g''}$, such that $p_i \in s_{g'}$, and $p_i \notin s_{g''}$. And $s_{g'}$ and $s_{g''}$ are not known to differ with respect to any predicate other than p_i. Accordingly, $s_g(p_i)$ can be said true in the choice of $s_g = s_{g'}$. And $\sim s_g(p_i)$ can be said true in the choice of $s_g = s_{g''}$. But if we force s_g to have either $p_i \in s_g$ or $p_i \notin s_g$, then we have both. Thus, we should not think that s_g can ever be completely resolved (for all predicates). s_g must always conflate $s_{g'}$ and $s_{g''}$, for otherwise, the membership of s_g is inconsistent. Of course, this means that – *technically* – s_g cannot qualify as a Leibnizian substance, for if it is any subset of P, then it conflates the membership of p_i with its non-membership.

But this disqualification (of s_g from substance-hood) need not be a problem. Leibniz has already said that only God knows the complete membership of a substance – since whereas God can, humans cannot know the membership of all substances. (Even where the $^1/_0$-string is known for all digit-places, we will eventually not know the predicate which corresponds to some distant digit-place.) Because humans cannot know the complete membership of a substance, all of those substances are ambiguous as to the membership of some substances. So human analysis of substance *must* conflate (in the sense of being *ambiguous among*) some of the substances. Usually, we do not know which extensions are conflated by the human notion of a given substance. In general,

103

there are always some conflated extensions, which we cannot access. Let us call the sorts of extension-conflations the *expressive conflations* – conflations due to the incapacity of humans to even *name* all of the substances.

However, via reasoning according to paradoxical fixed points, we can analyze a different kind of ambiguity. This second mode of conflation is due to the fact that substances such as s_g cannot have only one subset of P as its extension – because if it *did* have a single extension, then it must be inconsistent as to the membership of p_i. I propose that we interpret the inconsistency of s_g's extension as a symptom of ambiguity. In the context of these comments, when $s_g(p_i)$ is true, this is because s_g is conflated with $s_{g'}$, and $p_i \in s_{g'}$. When $s_g(p_i)$ is false, this is because s_g is conflated with $s_{g''}$, and $p_i \notin s_{g''}$. Clearly, $s_{g'} \neq s_{g''}$. Also clearly, it is no contradiction for an object, say a, to be a member of one set, say B, and to not be a member of another set, say C. Thusly, in the hypothesis that some substances conflate a (possibly transfinite) number of extensions, the human proof of the paradox, though it has the form of a contradiction, need not entail inconsistency. Once *de*-conflated, the paradoxical conclusion,

$$s_g(p_i) \leftrightarrow \sim s_g(p_i)$$

resolves to

$$s_{g'}(p_i) \leftrightarrow \sim s_{g''}(p_i),$$

which is not the form of a contradiction.

Finally, the substance of truth, s_T, need not imply a contradiction. Rather, s_T conflates (infinitely) many extensions. The liar sentence is true in some extensions, say $s_{T'}$, and false in others, say $s_{T''}$. Thus, with the liar predicate being p_L, we obtain,

$$s_{T'}(p_L) \leftrightarrow \sim s_{T''}(p_L),$$

which is not a contradiction. In this way, paradox is pacified.

CUMULATIVE BIBLIOGRAPHY

[ACZ2] Aczel, Peter, Non-Well-Founded Sets, 1988, C.S.L.I., Stanford.

[ACZ5] --------, *Schematic consequence, in* What is a Logical System?, D.M. Gabbay (ed.), Clarendon Press, Oxford, pp.261-272.

[BOO2] Boolos, George S. *and* Richard C. Jeffrey, Computability and Logic *(third edition)*, 1989, Cambridge University Press, New York.

[BOO3] --------, *On "seeing" the truth of the Gödel sentence*, 1990, *in* [JEF], pp.389-391.

[CAN1] Cantor, Georg, *Uber eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen, in* Journal fur die reine und angewandte Mathematik, 1874, vol.77, pp.258-262.

[CAN2] --------, *Beitrage zur Begrundung der transfiniten Mengenlehre, in* Mathematische Annalen, 1895, vol.46, pp.481-512.

[CHO1] Chomsky, Noam, Syntactic Structures, 1957, Moulton, The Hague.

