

# What *If*? The Exploration of an Idea

Graham Priest

Departments of Philosophy, the Graduate Center, City University of New York, and the University of Melbourne

## Contents

<b>1</b>	<b>Introduction</b>	<b>57</b>
<b>2</b>	<b>Logical Preliminaries</b>	<b>59</b>
2.1	<i>LP</i> . . . . .	59
2.2	Material Equivalence . . . . .	61
2.3	The Substitutivity of Equivalentents . . . . .	62
2.4	Monotonicity . . . . .	63
<b>3</b>	<b>A First Look at the Idea</b>	<b>64</b>
3.1	The Positives . . . . .	64
3.2	The Negatives . . . . .	66
<b>4</b>	<b>Some Advantages of the Proposal</b>	<b>66</b>
4.1	Curry Paradox . . . . .	67
4.2	Validity Curry . . . . .	68
4.3	Infinitary Curry . . . . .	70
4.4	An Extended Paradox . . . . .	72
4.5	Denotation Paradoxes . . . . .	72
4.6	Kripke's Paradox . . . . .	74
<b>5</b>	<b>Endorsement</b>	<b>75</b>
5.1	Transparent Truth . . . . .	76
5.2	Propositional Quantification . . . . .	77

<b>6</b>	<b>An Additional Conditional?</b>	<b>79</b>
6.1	Truth and a Detachable Conditional . . . . .	79
6.2	Quasi-Validity . . . . .	79
6.3	<i>LPm</i> . . . . .	81
6.4	The Material Conditional . . . . .	83
<b>7</b>	<b>A Consistency Operator</b>	<b>86</b>
7.1	What This is . . . . .	86
7.2	Classical Recapture . . . . .	88
7.3	True Only . . . . .	89
<b>8</b>	<b>Restricted Quantification</b>	<b>90</b>
8.1	The Problem . . . . .	90
8.2	Solution by the Consistency Operator . . . . .	91
<b>9</b>	<b>Interim Conclusions</b>	<b>93</b>
<b>10</b>	<b>The Comprehension Schema</b>	<b>94</b>
<b>11</b>	<b>Regaining Set Theory</b>	<b>97</b>
11.1	Type Lifting . . . . .	98
11.1.1	With a Partition . . . . .	98
11.1.2	With a Covering . . . . .	101
11.2	Expansions . . . . .	103
11.2.1	The Basic Construction . . . . .	103
11.2.2	Variations . . . . .	105
11.3	Meadows' Objections . . . . .	107
<b>12</b>	<b>Metatheory</b>	<b>109</b>
12.1	Model Theoretic Validity . . . . .	109
12.2	Defeasible Deductive Inference . . . . .	109
12.3	Extensional Equivalence . . . . .	110
<b>13</b>	<b><i>In Contradiction</i> Revisited</b>	<b>111</b>
<b>14</b>	<b>And to Conclude...</b>	<b>114</b>

<b>15 Appendix: Curry's Paradox</b>	<b>114</b>
15.1 The Inclosure Scheme . . . . .	115
15.2 What is a Paradox? . . . . .	115
15.3 The Liar and Curry . . . . .	117
15.4 Neo-Curry Paradoxes . . . . .	118
15.5 The Material Conditional . . . . .	120
15.6 Sorites . . . . .	122
15.7 The Bottom Line . . . . .	123

# 1 Introduction

The paradoxes of self-reference, such as the Liar Paradox, Russell’s Paradox, and König’s Paradox are *prima facie* sound arguments which give rise to contradictions. For the most part, solutions—both historical and contemporary—have tried to solve the paradoxes by breaking the arguments: some premise is not true, or some inference employed is invalid. The results have not been encouraging. After two and a half thousand years, there has been no successful solution, at least for the semantic paradoxes, if consensus is to be the mark of success. A dialethic solution is historically very recent, and stands out against the crowd. At its core is the thought that there is absolutely nothing wrong with the arguments, and that the contradictory conclusions are simply to be accepted.

Central to the paradoxes are very general principles governing semantic and set theoretic apparatus. These all deploy some notion of conditionality, or more accurately, biconditionality. Let us use  $\Rightarrow$  and  $\Leftrightarrow$  as generic symbols for a conditional and biconditional. (We assume that these are related in the usual fashion.)<sup>1</sup> Let  $T$  be the truth predicate,  $S$  the one-place satisfaction relation,  $D$  the denotation relation, and  $\langle \cdot \rangle$  a name-forming device. Then the principles are as follows. ( $A_x(y)$  indicates the formula  $A$  with every free occurrence of  $x$  replaced by  $y$ —relabelling bound variables if necessary, to avoid clashes).

- *T-Schema*: for all closed sentences,  $A$ ,  $T \langle A \rangle \Leftrightarrow A$
- *S-Schema*: for every formula,  $A$ , with one free variable,  $x$ ,  $S(y, \langle A \rangle) \Leftrightarrow A_x(y)$
- *D-Schema*: for every (closed) name,  $c$ :  $D(\langle c \rangle, y) \Leftrightarrow c = y$
- *Set Abstraction*: for every formula,  $A$ :  $y \varepsilon \{x : A\} \Leftrightarrow A_x(y)$
- *Property Abstraction*: for every formula,  $A$ :  $\lambda x A(y) \Leftrightarrow A_x(y)$

Let us call these *naive* principles.<sup>2</sup>

---

<sup>1</sup>In this essay, I will use  $\supset$  for the material conditional, and  $\rightarrow$  for a detachable intensional conditional. I use  $\Rightarrow$  when I want to be non-committal about what conditional is employed.

<sup>2</sup>From the Old French, *naïf*, meaning *natural*.

A crucial question here is what, exactly, the conditional,  $\Rightarrow$ , in the naive principles is. In ‘Logic of Paradox’,<sup>3</sup> I outlined two options. One is to take it to be the material conditional of the extensional paraconsistent logic  $LP$ ,  $\supset$ . Call this *Strategy 1*.  $LP$  is a relatively weak logic, however. In particular,  $\supset$  does not detach. (That is, the inference from  $A$  and  $A \supset B$  to  $B$  is not valid.) The other strategy is to take it to be some detachable conditional. Call this *Strategy 2*. Since there is no candidate for this in  $LP$ , it has to be added to the language, and appropriate semantics given. A natural candidate here is the conditional of some non-contracting relevant logic, in which  $LP$  can be embedded.

Dialetheism, and its applications to the paradoxes and other things, were articulated in *In Contradiction*, hereafter, *IC*.<sup>4</sup> In this, though there is certainly material relevant to Strategy 1, it was mainly Strategy 2 that was pursued; and the strategy of taking  $\Rightarrow$  to be a detachable conditional has been pursued by most people who have endorsed the naive principles.<sup>5</sup> By contrast, Strategy 1 has been advocated by few, a notable exception being Goodship (1996).

The aim of the present essay is to investigate Strategy 1. It is not to advocate it. The work is simply an extended exploration of the strategy, its strengths, its weaknesses, and the various different ways in which it may be implemented. In the first part of the paper I will set up the appropriate background details. In the second, I will look at the strategy as it applies to the semantic paradoxes. In the third I will look at how it applies to the set-theoretic paradoxes. One of the crucial issues to emerge will be that of whether Curry’s paradox is of the same kind as the Liar. This deserves an extended discussion, but such would take us off at something of a tangent to the main investigations; so I reserve the discussion for an appendix of the paper.<sup>6</sup>

---

<sup>3</sup>Priest (1979).

<sup>4</sup>Priest (1987). References are to the second edition.

<sup>5</sup>See, e.g., Beall (2009), Brady, (2006), Field (2009), Routley (1977). Brady and Field produce consistent formal theories by not endorsing the Law of Excluded Middle.

<sup>6</sup>Much of this paper was written in the second half of 2013, when I was a Humboldt Fellow at the University of Bochum. Many thanks go to the Humboldt Foundation for its generous support. Many thanks, too, go to Heinrich Wansing and his department for the congenial and stimulating atmosphere provided.

## 2 Logical Preliminaries

### 2.1 $LP$

In this section I will spell out some preliminary formal matters concerning the logic  $LP$ .<sup>7</sup>

Take a first-order language (without function symbols, for simplicity), with the connectives for negation, conjunction, disjunction ( $\neg$ ,  $\wedge$ ,  $\vee$ ), and the particular and universal quantifiers ( $\exists$  and  $\forall$ ).  $A \supset B$  is defined as  $\neg A \vee B$ , and  $A \equiv B$  as  $(A \supset B) \wedge (B \supset A)$ . We also assume a distinguished binary predicate of identity,  $=$ .

An interpretation,  $\mathfrak{M}$ , is a pair  $\langle D, \delta \rangle$ , where  $D$  is the non-empty domain of quantification; for every constant,  $c$ ,  $\delta(c) \in D$ ; and for every  $n$ -place predicate,  $P_n$ ,  $\delta(P_n) = \langle X, Y \rangle$ , where  $X \cup Y = D^n$  (the set of all  $n$ -tuples of  $D$ ). We will write  $X$  and  $Y$  as  $\delta^+(P_n)$  and  $\delta^-(P_n)$ , respectively.  $\delta^+(=) = \{\langle x, x \rangle : x \in D\}$ .