[CHO2] --------, Aspects of the Theory of Syntax, 1981, MIT Press, Cambridge.

[CHU] Church, Alonzo, *An unsolvable problem of elementary number theory*, 1936, American Journal of Mathematics, *vol.*58.

[CUR2] Curry, Haskell B., *The inconsistency of certain formal logics*, Journal of Symbolic Logic, 1942, #7, pp.115-117.

[DAW2] Dawson, John W., *Gödel and the limits of logic, in* Scientific American, June 1999, pp.76-81.

[DID] Didier, Dubois *and* Henri Prade (eds.), Fundamentals of Fuzzy Sets, 2000, Kluwer Academic Publishers, Boston.

[FRE1] Frege, Gottlob, *Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought*, 1879, *in* [vHE1], pp.???

[FRE3] --------, *Uber Sinn und Bedeutung, in* Zeitchrift fur philosophie und philosopischie kritik, n.s. vol.100, 1892.

[GEN] Gentzen, Gerhard, *The consistency of elementary number theory, in* [SZA, pp.132-201].

106

[GOD1] Gödel, Kurt, *On formally undecidable propositions of Principia Mathematica and related systems I*, 1931, *in* [vHE1].

[GUM1] Gumanski, Leon, *On decidability of the first order functional calculus*, 1983, *presented to* 7th International Congress of Logic, Methodology and Philosophy of Science, Salzburg.

[GUM2] --------, *Remarks on Cantor's diagonal method and some related topics*, *in* Types of Logical Systems and the Problem of Truth, 1988, Bulgaria Academy of Sciences, pp.45-56.

[HEN] Henkin, Leon, *A problem concerning provability*, *in* Journal of Symbolic Logic, vol.17, 1952.

[HER1] Herzberger, Hans, *New paradoxes for old*, Proceedings of the Aristotelian Society, 1980, n.s. Vol.LXXXI, pp.109-123.

[HIN] Hintikka, Jaakko, *Semantics: A revolt against Frege*, *in* Contemporary Philosophy: A New Survey *(vol.1)*, 1981, Martinus Nijhoff, Boston, pp.57-82.

[HOD] Hodges, Wilfrid, *An editor recalls some hopeless papers*, Bulletin of Symbolic Logic #4, 1998, pp.1-16. *(Note: Need the published paper itself to make the pagination match w/ citations.)*

[HOF] Hofstadter, Douglas, Gödel, Escher, Bach, Basic Books, 1979.

[JAC] Jackendoff, Ray, Consciousness and the Computational Mind, 1987, M.I.T. Press, Cambridge.

[JEF] Jeffrey, Richard *ed.*, Logic, Logic and Logic, 1998, Harvard University Press, Cambridge.

[JOH] Johnson, Mark, Attribute-Value Logic and the Theory of Grammar, CSLI, Stanford, 1988.

[KIN] King, James D., *A problem concerning the first incompleteness theorem*, 1999, *in* Logical Consequence: Rival Approaches, John Woods and Bryson Brown *eds.*, 2001, Hermes Science Publishing, Oxford.

[KLE1] Kleene, Stephen Cole, Introduction to Metamathematics, 1952, D. van Nostrand, Toronto.

[KLE2] --------, Mathematical Logic, 1967, John Wiley & Sons, New York.

[KLI] Klir, G.J., *Foundations of fuzzy set theory and fuzzy logic*, *in* International Journal of General Systems, 2001, v.30, no.2, pp.91-132.

[KNE] Kneale, William *and* Martha Kneale, The Development of Logic, 1962 (1984 printing), Clarendon Press, Oxford.

[LOB] Löb, M.H., *Solution of a problem of Henkin*, 1955, Journal of Symbolic Logic, *vol.*20, #2, pp.115-118.

[MEN] Mendelson, Elloitt, Introduction to Mathematical Logic, 1964, D. van Nostrand, Toronto.

[PART] Partee, Barbara H., Alice ter Meulen *and* Robert E. Wall, Mathematical Methods in Linguistics, 1990, Kluwer Academic Publishers, Boston.

[RIC] Richard, Jules, *The principles of mathematics and the problem of sets*, 1905, *in* [vHE1], pp.142-144.