If we were to drop the constraint that for all predicates,  $P_n$ ,  $\delta^+(P_n) \cup \delta^-(P_n) = D^n$ , we would have the logic of first degree entailment,  $FDE$  (which does not have the Law of Excluded Middle). If we were to add the constraint that for all predicates,  $P_n$ ,  $\delta^+(P_n) \cap \delta^-(P_n) = \emptyset$ , we would have classical logic,  $CL$ .

We can now define what it is for a (closed) formula to be true,  $\Vdash^+$ , or false,  $\Vdash^-$ , in  $\mathfrak{M}$ . For the quantifiers, we assume that the language has been extended by a bunch of new constants,  $k_d$ , one for every  $d \in D$ , such that  $\delta(k_d) = d$ . We will call this extension the *language* of the interpretation.

For closed formulas of the language of  $\mathfrak{M}$ :

- $\Vdash^+ P_n c_1 \dots c_n$  iff  $\langle \delta(c_1), \dots, \delta(c_n) \rangle \in \delta^+(P_n)$
- $\Vdash^- P_n c_1 \dots c_n$  iff  $\langle \delta(c_1), \dots, \delta(c_n) \rangle \in \delta^-(P_n)$
- $\Vdash^+ \neg A$  iff  $\Vdash^- A$
- $\Vdash^- \neg A$  iff  $\Vdash^+ A$
- $\Vdash^+ A \wedge B$  iff  $\Vdash^+ A$  and  $\Vdash^+ B$
- $\Vdash^- A \wedge B$  iff  $\Vdash^- A$  or  $\Vdash^- B$

---

<sup>7</sup>See Priest (2008), chs 7 and 21.

- $\Vdash^+ A \vee B$  iff  $\Vdash^+ A$  or  $\Vdash^+ B$
- $\Vdash^- A \vee B$  iff  $\Vdash^- A$  and  $\Vdash^- B$
- $\Vdash^+ \exists x A$  iff for some  $d \in D$   $\Vdash^+ A_x(k_d)$
- $\Vdash^- \exists x A$  iff for all  $d \in D$   $\Vdash^- A_x(k_d)$
- $\Vdash^+ \forall x A$  iff for all  $d \in D$   $\Vdash^+ A_x(k_d)$
- $\Vdash^- \forall x A$  iff for some  $d \in D$   $\Vdash^- A_x(k_d)$

An interpretation is a *model* of a formula,  $A$ , iff  $\Vdash^+ A$ ; it is a model of a set of formulas,  $\Sigma$ , if it is a model for every formula in  $\Sigma$ . And an inference is valid,  $\Sigma \vDash_{LP} A$ , iff every interpretation which is a model of  $\Sigma$  is a model of  $A$ . (In what follows, I will often omit the subscript when no confusion can arise.)

In order to visualise how the truth functions work, it is sometimes helpful to represent the information as follows. Let us write  $t$ ,  $f$ , and  $b$ , to mean, respectively, that a formula relates only to truth, only to falsity, or to both. Then the truth and falsity conditions may be depicted in the following tables:

$\vee$	$t$	$b$	$f$	$\wedge$	$t$	$b$	$f$	$\neg$	$t$	$f$
	$t$	$t$	$t$		$t$	$t$	$b$		$t$	$f$
	$b$	$t$	$b$		$b$	$b$	$b$		$b$	$b$
	$f$	$t$	$b$		$f$	$f$	$f$		$f$	$t$

It is easy to check that the tables for  $\supset$  and  $\equiv$  are then:

$\supset$	$t$	$b$	$f$	$\equiv$	$t$	$b$	$f$
	$t$	$t$	$b$		$t$	$t$	$b$
	$b$	$t$	$b$		$b$	$b$	$b$
	$f$	$t$	$t$		$f$	$f$	$b$

$\forall$  and  $\exists$  are the infinitary analogues of  $\wedge$  and  $\vee$ . If one thinks of  $LP$  as a three-valued logic, then the designated values are  $t$  and  $b$ .

Clearly, if  $\Sigma \vDash_{LP} A$  then  $\Sigma \vDash_{CL} A$ , but not conversely, since  $LP$  is paraconsistent. However, it is also the case that if  $\vDash_{LP} A$  then  $\vDash_{CL} A$ .<sup>8</sup>

---

<sup>8</sup>See *IC*, ch. 6.

## 2.2 Material Equivalence

Since material equivalence will play a central role in what follows, it will pay to think about it a little more. Given the above tables, it is easy to check the following inferences:

- $A \wedge B \vDash A \equiv B$
- $\neg A \wedge \neg B \vDash A \equiv B$
- $A \equiv B \vDash (A \wedge B) \vee (\neg A \wedge \neg B)$
- $A \wedge \neg B \vDash \neg(A \equiv B)$
- $\neg A \wedge B \vDash \neg(A \equiv B)$
- $\neg(A \equiv B) \vDash (A \wedge \neg B) \vee (\neg A \wedge B)$

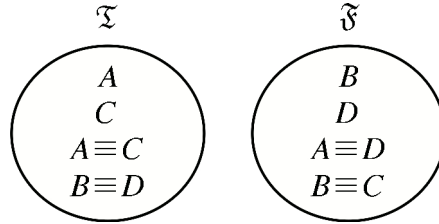
The first three inferences say that  $A \equiv B$  is true iff  $A$  and  $B$  are both true or both false. The second three say that  $A \equiv B$  is false iff one of  $A$  and  $B$  is true, and the other false. The following are also easy to check, and useful to note, where  $A!$  is  $A \wedge \neg A$ :

- $\vDash A \equiv A$
- $A \equiv B \vDash B \equiv A$
- $A \equiv B, B \equiv C \vDash (A \equiv C) \vee B!$
- $A! \vDash A \equiv B$
- $A! \vDash \neg(A \equiv B)$

Perhaps the most surprising of these initially is the fourth. The rationale is as follows.  $B$  is either true or false. Suppose that it is true. (The other case is similar.) Given that  $A!$  is true,  $A$  is true; so  $A$  and  $B$  have the same truth value, and  $A \equiv B$  is true. Of course, given that  $A!$  is true,  $\neg A$  is true too, so  $A$  and  $B$  have different truth values, and  $\neg(A \equiv B)$  is true too—which is what the fifth inference tells us. Note also, as is easy to check, that  $A, A \equiv B \not\vDash B$ , though  $A, A \equiv B \vDash B \vee A!$ .

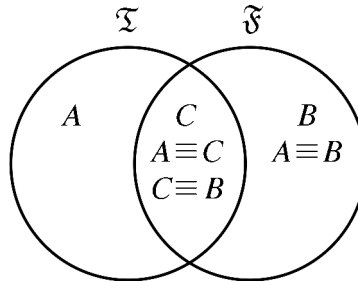
Finally, it may be illuminating to look at matters diagrammatically. Classically, every interpretation partitions sentences of the language into two zones, the truths ( $\mathfrak{T}$ ) and the falsehoods ( $\mathfrak{F}$ ), the two zones being mutually exclusive and exhaustive:





If two sentences are both in the same zone, their material equivalence is in the  $\mathfrak{T}$  zone; whilst if one is in one zone, and the other is in the other zone, their material equivalence is in the  $\mathfrak{F}$  zone. (See the diagram above.)

In *LP*, everything is the same except that the  $\mathfrak{T}$  and the  $\mathfrak{F}$  zones may overlap. Thus we have the following picture:



As before, the material equivalence of two sentences is in the  $\mathfrak{T}$  zone iff both are in the same zone ( $\mathfrak{T}$  or  $\mathfrak{F}$ ), and in the  $\mathfrak{F}$  zone if they are in different zones. But now, of course, a sentence—and so a material equivalence—can be in both zones. (See the next diagram.)

### 2.3 The Substitutivity of Equivalents

We can also see that the material conditional satisfies the substitutivity of equivalents. First, note the following facts about *LP*. Given any interpretation,  $\mathfrak{M}$ :

- if  $\Vdash^+ A \equiv B$  then  $\Vdash^+ \neg A \equiv \neg B$
- If  $\Vdash^+ A \equiv B$  and  $\Vdash^+ C \equiv D$  then  $\Vdash^+ (A \wedge C) \equiv (B \wedge D)$
- If  $\Vdash^+ A_x(k_d) \equiv B_x(k_d)$  for all  $k_d$  in the language of the interpretation, then  $\Vdash^+ \forall x A \equiv \forall x B$

The first is easy to see. For the second, suppose that, in  $\mathfrak{M}$ ,  $(A \wedge C) \equiv (B \wedge D)$  is  $f$ . Then  $A \wedge C$  is  $t$  and  $B \wedge D$  is  $f$ —or vice versa, the result of which is the same. Then  $A$  and  $C$  are  $t$ , and at least one of  $B$  and  $D$  is  $f$ . So either  $A \equiv B$  or  $C \equiv D$  is  $f$ . For the third, suppose the antecedent. Then by the truth conditions of the universal quantifier,  $\Vdash^+ \forall x(A \equiv B)$ . But  $\forall x(A \equiv B) \models \forall xA \equiv \forall xB$ . For suppose that in an interpretation the conclusion is  $f$ . Then  $\forall xA$  is  $t$  and  $\forall xB$  is  $f$ —or vice versa, the result of which is the same. Hence every instance of  $A$  is  $t$ , and some instance of  $B$  is  $f$ . Hence, some instance of  $A \equiv B$  is  $f$ , and so is  $\forall x(A \equiv B)$ . The result follows.

It is now easy to show that in any interpretation, if  $\Vdash^+ [\forall](A \equiv B)$ , then  $\Vdash^+ D_C(A_s) \equiv D_C(B_s)$ . Here,  $[\forall]$  indicates universal closure,  $s$  is any map from variables to the constants of the language of the interpretation,  $C_s$  is  $C$  with the free variables substituted according to  $s$ , and  $D_C(E)$  is  $D$  with every occurrence of the subformula  $C$  replaced by  $E$ . If  $C$  does not occur in  $D$ , the result holds vacuously. If it does,  $D$  is made up of truth functions and quantifiers applied to  $C$ . We can now prove the result by induction on the construction. The basis case follows by universal instantiation, and the cases for  $\neg$ ,  $\wedge$ , and  $\forall$  are given by the above results. Disjunction and the particular quantifier can be thought of as defined in the usual ways. It follows that provided no variable free in  $A$  or  $B$  is bound when substituted into  $D$ ,  $[\forall](A \equiv B) \models [\forall](D_C(A) \equiv D_C(B))$ .

## 2.4 Monotonicity

Finally, one technical result will be useful in what follows. The connectives and quantifiers of  $LP$  are *monotonic*—in the sense that if one takes a formula and increases the truth values of its atomic components (say, from just true to true and false), the truth values of the formula do not decrease. We can make the point precise as follows. Given two  $LP$  interpretations,  $\mathfrak{M}_1 = \langle D_1, \delta_1 \rangle$  and  $\mathfrak{M}_2 = \langle D_2, \delta_2 \rangle$ , say that:

- $\mathfrak{M}_1 \leq \mathfrak{M}_2$  iff  $D_1 = D_2$ ; and for every predicate,  $P$ ,  $\delta_1^+(P) \subseteq \delta_2^+(P)$  and  $\delta_1^-(P) \subseteq \delta_2^-(P)$

The *Monotonicity Lemma* then tells us that if  $\mathfrak{M}_1 \leq \mathfrak{M}_2$  then for every sentence,  $A$ , in the language of  $\mathfrak{M}_1$  (or  $\mathfrak{M}_2$ ):

- if  $\mathfrak{M}_1 \Vdash^+ A$  then  $\mathfrak{M}_2 \Vdash^+ A$

- if  $\mathfrak{M}_1 \Vdash^- A$  then  $\mathfrak{M}_2 \Vdash^- A$

The Lemma may be proved by a straightforward induction.

Looking at things in terms of  $t$ ,  $b$ , and  $f$ , and letting the values of  $A$  in  $\mathfrak{M}_1$  and  $\mathfrak{M}_2$ , respectively, be  $|A|_1$  and  $|A|_2$ , it follows that the value  $b$  is preserved upwards:

- if  $\mathfrak{M}_1 \leq \mathfrak{M}_2$  then if  $|A|_1 = b$  then  $|A|_2 = b$

and that classical values are preserved downwards. That is:

- if  $\mathfrak{M}_1 \leq \mathfrak{M}_2$  then, if  $|A|_2 = t$ ,  $|A|_1 = t$
- if  $\mathfrak{M}_1 \leq \mathfrak{M}_2$  then, if  $|A|_2 = f$ ,  $|A|_1 = f$

### 3 A First Look at the Idea

#### 3.1 The Positives

Thus primed, let us turn to the basic idea of Strategy 1 with respect to the semantic paradoxes. I will direct my remarks, in the first instance, to truth, though similar comments apply to the other semantic notions.

The idea that the biconditional involved in the  $T$ -Schema is the material conditional is certainly a very simple and natural one. As we have just seen, a material conditional is true, just if both of its flanks are in the same zone: both true or both false. Now, the prime idea that informs the natural view about truth is exactly that  $A$  and  $T \langle A \rangle$  have the same truth value; that is, that they are in the same zone (possibly necessarily so; but we do not need to take a stand on this here). This is exactly what the material conditional expresses. So the prime idea about truth is expressed exactly by the Schema  $T \langle A \rangle \equiv A$ .<sup>9</sup> We may define a falsity predicate,  $F \langle A \rangle$  as  $T \langle \neg A \rangle$ . It then follows that  $F \langle A \rangle \equiv \neg A$ .

Note three things. (i) By substitutivity,  $T \langle A \rangle \equiv A \vDash D_C(T \langle A \rangle) \equiv D_C(A)$  (and similarly for  $F$ .) (ii) The  $T$ -Schema in this form contraposes, since  $A \equiv B \vDash \neg A \equiv \neg B$ . This does not mean that truth commutes with negation, however. We have  $T \langle \neg A \rangle \equiv \neg A$  and  $\neg T \langle A \rangle \equiv \neg A$  but this entails

---

<sup>9</sup>It might be thought that  $A$  and  $T \langle A \rangle$  should have *exactly* the same truth values, but this is moot. If  $A$  is true and false, then *ex hypothesi*,  $T \langle A \rangle$  is true, but it may not be false as well. It may express the truth of  $A$  consistently. See *IC*, 5.4.

only that  $(T \langle \neg A \rangle \equiv \neg T \langle A \rangle) \vee A!$ . (iii) The material biconditional does not determine the extension of  $T$  completely. If  $T_1 \langle A \rangle \equiv A$  and  $T_2 \langle A \rangle \equiv A$ , it follows that  $(T_1 \langle A \rangle \equiv T_2 \langle A \rangle) \vee A!$ . So what happens when  $A$  is a dialetheia is not determined. The first disjunct says that  $T_1 \langle A \rangle$  and  $T_2 \langle A \rangle$  are both true or both false, but even this leaves open the possibility that one of these, but only one, is both.

Next, the Schema does, indeed, give rise to contradictions when self-reference is present. With appropriate self-reference, we can find a sentence,  $L$ , of the form  $\neg T \langle L \rangle$ . The Schema then delivers a sentence of the form  $\neg L \equiv L$ , that is,  $(\neg L \supset L) \wedge (L \supset \neg L)$ ; that is,  $(\neg \neg L \vee L) \wedge (\neg L \vee \neg L)$ . And given the law of double negation, this is just  $L!$ . Note that the argument does not even make use of the LEM.<sup>10</sup>

Next, the theory, even when annexed to arithmetic, and so self-reference, is non-trivial. In particular, take the set of all sentences true in the standard model of arithmetic. Use gödel coding for naming; this delivers self-reference. The addition of the material  $T$ -Schema produces a conservative extension. The proof of this is almost trivial.<sup>11</sup> Take the standard model of arithmetic, with domain  $N$ , and add a denotation for the truth predicate which makes  $\delta^+(T) = \delta^-(T) = N$ . This does not disturb the arithmetic vocabulary, but renders all instances of the  $T$ -Schema true, since  $A! \models A \equiv B$ . Of course, this is not a particularly interesting model, and there is no reason to believe that the truth predicate actually works like this—indeed, there are reasons to suppose that it does not. There are models where the truth predicate is much more consistent than this. For example, as we will see in due course, there are models of the theory in which  $A$  and  $T \langle A \rangle$  have exactly the same truth values, and  $T \langle A \rangle$  is consistent if  $A$  is grounded (in pretty much Kripke's sense).

Finally, other things being equal, this approach to the paradoxes is an exceptionally simple solution to the paradox. The  $T$ -Schema +  $LP$  is simpler by orders of magnitude than virtually all the other extant accounts of truth

---

<sup>10</sup>More generally, it is not difficult to show that if  $\Sigma \models_{CL} A$  then for some  $B$ ,  $\Sigma \models_{LP} A \vee B!$ . (See *IC*, 8.6.) Hence, if  $\Sigma$  classically entails a contradiction,  $A!$ , it entails a contradiction in  $LP$ ,  $A! \vee B!$ .

<sup>11</sup>See Priest (2002a), 8.1. It is worth noting that with the usual language of formalised arithmetic (where the only function symbols are those for successor, addition, and multiplication), we cannot find a sentence,  $L$ , that *is*  $\neg T \langle L \rangle$  (as opposed to being equivalent to it). The standard fixed-point construction delivers a sentence such that  $L \equiv \neg T \langle L \rangle$ . But  $(L \equiv \neg T \langle L \rangle) \wedge (L \equiv T \langle L \rangle)$  entails  $(T \langle L \rangle \equiv \neg T \langle L \rangle) \vee L!$ . So we still have  $T \langle L \rangle! \vee L!$ .

and paradox.

### 3.2 The Negatives

The main objection to the core idea, is, however, obvious: the material conditional (and biconditional) do not detach. So, given only the material  $T$ -scheme, one cannot infer  $A$  from  $T \langle A \rangle$ , or vice versa. Now, we certainly do seem to reason using detachment in this way. Thus, for example, as many have noted,<sup>12</sup> one function of the  $T$ -scheme would appear to be to make endorsement possible in contexts where this cannot be done explicitly.

One such context is that of blind endorsement, as where one might say: Everything the Bible says is true. Then, it is natural to reason as follows:

Everything the Bible says is true.

The Bible says that there is a time for every purpose under heaven.

So there is a time for every purpose under heaven.

Assuming that we can conclude from the premises that it is true that there is a time for every purpose under heaven, it would appear that we need the  $T$ -Schema and detachment to infer the last line.

Similarly, we might reason:

Everything true is said in the Bible.

People have evolved.

So the Bible says that people have evolved.

Here we need detachment with the  $T$ -Schema in the other direction.