[ROS] Ross, T.J. *and* W.J. Parkinson, *Fuzzy set theory, fuzzy logic and fuzzy systems*, *in* Fuzzy Logic and Probability Applications, 2002, pp.29-53.

[SIK] Sikic, Zvonimir, *The diagonal argument – a study of cases*, International Studies in the Philosophy of Science, 1992, *vol.*6, #3, pp.191-203.

[SIM] Simmons, Keith, Universality and the Liar: An essay on truth and the diagonal argument, 1993, Cambridge University Press.

[SMU1] Smullyan, Raymond M., Gödel's Incompleteness Theorems, 1992, Oxford University Press, New York.

[SMU2] --------, Diagonalization and Self-Reference, 1994, Clarendon Press, Oxford.

[TAR2] Tarski, Alfred, *The semantic conception of truth*, 1944, *in* Semantics and the Philosophy of Language, L. Linsky *ed.*, Illinois.

[vHE1] van Heijenoort, Jean, From Frege to Gödel, 1967, Harvard University Press, Cambridge.

[vHE2] --------, *Logic as calculus and logic as language*, Synthese #17, 1967, pp.324-330.

[WHI] Whitehead, Alfred North *and* Bertrand Russell, Principia Mathematica *vol.*1 *(second edition)*, 1963, Cambridge University Press.

BIBLIOGRAPHY OF UNCITED WORKS

[SHA] Shapiro, Stewart, *Induction and indefinite extensibility: The Gödel sentence is true, but did someone change the subject?*, 1998, Mind, *vol.*107, #427, pp.596-624.

[ROU] Rounds, William C., *The complexity of paradox*, *in* Situation Theory and Its Applications, 1990, Robin Cooper, Kuniaki Mukai *and* John Perry *eds.*, Stanford (U. Press?).

[PRE] Preist, Graham, *Semantic closure, descriptions and non-triviality*, 1998, *manuscript.*

[PUT1] Putnam, Hilary, *Mathematics without foundations*, 1967, *in* [BEN], pp.295-311.

[PUT2] --------, *Models and reality*, 1980, *in* [BEN], pp.421-444.

[NEA] Neale, Stephen *and* Josh Dever, *Slingshots and boomerangs*, January 1997, Mind, *vol.*106, #421, pp.143-168.

[OPP] Oppy, Graham, *The philosophical insignificance of Gödel's slingshot*, January 1997, Mind, *vol.*106, #421, pp.121-141.

[PARS] Parsons, Terence, *Assertion, denial and the liar paradox*, Journal of Philosophical Logic, 1984, #13, pp.137-152.

[LOR] Lorenzen, P., *Logical reflections and formalism*, 1958, Journal of Symbolic Logic, *vol.*23, #3, pp.241-249.

[MAR] Martin, Robert L., *Are natural languages universal?*, Synthese #32, 1976, pp.271-291.

[McC] McCall, Storrs, *A true but unprovable sentence, (undated manuscript, but apparently published sometime after 1998)*

[KOO] Koons, Robert, *Three indexical solutions to the liar paradox*, *in* Situation Theory and Its Applications, 1990, Robin Cooper, Kuniaki Mukai *and* John Perry *eds.*, Stanford (U. Press?).

[KAL] Kalderon, Mark Eli, *The transparency of truth*, Mind, July 1997, *vol.*106, #423, pp.475-497.

[ISA] Isaacson, Daniel, *Arithmetical truth and hidden higher-order concepts*, 1987, *in* Philosophy of Mathematics, W.D. Hart *ed.*, 1996, Oxford University Press, New York, pp.203-224.

109

[GUP] Gupta, Anil, *Truth and Paradox*, 1982, <u>Journal of Philosophical Logic</u>, #11, pp.1-60.

[FIN] Finsler, Paul, *Formal proofs and undecidability*, 1926, *in* [vHE1], pp.438-445.

[DEU] Deutsch, David *and* Artur Ekert, *Machines, logic and quantum physics*, 2000, <u>Bulletin of Symbolic Logic</u>, *vol.*6, #3, pp.265-283.