Another context is where we would endorse all of an infinite number of claims. Thus, we might say ‘All of the axioms of Peano Arithmetic are true’ (the axioms being infinite in number). Given that ‘ $\forall x x' \neq 0$ ’ is an axiom of Peano Arithmetic, we may infer that  $T \langle \forall x x' \neq 0 \rangle$ , but with a non-detachable  $T$ -Schema, one cannot infer that  $\forall x x' \neq 0$ .

## 4 Some Advantages of the Proposal

So much for a first look at matters. Let us now examine the positives and negatives in more detail, starting with the positives in this section. One of

---

<sup>12</sup>I do so in *IC*, 4.2.

the most striking thing about Strategy 1 is that it handles not just the Liar paradox, but a plethora of others as well, as we will now see.

## 4.1 Curry Paradox

For Goodship, a major advantage of Strategy 1 concerns the Curry Paradox. One standard version of this concerns a sentence,  $C$ , of the form  $T \langle C \rangle \Rightarrow \perp$ , where  $\perp$  is a logical constant which entails everything. Putting the argument in natural deduction terms, its core is as follows.

$$\frac{\frac{\frac{\overline{T \langle C \rangle} \quad \overline{T \langle C \rangle \Leftrightarrow C}}{C}}{\overline{T \langle C \rangle} \quad \overline{T \langle C \rangle \Rightarrow \perp}}}{\perp}}{\overline{T \langle C \rangle \Rightarrow \perp}}$$

The argument uses *modus ponens* (MP) for the biconditional of the  $T$ -Schema (and for the conditional of the Curry sentence itself), and the last line discharges both occurrences of the assumption  $T \langle C \rangle$ . This argument gives us an unconditional proof of  $T \langle C \rangle \Rightarrow \perp$ , that is,  $C$ . This and the  $T$ -Schema then deliver  $T \langle C \rangle$  by MP, whence applying MP again delivers  $\perp$ .

Most of the solutions to the Paradox which endorse the  $T$ -Schema deny Contraction. That is, an application of  $\Rightarrow$ -introduction cannot discharge multiple assumptions of the antecedent. This solution may be implemented in a contraction-free logic, such as a weak relevant logic. In the axiomatic forms of these logics, Contraction is equivalent to the validity of the inference:  $A \Rightarrow (A \Rightarrow B) \vdash A \Rightarrow B$ .

Now, to many eyes it would appear to be the case that the Liar Paradox and Curry's Paradox are of a piece. The Principle of Uniform Solution (*Same Kind of Paradox, Same Kind of Solution*)<sup>13</sup> would therefore dictate that they have the same kind of solution. But rejecting Contraction has nothing to do with a dialethic solution to the Liar Paradox. So there seems to be a problem here.

Whether or not the Curry paradox and the Liar paradox are of a kind is a thorny question, but given the importance of the Curry paradox to the matters at hand, it is an important one. Its discussion would, however, take

<sup>13</sup>See Priest (2002b), 11.5.

us off at a tangent to matters immediately to hand. So I will reserve a discussion of the matter till an appendix to this paper. Here, I simply note that Strategy 1 solves the Curry paradox as well as the Liar: the (bi)conditional of the  $T$ -scheme does not detach. The non-triviality argument of Section 3 demonstrates the non-triviality of the Curry sentence just as much as the Liar sentence (even if the language contains a detachable conditional, as we shall see). So if the Liar and Curry are of a kind, Strategy 1 gives the two the same solution, as required by the Principle of Uniform Solution. In particular, we can just take  $T\langle C \rangle$  to be both true and false. The failure of detachment does the rest.

Moreover, even if the two paradoxes are not the same kind, this solution is simpler than the non-contracting solution. The explanations of the failure of conditional contraction—even if they are right—all involve a world semantics, indeed a semantics with impossible worlds; it is clear that this is much more complex. And there is, presumably, nothing problematic about different kinds of paradoxes having the same kind of solution. That is, one should not expect to have *Different Kind of Paradox, Different Kind of Solution*. The same drug can, after all, cure different illnesses. Thus,  $IC$  argues that *both* the paradoxes of self-reference *and* Zeno’s Arrow Paradox—quite different sorts of paradox—may be solved by endorsing dialetheism. Even if one holds that the Arrow Paradox is not to be solved by deploying dialetheism, one would hardly argue that this cannot be the solution since dialetheism solves the paradoxes of self-reference.

## 4.2 Validity Curry

There are versions of the Curry Paradox which concern validity. These, too, are handled by Strategy 1.

Given some very weak assumptions about the conditional,<sup>14</sup> the contraction inference,  $A \Rightarrow (A \Rightarrow B) \vdash A \Rightarrow B$ , is equivalent to the validity of “pseudo *modus-ponens*” (PMP):  $(A \wedge (A \Rightarrow B)) \Rightarrow B$ . For suppose that  $A \Rightarrow (A \Rightarrow B)$ . Since  $A \Rightarrow A$ , it follows that  $A \Rightarrow (A \wedge (A \Rightarrow B))$ . Given PMP and the transitivity of  $\Rightarrow$ ,  $A \Rightarrow B$  follows. Conversely, let us write  $C$  for  $A \wedge (A \Rightarrow B)$ . Since  $C \Rightarrow A$ ,  $(A \Rightarrow B) \Rightarrow (C \Rightarrow B)$ . But  $C \Rightarrow (A \Rightarrow B)$ . Hence by transitivity,  $C \Rightarrow (C \Rightarrow B)$ . By the contraction inference,  $C \Rightarrow B$ .

---

<sup>14</sup>Namely, that it is at least as strong as the basic relevant logic  $B$ . See Priest (2008), ch. 9.

Now, let  $\rightarrow$  express the relationship that  $A$  bears to  $B$  when  $A$  entails  $B$  (that  $A$  entails that  $B$ ).<sup>15</sup> Then if MP is valid for the conditional of the  $T$ -Schema,  $\Rightarrow$ , we have  $(A \wedge (A \Rightarrow B)) \rightarrow B$ . It is natural to suppose that valid inferences are truth preserving. That is,  $C \rightarrow D \vdash C \Rightarrow D$ , and if this is the case, we have  $(A \wedge (A \Rightarrow B)) \Rightarrow B$ , and we are back with the Curry paradox. One solution to the paradox is to reject the necessary connection between  $\rightarrow$  and  $\Rightarrow$ . But if the conditional of the  $T$ -Schema is material, the failure of MP breaks the argument at the first step.

There are also versions of the Curry paradox involving a validity predicate directly.<sup>16</sup> These may also be solved by Strategy 1.

Let us write  $V(\langle A \rangle, \langle B \rangle)$  to express the claim that the one-premise inference from  $A$  to  $B$  is valid. Corresponding to the  $T$ -Schema, we have what we might call the *naive V-Schema*:

$$V(\langle A \rangle, \langle B \rangle) \Leftrightarrow (A \rightarrow B)$$

Given techniques of self-reference, we can construct a sentence  $D$ , of the form  $V(\langle D \rangle, \langle \perp \rangle)$ . We then have an argument whose core is as follows:

$$\frac{\overline{D} \quad \frac{\overline{V(\langle D \rangle, \langle \perp \rangle)} \quad \overline{V(\langle D \rangle, \langle \perp \rangle) \Leftrightarrow (D \rightarrow \perp)}}{D \rightarrow \perp}}{\overline{D} \quad \frac{\perp \quad \overline{D \rightarrow \perp} \quad \overline{V(\langle D \rangle, \langle \perp \rangle) \Leftrightarrow (D \rightarrow \perp)}}{V(\langle D \rangle, \langle \perp \rangle)}}$$

The line after the naked  $\perp$  is justified by the deduction of  $\perp$  from the assumptions  $D$ , and discharges both of them. Hence we have an unconditional

---

<sup>15</sup>Formally,  $\rightarrow$  may be taken to satisfy the two conditions:

$$\frac{\overline{A} \quad \vdots \quad B}{A \rightarrow B} \quad \frac{A \quad A \rightarrow B}{B}$$

where the deduction in the first rule discharges  $A$ , and there are no other assumptions. The rules are exactly what one would expect if  $A \rightarrow B$  is simply an object-language way of expressing what a metalanguage deduction from  $A$  to  $B$  expresses.

<sup>16</sup>See, e.g., Beall and Murzi (2013).



proof of  $V(\langle D \rangle, \langle \perp \rangle)$ , which is  $D$ . Applying the  $V$ -Schema again gives us  $D \multimap \perp$ , and this, together with  $D$  gives us  $\perp$ .

As in the standard Curry paradox, one kind of solution rejects Contraction. The line after the naked  $\perp$  can discharge only one occurrence of  $D$ . How to articulate this solution, and, in particular, how to relate it to a semantic definition of validity, is, however, a sensitive matter.<sup>17</sup> A simpler solution is to take the conditional of the  $V$ -Schema to be material. This breaks the arguments where *modus ponens* is applied to it.

### 4.3 Infinitary Curry

Another version of the Curry Paradox which Strategy 1 easily solves is given in Restall (2006). It is assumed that one can formulate the  $T$ -Schema as two inferences:

$$\tau \wedge T \langle A \rangle \vdash A \qquad \tau \wedge A \vdash T \langle A \rangle$$

Thus, the  $T$ -Schema is taken to be enthymematically valid: some true background information,  $\tau$ , may be necessary. (Of course, if one takes it to hold non-enthymematically, then the enthymematic form will hold for any  $\tau$ .) Restall also assumes the existence of a conditional,  $\rightarrow$ , satisfying the inferences:

$$\frac{A \wedge B \vdash C}{A \vdash B \rightarrow C} \qquad \frac{A \vdash B \rightarrow C}{A \wedge B \vdash C}$$

Now let  $C$  be  $T \langle C \rangle \rightarrow \perp$ . Then we have:

$$\frac{\tau \wedge T \langle C \rangle \vdash C}{\tau \wedge T \langle C \rangle \vdash T \langle C \rangle \rightarrow \perp} \\ \frac{\tau \wedge T \langle C \rangle \wedge T \langle C \rangle \vdash \perp}{\tau \wedge T \langle C \rangle \vdash \perp} \\ \tau \vdash T \langle C \rangle \rightarrow \perp$$

The rest of the argument then proceeds much as before, and we end up with a proof of the unacceptable  $\tau \vdash \perp$ .

Perhaps the most obvious way to object to this argument is to deny the existence of a meaningful connective,  $\rightarrow$ , satisfying these introduction and elimination rules. However, Restall points out that it would appear that one can *define* such a conditional,  $B \rightarrow C$ , as:

---

<sup>17</sup>See Priest (2015a).

- $\bigvee\{A : A \wedge B \vdash C\}$

Here,  $\bigvee$  is an infinitary disjunction. We may show that this satisfies the rules as follows. Suppose that  $A' \wedge B \vdash C$ . Then  $A' \vdash \bigvee\{A : A \wedge B \vdash C\}$ , since  $A' \in \{A : A \wedge B \vdash C\}$ . Conversely, suppose that  $A' \vdash B \rightarrow C$ ; that is,  $A' \vdash \bigvee\{A : A \wedge B \vdash C\}$ . Then  $A' \wedge B \vdash B \wedge \bigvee\{A : A \wedge B \vdash C\}$ , and by the infinitary distribution of conjunction over disjunction,  $A' \wedge B \vdash \bigvee\{A \wedge B : A \wedge B \vdash C\}$ . By the usual properties of disjunction,  $\bigvee\{A \wedge B : A \wedge B \vdash C\} \vdash C$ . So  $A' \wedge B \vdash C$ , by transitivity.

Now, there are several ways one might object to this argument. The most obvious is to reject infinitary operators altogether.  $\rightarrow$  is just not a humanly meaningful connective. One might reply that it is meaningful in a more abstract sense. But this reply is also moot, since the definition of  $\rightarrow$  is circular. The set of which  $B \rightarrow C$  is the infinite disjunction is defined in terms of  $\vdash$ . It therefore presupposes that this is well defined. But the rules for  $\vdash$  include the rules for  $\rightarrow$ , and therefore presuppose, amongst other things, that the set which defines a conditional is well defined. The definition of  $\rightarrow$  is therefore a creative one, and does not guarantee it sense. As far as this goes, it could be a tonk-like connective.

These considerations provide a solution to this Curry problem; but one can, of course, debate them further. It is clear that the matter is not a simple one. By contrast, Strategy 1 provides a simple and straightforward solution. We have only  $T \langle A \rangle \equiv A$ , or maybe enthymematically,  $\tau \vdash T \langle A \rangle \equiv A$ . We do not have  $(\tau \wedge)T \langle A \rangle \dashv\vdash A(\wedge\tau)$ , so the argument fails at the first step since  $\equiv$  does not detach.

Indeed, given this solution, non-triviality can easily be proved. Take the language to be that of the relevant logic  $R$  with the logical constant  $t$ , plus arithmetical vocabulary and the truth predicate. Consider a Routley/Meyer interpretation, where the interpretation of the arithmetic language at the base world is classical, and where the base world also verifies all sentences of the form  $T \langle A \rangle!$ . This validates the material  $T$ -Schema. In  $R$ ,  $\rightarrow$  is detachable and contracting, and the conditional  $A \mapsto B$ , defined as  $(A \wedge t) \rightarrow B$ , satisfies the Restall rules.<sup>18</sup>

---

<sup>18</sup>As shown in Meyer (1973), this is the conditional of intuitionist logic.

## 4.4 An Extended Paradox

Yet another paradox which is handled simply by Strategy 1, is an “extended paradox” first formulated by Smiley.<sup>19</sup> Define a function,  $Val$ , as follows, where  $T$  is the truth predicate and  $F$  is the falsity predicate:

- $Val \langle A \rangle = \{1\}$  iff  $T \langle A \rangle \wedge \neg F \langle A \rangle$
- $Val \langle A \rangle = \{1, 0\}$  iff  $T \langle A \rangle \wedge F \langle A \rangle$
- $Val \langle A \rangle = \{0\}$  iff  $\neg T \langle A \rangle \wedge F \langle A \rangle$

Given that  $T \langle A \rangle \vee \neg T \langle A \rangle$  and  $F \langle A \rangle \vee \neg F \langle A \rangle$ , one of these cases must hold; and the three cases seem to be disjoint. Thus, e.g., if the first and the second held, we would have  $F \langle A \rangle \wedge \neg F \langle A \rangle$ . Now, by techniques of self-reference, form a sentence  $S$ , of the form  $Val \langle S \rangle = \{0\}$ . The  $T$ -Schema gives us that:

- $T \langle S \rangle \Leftrightarrow Val \langle S \rangle = \{0\}$

We have that  $Val \langle S \rangle = \{1\} \vee Val \langle S \rangle = \{1, 0\} \vee Val \langle S \rangle = \{0\}$ . In the first two cases,  $T \langle S \rangle$ . So by the  $T$ -Schema,  $Val \langle S \rangle = \{0\}$ . So either  $\{1\} = \{0\}$ , or  $\{1\} = \{1, 0\}$ , and so  $0 = 1$ . In the third case, by the  $T$ -Schema,  $T \langle S \rangle$ . Hence,  $Val \langle S \rangle = \{1\} \vee Val \langle S \rangle = \{1, 0\}$ , so again  $0 = 1$ . Hence,  $0 = 1$ , and all sentences have the same value.

Priest (1993) and *IC*, 20.3, argue that the argument fails, since the three cases in the definition of  $Val$  are not exclusive. In paradoxical cases, two or more of the defining cases may hold. (Thus, if  $L$  is  $\neg T \langle L \rangle$ , we have  $T \langle L \rangle \wedge \neg T \langle L \rangle$ . But if  $L$  is not true, it is false, so we have  $F \langle A \rangle$  as well, and we are in the second and third cases.) So  $Val$  is ill-defined. Strategy 1 gives an even simpler solution, however. If  $\Leftrightarrow$  is  $\equiv$ , the detachments involved in applying the  $T$ -Schema break down.

## 4.5 Denotation Paradoxes

Let us now turn to the paradoxes of denotation. These are a very distinctive subclass of the self-referential paradoxes: first, because they require, essentially, description terms of some kind, and some principle concerning how these behave; secondly, because the standard paradoxical arguments in these cases do not go via biconditionals of the form  $A \Leftrightarrow \neg A$ , but give *independent*

---

<sup>19</sup>Smiley (1993).

arguments for each of  $A$  and  $\neg A$ . Such arguments do not then need to invoke the LEM essentially.<sup>20</sup>

For the most part, the paradoxes of denotation are handled very easily by a dialethic strategy. The arguments are sound, and the contradictory conclusions simply true. But there is a paradox in this family which cannot obviously be solved in this way. This is an argument first produced by Hilbert and Bernays to show that a consistent theory cannot contain its own denotation function (just as Tarski employed the Liar Paradox to show that a consistent theory cannot contain its own truth predicate).<sup>21</sup>

An informal version of the paradox goes as follows. Consider a term,  $t$ , of the form: the denotation of ‘ $t$  plus one’. ‘ $t$ ’ denotes  $t$ , but it also denotes  $t + 1$ . Hence  $t = t + 1$ , and so  $0 = 1$ .

Formally, let  $f$  be any one-place functional term. By techniques of self-reference, we can find a term,  $t$ , such that  $t = f(\langle t \rangle)$ .<sup>22</sup> Now let  $\epsilon$  be an indefinite description operator, satisfying the *Description Principle*:  $\exists xA \vdash A_x(\epsilon xA)$ .<sup>23</sup> Let  $D$  be the denotation relation, and consider the functional term  $f(\epsilon yD(x, y))$ . By what we have just seen, we can find a term,  $t$ , such that:

$$(0) \quad t = f(\epsilon yD(\langle t \rangle, y))$$

We can now reason as follows:

---

<sup>20</sup>See Priest (2006a).

<sup>21</sup>For full discussion and references, see Priest (2005), ch. 8.

<sup>22</sup>Thus, suppose that we are operating in arithmetic, and naming is given by gödel coding. Suppose, also, that we have a function symbol,  $\delta$ , such that for any term,  $t$ ,  $\delta(\langle t \rangle)$  is the diagonalisation of  $t$ —that is, (the code of) the term  $t$  with the code of term  $t$  replacing every free variable. Consider the term  $f(\delta(x))$ ; call this  $s$ . Its diagonalisation is  $f(\delta(\langle s \rangle))$ ; call this  $t$ . Then  $\delta(\langle s \rangle) = \langle t \rangle$ . So  $f(\delta(\langle s \rangle)) = f \langle t \rangle$ . So  $t = f \langle t \rangle$ .

<sup>23</sup>For the application at hand, it might well be thought that a definite description operator could be used. However, to keep matters as simple as possible, we use indefinite descriptions.