[ACZ1] Aczel, Peter, *Frege structures and the notions of proposition, truth and set*, *in* <u>The Kleene Symposium</u>, 1980, Jon Barwise, H.J. Keisler *and* K. Kunen *eds.*, North-Holland Publishing, pp.31-59.

[ACZ3] -------- *and* Rachel Lunnon, *Universes and parameters*, *in* <u>Situation Theory and Its Applications</u> (*vol.*2), 1991, Jon Barwise, Jean Mark Gawron, Gordon Plotkin *and* Syun Tutiya *eds.*, Stanford University (Press?), pp.3-24.

[ACZ4] --------, *Generalized set theory*, *in* <u>Logic, Language and Computation</u>, 1996, Jerry Seligmann *and* Dag Weterstahl *eds.*, Stanford (U. Press?), pp.1-17.

[BEN] Benacerraf, Paul *and* Hilary Putnam *eds.*, <u>Philosophy of Mathematics</u> *(second edition)*, 1983, Cambridge University Press, Cambridge.

[BOO1] Boolos, George S., *The iterative conception of set*, 1971, *in* [BEN], pp.486-502.

[BOO4] --------, *Constructing Cantorian Counterexamples*, 1997, *in* [JEF], pp.339-341. *(Includes editorial note by Vann McGee.)*

[BRO] Brown, J.R., *Proofs and pictures*, *in* <u>British Journal for the Philosophy of Science</u>, Je97; 48(2): 161-180.

[BUR] Burge, Tyler, *Sinning against Frege*, <u>Philosophical Review</u>, July 1979, *vol.*88, pp.398-432.

[BUJ] Burgess, John P., *Fundamental theorems of provability logic*, *in* [JEF], pp.415-419.

[CUR3] Curry, Haskell B., *Remarks on the definition and nature of mathematics*, 1954, *in* [BEN], pp.202-206.

[CUR4] --------, <u>Foundations of Mathematical Logic</u>, 1963, McGraw-Hill, Toronto.

[CUR5] --------, J. Roger Hindley *and* Jonathan P. Seldin, <u>Combinatory Logic</u> (*vol.*II), 1972, North-Holland Publishing, London.

[DAW1] Dawson, John W., *Facets of incompleteness*, *in* <u>Mathematical Logic and its Applications</u>, Dimiter G. Skodev *ed.*, Plenum Press, New York, pp.9-21.

[DUN1] Dunn, J. Michael, *Partial gaggles applied to logics with restricted structural rules*, in Substructural Logics, Tubengen, 1990, pp.63-108.

[DUN2] --------, *Gaggle theory: an abstraction of Galois connections and residuation, with applications to negation, implication and various logical operators*, in Logics and AI, Amsterdam, 1990, pp.31-51.

[GOD2] Gödel, Kurt, *On intuitionistic arithmetic and number theory*, 1933, in Kurt Gödel: Collected Works (*vol.*1), 1986, Solomon Feferman *et.al. eds.*, Oxford University Press, New York, pp.287-295.

[GOD4] --------, *Russell's mathematical logic*, 1944, *in* [BEN], pp.447-469.

[HER2] Herzberger, Hans, *Notes on naive semantics*, 1982, Journal of Philosophical Logic, #11, pp.61-102.

[RUS] Russell, Bertrand, Introduction to Mathematical Philosophy, 1919, George Allen & Unwin, London.

[SHO] Shoenfield, Joseph R., Mathematical Logic, 1967, Addison-Wesley Publishing, Don Mills (Ontario).

[SMU3] Smullyan, Raymond, Forever Undecided, Alfred A. Knopf, 1987.

[SOR] Sorensen, Roy A., *Yablo's paradox and kindred infinite liars*, 1998, Mind, *vol.*107, #425, pp.137-155.

[SZA] Szabo, M.E. *ed.*, The Collected Papers of Gerhard Gentzen, 1969, North-Holland Publishing, London.

[TAR1] Tarski, Alfred, *Der Wahreitsbegriff in den formalisierten Sprachen*, in Studia Philosophica, 1936, vol.1, pp.261-405.

[WOO1] Woods, John H., Paradox and Paraconsistency: Whither Logic?, 1990, *manuscript*.

[YAB] Yablo, Steve, *Grounding, dependence, and paradox*, 1982, Journal of Philosophical Logic, #11, pp.117-137.