(1)	$t = t$	Identity
(2)	$D(\langle t \rangle, t)$	Denotation Schema (and MP)
(3)	$\exists y D(\langle t \rangle, y)$	Generalisation
(4)	$D(\langle t \rangle, \epsilon y D(\langle t \rangle, y))$	Description Principle
(5)	$t = \epsilon y D(\langle t \rangle, y)$	Denotation Scheme (and MP)
(6)	$t = f(t)$	Substituting (5) in (0)

In particular, suppose we are quantifying over natural numbers, and for  $f$  take the “zero function”,  $\zeta$ :  $\zeta(0) = 1$ , and  $\zeta(n) = 0$  if  $n > 0$ . Then  $t = \zeta(t)$ . If  $t = 0$  then  $\zeta(t) = 1$ ; so  $0=1$ . If  $t > 0$  then  $\zeta(t) = 0 = t$ , so  $0 = 1$ , again.

Now, there are various solutions to this paradox one might attempt. One is to reject the  $D$ -Schema, by stratification, or in some other way. That is not an option if we accept the naive semantic principles. Another is to hold that  $t$  has no denotation. However, this does not work, since we can arrange for all  $\epsilon$ -terms to have a denotation (*à la* Hilbert). Priest (2005), ch. 8, argues that the correct solution is to take  $t$  to have multiple denotations. The natural semantics shows that the Substitutivity of Identicals (step (6)) then fails. Taking the  $D$ -Schema to be a material conditional provides a much simpler solution, however. The argument fails at lines (2) and (5).

Indeed, the non-triviality proof of this approach is as simple as before. Take a classical model of arithmetic plus the denotation predicate, in which arithmetic is standard and the extension and anti-extension of  $D$  are both the set of all pairs of objects in the domain. This validates the  $D$ -Schema. Now extend this to a model of Hilbert’s  $\epsilon$ -operator (by the addition of a choice function).<sup>24</sup> This still validates the  $D$ -Schema (however  $\epsilon$ -terms are coded); but it also validates the Description Principle, and is a conservative extension.

## 4.6 Kripke’s Paradox

Another paradox of self-reference to note in this context is one proposed by Kripke (2011). Let  $k$  be the set of all times,  $t$ , when I am thinking about a set of times of which  $t$  is not a member. I can think about this at time  $\tau$ .  $\tau$  is and is not, then, a member of  $k$ . This is obviously a paradox of a kind similar to Russell’s paradox, but it uses an intentional notion, *think of*, and standard

---

<sup>24</sup>As in Leisenring (1969).

solutions to the set-theoretic paradoxes do not apply to it. In particular, since—identifying times with real numbers— $k$  is a set of real numbers, one cannot say that it does not exist. And even if it did, this appears to be irrelevant, since I obviously can think about non-existent objects, such as Zeus.

Though we will not turn to set theory till the next part of this essay, let us look at this paradox and its relation to Strategy 1 now. Much of the detail of Kripke’s example is, in fact, unnecessary to the paradox. In Priest (2015b), I strip it down to the following.

Let  $\theta(x)$  be ‘I am thinking about  $x$  (now)’. Let  $k = \{n < 1 : \exists s \subseteq 1(\theta(s) \wedge n \notin s)\}$ , where the variable  $n$  ranges over natural numbers. Clearly either  $k = \emptyset$  or  $k = \{0\}$ . We have:

- $0 \in k \Leftrightarrow \exists s \subseteq 1(\theta(s) \wedge 0 \notin s)$

Given the empirical premise that I am (now) thinking about  $k$  and only  $k$ , then  $\theta(x)$  iff  $x = k$ . Hence:

- $0 \in k \Leftrightarrow \theta(k) \wedge 0 \notin k$

Given that  $\Leftrightarrow$  is detachable, we then have that  $0 \in k!$ . So either  $0 \in \emptyset$  or  $0 \notin \{0\}$ .

Now, while there are ways in which we can make sense of this,<sup>25</sup> it is hard to suppose that such simple and small sets are contradictory. However, if  $\Leftrightarrow$  is the material biconditional, we have only that  $0 \in k! \vee \neg\theta(k)$ . And given the empirical premise:  $0 \in k! \vee \neg\theta(k)!$ . The purely set-theoretic statement can, then, have a quite consistent value (*false*), provided that it is the second disjunct which is true: I am and am not (now) thinking about  $k$ .

Odd though this may seem at first, it is of a piece with many of the other paradoxes of self-reference. Thus, König’s paradox tells us that a certain number both can and cannot be referred to. To think of something is to refer to it mentally. So, in a similar way,  $k$  both can and cannot be thought of.

## 5 Endorsement

Let us now look at the negative side of the ledger more closely: the failure of  $\supset$  to detach, as it would seem to have to, if truth is to play a role in

---

<sup>25</sup>See Priest (2015b).

endorsement.

## 5.1 Transparent Truth

A very simple way of solving the problem is to endorse the transparency of truth, that is, the view that  $A$  and  $T\langle A \rangle$  are intersubstitutable in all contexts:

$$\frac{B_C(A)}{\overline{B_C(T\langle A \rangle)}}$$

where the double line indicates a two-way inference. Let us call this pair of rules *Transparency*. An obvious special case of Transparency is what we might call the *T*-rules:

$$\frac{A}{\overline{T\langle A \rangle}}$$

These rules are weaker, though. They do not entail the “negative” *T*-rules:

$$\frac{\neg A}{\overline{\neg T\langle A \rangle}}$$

As is not hard to show, the *T*-rules and negated *T*-rules do deliver intersubstitutivity in all extensional contexts (i.e., those of *LP*). In particular, since  $\models A \equiv A$ , they deliver the material *T*-Schema  $\models A \equiv T\langle A \rangle$ . They do not deliver a transparent truth predicate if the contexts involved are intentional, however, such as  $\Box$  or a relevant  $\rightarrow$ . The rules require  $A$  and  $T\langle A \rangle$  to relate to the same truth values at the “base world” of a world-semantics. This does not guarantee intersubstitutivity in intensional contexts (i.e., ones whose semantic evaluation at a world requires a world-shift). For that, we need to enforce the connection between  $A$  and  $T\langle A \rangle$  at other worlds.

Now, Transparency—or just the positive *T*-rules—does solve the problem of endorsement. Thus, in the first example of 3.2, once we have inferred that ‘There is a time for every purpose under heaven’ is true, the result follows. Similarly for the second example of 3.2.

Matters are the same when the truth predicate is used in an infinitary act of endorsement. Thus, in the third example of 3.2, once we have inferred  $T\langle \forall x x' \neq 0 \rangle$ , we can use the rules to infer  $\forall x x' \neq 0$ .

The addition of Transparency, moreover, produces a non-trivial extension in the extensional language. This can be proved by a standard fixed-point

construction. We start with a model of arithmetic, where both the extension and anti-extension of  $T$  are the whole domain,  $N$ . We then ascend the ordinals, at successor stages setting the extension and anti-extension of  $T$  to be the set of (codes of) things true/false at the previous stage; and at limit ordinals taking intersections. Because of the Monotonicity Lemma, the extension and anti-extension of the truth predicate are non-increasing. Hence, we reach a fixed-point interpretation, where the extension and anti-extension of  $T$  are exactly the (codes of) the things true/false in the interpretation. Full details can be found in Priest (2002a), 8.1.

What effect does the addition of Transparency have on the positives for Strategy 1 discussed in Section 4? The Curry paradox and the Validity Curry are solved exactly as before, as is shown by the fixed-point construction: detachment fails for the relevant conditional. The addition has no relevance to Kripke's paradox. It does spoil the solution to the paradoxes of 4.3-4.5, however, since these arguments can all be run using the positive  $T$ -rules, or their analogue for denotation:

$$\frac{t = c}{\overline{D(\langle t \rangle, c)}}$$

The only solutions for these would then seem to be of a different kind: the ones indicated in the discussion of those sections.

## 5.2 Propositional Quantification

A rather different way of addressing the concerns engendered by endorsement is to cut out the middle-man entirely, and use propositional quantification.

For example, in the first case of 3.2, we can then say:

- $\forall p(\text{the Bible says that } p \Rightarrow p)$

from which, by universal instantiation, we can infer: The Bible says that there is a time for every purpose under heaven  $\Rightarrow$  there is a time for every purpose under heaven. Given that we may move from the antecedent to the consequent in some way, we can then infer that there is a time for every purpose under heaven.

The second example of 3.2 can be treated in the same way. We have:

- $\forall p(p \Rightarrow \text{the Bible says that } p)$ .



From which we can infer: People have evolved  $\Rightarrow$  the Bible says that people have evolved. *Modus ponens* does the rest. The third example of 3.2 can also be dealt with in the same way. The endorsement is expressed as:

- $\forall p$ (some axiom of  $PA$  says that  $p \Rightarrow p$ )

Given only that an axiom of  $PA$  says that  $\forall x x' \neq 0$ , the result follows.

The machinery of propositional quantification is easily added to that at hand.<sup>26</sup> We add two propositional constants, 1 and 0, a stock of propositional variables, and the propositional quantifiers,  $\forall$  and  $\exists$ . The truth/falsity conditions are then extended as follows:

- $\Vdash^+ 1$  and  $\Vdash^- 0$
- $\Vdash^+ \exists pA$  iff  $\Vdash^+ A_p(1)$  or  $\Vdash^+ A_p(0)$
- $\Vdash^- \exists pA$  iff  $\Vdash^- A_p(1)$  and  $\Vdash^- A_p(0)$
- $\Vdash^+ \forall pA$  iff  $\Vdash^+ A_p(1)$  and  $\Vdash^+ A_p(0)$
- $\Vdash^- \forall pA$  iff  $\Vdash^- A_p(1)$  or  $\Vdash^- A_p(0)$

Note that we do not insist that 1 is not false (as well as true), or that 0 is not true (as well as false). The proof of substitutivity for material equivalents is easily extended to the new context.

It is now easy enough to establish that the quantifiers behave as one would wish. Let  $B$  be any formula. Then in any interpretation  $\Vdash^+ B \equiv 1$  or  $\Vdash^+ B \equiv 0$ . Suppose the former; the latter is similar. Then if  $\Vdash^+ \forall pA$ ,  $\Vdash^+ A_p(1)$ . So by the substitutivity of equivalents,  $\Vdash^+ A_p(B)$ . Hence,  $\forall pA \models A_p(B)$ . By an analagous argument,  $A_p(B) \models \exists pA$ .<sup>27</sup>

The addition of the machinery of propositional quantification obviously produces a conservative extension (since any counter-model of an inference in the original language can be extended to a counter-model in the extended language). Hence, the non-triviality arguments already given carry over to the new context. And, unlike the addition of  $T$ -rules the extension does nothing to destabilise any of the paradox solutions discussed in Section 4.

---

<sup>26</sup>Many people feel uncomfortable about propositional quantification. Some of this may be eased just by noting that the propositional quantifiers are just zero-adic second-order quantifiers.

<sup>27</sup>If there are intensional operators in the language, a more sophisticated semantics is required. This can be found in Priest (2009b), 2.1.

## 6 An Additional Conditional?

So much for the obvious issues. There are, however, other, less obvious aspects of Strategy 1 to be looked at. In this and the next two sections we consider three of these.

### 6.1 Truth and a Detachable Conditional

The conditional of the  $T$ -Schema is, according to Strategy 1, material. This does not prevent there being a detachable conditional,  $\rightarrow$ , in the language (such as the connective  $\succrightarrow$  of 4.2). The Curry argument still fails, even if  $\rightarrow$  contracts. Thus, we can formulate the sentence,  $C$ , of the form  $T \langle C \rangle \rightarrow \perp$ , but the Curry argument of 4.1 still fails at the MP steps for the  $T$ -Schema.

Indeed, such a conditional can be added conservatively to the language. We take the interpretation of 3.1 as the base world of model for  $S5$ . The strict conditional (which detaches and contracts) is given truth conditions in the usual way ( $A \rightarrow B$  is true at a world iff at every world where  $A$  is true,  $B$  is true); but only the material  $T$ -Schema is delivered. (In particular, it is not delivered in the form of a strict biconditional.)

A fully transparent truth predicate does, however, rule out a detachable conditional—at least if it validates  $A \leftrightarrow A$ . For then, transparency delivers  $A \leftrightarrow T \langle A \rangle$ , and we move from Strategy 1 to Strategy 2. Thus, Strategy 1 and a transparent truth predicate *requires* that there be no such conditional.

### 6.2 Quasi-Validity

Having no other conditional raises a substantial problem, however. The material conditional does not detach. Transparency may solve the problem as concerns the  $T$ -Schema, but what of all the other places where we deploy a conditional? We can never apply *modus ponens*!

There is a way in which one might attempt to get around this problem, however.<sup>28</sup> This is to take material detachment, and, more generally, classically valid inferences that are not valid in  $LP$ —*quasi-valid* inferences—as acceptable *default* inferences. Counterexamples to these inferences arise, we know, only when dialetheias occur at strategic places in the reasoning. Given that in ordinary reasoning such things are statistically uncommon, they have

---

<sup>28</sup>Which might also be used to handle some occasions where one might want to detach with a material  $T$ -Schema.

low *a priori* probability; we are therefore justified in proceeding on the assumption that we are not dealing with them, and using classical logic. Given some dog which we have no reason to believe to be unusual, we are entitled to take it to have four legs—until we learn otherwise—since dogs are normally four-legged. Similarly, when reasoning about a situation, we are entitled to take it to be consistent, until we learn otherwise.<sup>29</sup>

Of course, sometimes we will learn otherwise: we may obtain evidence that dialetheias are involved in the reasoning. We will then have to review and possibly revise our belief in the things established with quasi-valid reasoning. What has been established is therefore fallible. We do not have to reject everything established in this way, however. Thus, if we had inferred  $B$  from  $A$  and  $A \supset B$ , and we have no particular reason to suppose that  $A$  is a dialetheia, there is no reason to revise our belief in  $B$ . In other words, we may move from an assumption of *global* consistency to one of *local* consistency. Of course, disentangling where and how things we have reason to believe dialetheic are involved in our reasoning will be a messy and, no doubt, itself fallible matter. Perhaps all one can do is leave things as a matter of informal judgment.<sup>30</sup>

However, there is a way of turning this whole process of reasoning into something more determinate. This is to use a non-monotonic logic which implements the default of consistency (both locally and globally). A non-monotonic logic is one where a valid inference may fail, given further information. That is, we may have  $\Sigma \vdash A$  but  $\Sigma \cup \Pi \not\vdash A$ . Fallibility due to new information is, then, built into the very notion of logical consequence. I will describe one such logic, *LPm*, in the next subsection.<sup>31</sup>

---

<sup>29</sup>The strategy is defended in detail in *IC*, ch. 8, and 19.10. If one endorses a dialetheic solution to the sorites paradox, as I will discuss in 15.6, it might be objected that since vague predicates are a commonplace in natural language, so would dialetheias be; and contradiction-freedom is not the norm. This does not follow, however. Take a vague predicate, such as ‘red’. The vast majority of objects are not red, and consistently so. It follows that the collection of objects which are red and not red are a very small proportion of the total. Again, according to *IC*, motion produces contradictions; since motion is everywhere, it might be argued that contradictions are ubiquitous. The contradictions in motion concern instantaneous states, however—or at least states of very short duration; and normal reasoning rarely concerns such states.

<sup>30</sup>If I understand it right, this is the view recently adopted by Beall. See, e.g., Beall (2011) and (2013b).

<sup>31</sup>For what follows, see *IC*, ch. 16.

### 6.3 $LPm$

Given any  $LP$  interpretation,  $\mathfrak{M}$ , let:

- $\mathfrak{M}! = \{A : A \text{ is of the form } Pk_{d_1} \dots k_{d_n}, \text{ and } \mathfrak{M} \text{ is a model of } A!\}$

Here the  $k_d$ 's are the canonical names for the members of the domain.  $\mathfrak{M}!$  records the set of contradictory atomic “facts” in  $\mathfrak{M}$ . This provides a simple and natural measure of its inconsistency. Given two  $LP$  interpretations,  $\mathfrak{M}$ ,  $\mathfrak{N}$ , define  $\mathfrak{M} <_m \mathfrak{N}$  to mean that  $\mathfrak{M}! \subset \mathfrak{N}!$ . Here, ‘ $\subset$ ’ is *proper* subsethood.  $\mathfrak{M} <_m \mathfrak{N}$  expresses the fact that that  $\mathfrak{M}$  is *more* consistent than  $\mathfrak{N}$ . Say that  $\mathfrak{M}$  is a *minimally inconsistent* (mi) model of  $\Sigma$  if:

- $\mathfrak{M}$  is a model of  $\Sigma$
- and if  $\mathfrak{N} <_m \mathfrak{M}$  then  $\mathfrak{N}$  is not a model of  $\Sigma$

Finally, say that  $A$  is a *minimally inconsistent consequence* of  $\Sigma$ ,  $\Sigma \models_m A$ , iff:

- every mi model of  $\Sigma$  is a model of  $A$

Thus, to determine the mi consequences of a set,  $\Sigma$ , we ignore all those interpretations that are more inconsistent than  $\Sigma$  requires us to suppose. In this way, we implement the default assumption of consistency.

$LPm$  validates inferences that  $LP$  does not. Thus, let us write  $p$  for  $Pa$ ,  $q$  for  $Qb$ , and  $r$  for  $Rc$ . Then  $\{p, \neg p \vee q\} \models_m q$ , since the mi models of the premises are just the classical models. Similarly,  $r$  is irrelevant to the inference, so we have  $\{r!, p, \neg p \vee q\} \models_m q$ . The premises require  $r$  to be inconsistent, but not  $p$ .  $\models_m$  bears all the marks of a non-monotonic consequence relation. Additional information can invalidate an inference, and it is not closed under uniform substitution. Thus, it is not difficult to show that  $\{\neg p, p, \neg p \vee q\} \not\models_m q$  and  $\{p, \neg p \vee q!\} \not\models_m q!$  Crucially, however, if  $\Sigma$  is consistent (either classically or in  $LP$ , it makes not difference, as we observed in fn. 10) then  $\Sigma \models_m A$  iff  $\Sigma \models_{CL} A$ , since the mi models of  $\Sigma$  are just its classical models. ( $\mathfrak{M}$  is classical iff  $\mathfrak{M}! = \emptyset$ .) We see, at any rate, that, using  $LPm$ , we can use the disjunctive syllogism (detachment for  $\supset$ ) when reasoning from (locally) consistent information.

On a more technical note, *IC*, ch. 16, claims that if  $\Sigma$  is non-trivial under  $LP$  consequence, it is non-trivial under  $LPm$  consequence. This is called *Reassurance*. The claim is false, as Crabbé (2011) shows. Thus, if  $\Sigma =$

$\{\forall x(Px! \vee Qx!), \exists x(Px! \wedge Qx!)\}$ ,  $\Sigma$  does not entail  $\forall xPx$  in  $LP$ , but its only mi model (up to isomorphism) has one element which satisfies both  $Px!$  and  $Qx!$ . It is not difficult to show that in this interpretation all sentences are true.<sup>32</sup>

Now, even without Reassurance,  $LPm$  would seem to do everything that one would like: it delivers a more generous notion of consequence than  $LP$ , where irrelevant contradictions do not invalidate classical inferences, and which delivers all classical consequences given consistent premises.<sup>33</sup> None the less, there is a very simple way to obtain Reassurance.<sup>34</sup> One can simply change the definition of mi consequence to be as follows.  $\Sigma \vDash_m A$  iff:

- some mi model of  $\Sigma$  is non-trivial, and every mi model of  $\Sigma$  is a model of  $A$
- every mi model of  $\Sigma$  is trivial,<sup>35</sup> and  $\Sigma \vDash A$

Reassurance follows simply, and  $\vDash_m$  still has all the crucial properties.

I note that an approach to minimal inconsistency, slightly different from the one just described, is endorsed by Batens (who founded this whole area of research). This has consequences which might well be thought to make it more attractive. The approach defines  $\mathfrak{M}!$ , not as above, but as  $\{[\exists](A \wedge \neg A) : \mathfrak{M} \text{ is a model of } [\exists](A \wedge \neg A)\}$ , where, here,  $[\exists]B$  is the sentence obtained by prefixing  $B$  with a particular quantifier,  $\exists x$ , for every variable,  $x$ , free in  $B$ .

---

<sup>32</sup>The mistake arose as follows. Originally (Priest (1991)), the definition of  $\mathfrak{M} <_m \mathfrak{N}$  required  $\mathfrak{M}$  and  $\mathfrak{N}$  to have the same domain. For this definition of  $<_m$ , the proof is correct. However, as Batens observed, given this definition, there can be consistent sets which have non-classical mi models. Thus,  $\{\exists xPx, \exists x\neg Px\}$  has a one-element model in which the only object satisfies  $Px!$ . It is not difficult to see that this is mi. When, to avoid this consequence, the definition of  $<_m$  was changed to the present one in the second edition of *IC*, I failed to observe that the proof of Lemma 3 (p. 229) now fails.

<sup>33</sup>The reason given in *IC*, 16.6, for the desirability of Reassurance is as follows. Taking triviality to be a mark of incoherence, Reassurance guarantees that a coherent situation will never be turned into an incoherent one under  $LPm$ . This may be more than is required, though. It might be quite sufficient if mostly, or normally,  $LPm$  does not turn a non-trivial situation into a trivial one. If there are some exceptions, and  $LPm$  is otherwise robust, we might take the triviality exposed to speak against the coherence of the original situation.

<sup>34</sup>Crabbé (2011) and (2012) also contain observations concerning how Reassurance can be restored.

<sup>35</sup>Which of course includes the case in which it has none.

Using this definition, one can prove that for any model,  $\mathfrak{M}$ , there is an  $\text{mi } \mathfrak{N} <_m \mathfrak{M}$ ; so, in particular, Reassurance holds.<sup>36</sup>

Let me end this section with a comment on truth. As we saw in 3.1, there is a model of the material  $T$ -schema which guarantees its non-triviality. In this, the extension and the anti-extension of the truth predicate are both the whole domain. In other words, every sentence of the form  $T \langle A \rangle!$  is true. As an account of truth itself, this model is completely implausible: truth should not be that inconsistent! What is a more plausible model like? One answer is provided by the construction of 5.1. We saw there how to construct a model in which truth is inconsistent only for ungrounded sentences. This construction does validate the  $T$ -inferences, however—which one may not want for reasons, some of which we have already seen, and others of which are soon to appear.

Another approach to the matter is provided by the material in the present section. A natural thought is that our theory of truth (or anything else, for that matter) should minimise inconsistency. Contradictions should not be multiplied beyond necessity (as *IC*, 8.4 puts it). If this is the case, then the natural interpretations of the truth predicate are the  $\text{mi}$  ones (at least with respect to the truth predicate). Whether *LPm* has such interpretations is still an open question; on the Batens approach we know that there are. Of course, if there are  $\text{mi}$  interpretations, there may be more than one. In this case, we would require some further constraint on truth to select from amongst them. Alternatively, however, one might just accept that the behaviour of the truth predicate is radically under-determined: truth may behave differently in different “universes”. This view of truth is of a piece with the “multiverse” view of set theory: there is no unique universe of sets: there is a multiverse of such. We will return to this matter when we turn to set theory in the third part of this essay.

## 6.4 The Material Conditional

Using *LPm* or something like it may, then, get around *some* of the issues to do with the non-detachability of the material conditional. However, there are other issues arising from having the material conditional as one’s only conditional connective.

The first concerns fallibilism. If we have no detachable conditional avail-

---

<sup>36</sup>See Batens (2000).

able, there is then no way of expressing indefeasible conditionality. We can apply detachment to any conditional claim only using *LPm* or some other defeasible mechanism. But there is no way of ensuring a non-defeasible conditional connection.

This may well be seen as a problem. However, it is not clear to me that it is. The whole trend of the philosophy of science in the last 80 years has been towards fallibilism.<sup>37</sup> The positivists were forced to accept that verification was always a fallible process. A well corroborated theory could turn out to be false. The fallibility of our empirical observations forced the realisation that falsification is also fallible. In the end, it might be our data that is wrong. This is theory-laden and itself fallible. Thus, all empirical belief is fallible. Indeed, any area in which we theorise is fallible for exactly the same reasons: theory always outruns data, which is, in any case, soft. The matter is no different in our theorisation about logic.<sup>38</sup>

Dialetheism already reinforces the fallibilist picture, since it allows for the possibility that it may not always be rational to discard an inconsistent theory.<sup>39</sup> Having only the material conditional at our disposal just ratchets the story up a notch. All detachment is fallible too. This is not to say that it is useless. Fallible conditions (such as: if there are thunder clouds, it is going to rain) are conditions none the less. But it *is* to say that we may always have to go back and retract the conclusions reached by detachment in the light of new information—specifically, in this case, that we are in (or may well be in) an inconsistent situation. This is not something that should upset any good fallibilist, though. Indeed, they should welcome the support for their general view.

Unfortunately, there is a second, and I think, more significant problem with having only the material conditional at one's disposal. This is that the material conditional is a poor candidate for representing a natural language *if*. The problem is not that it is too *weak*, but that it is too *strong*. It validates too many inferences which are not plausible for an indicative conditional, let alone a subjunctive conditional (assuming there to be a difference, which I doubt<sup>40</sup>). And going paraconsistent does little to help.

Standard discussions focus on the old war horses:

---

<sup>37</sup>The story is nicely told in Chalmers (1999).

<sup>38</sup>See Priest (2014), (2016).

<sup>39</sup>See *IC*, p. 106-7, and Priest (2006b), Part 3.

<sup>40</sup>See Priest (201+).

- $A \vDash B \supset A$
- $\neg A \vDash A \supset B$

suggesting why these inferences are not as bad as they appear. These suggestions hardly work in their own terms.<sup>41</sup> But they can deal at best with “naked conditionals”—ones not in the scope of some other operator. The most telling objections concern non-naked conditionals. There is a whole host of problematic inferences (which are strangely absent in the standard discussions of the matter). Routley *et al* (1982), 1.2, has a list of 16 such inferences. Most of these are equally valid for the paraconsistent material conditional. I reproduce a few of these here.

- $\neg(A \supset B) \vDash A$  [It’s not the case that if there is a good god then the prayers of evil people will be answered. So there is a good god.]
- $\neg(A \supset B) \vDash \neg A \supset B$  [It’s not the case that if you break the mirror you will have bad luck. So if you don’t break the mirror you will have bad luck.]
- $(A \supset B) \wedge (C \supset D) \vDash (A \supset D) \vee (C \supset B)$  [If John is in Paris then he is in France, and if John is in Istanbul then he is in Turkey. So either, if he is in Paris then he is in Turkey, or if he is in Istanbul then he is in Paris.]
- $(A \wedge B) \supset C \vDash ((A \wedge \neg B) \supset C) \vee ((\neg A \vee B) \supset C)$  [If  $a$  (a quadrilateral) has equal angles and equal sides it is a square. So either, if  $a$  has equal angles and does not have equal sides it is a square, or if  $a$  does not have equal angles and has equal sides it is a square.]
- $C \vDash (A \supset B) \vee (B \supset A)$  [He’s an anarchist. So if he is smiling he has a bomb in his pocket, or if he has a bomb in his pocket he is smiling.]

The indicative conditional is not, then, a material conditional. Real conditionals are not truth functional. And if we do need a non-material conditional, why not suppose it to be detachable if one can indeed have such a thing?

---

<sup>41</sup>See, for example, Priest (2009), which discusses Jackson’s account of the matter—one of the most sophisticated.



## 7 A Consistency Operator

### 7.1 What This is

The next relevant issue concerns a consistency operator.<sup>42</sup> Consider a monadic operator,  $\circ$ , with the following truth/falsity conditions:

- if  $\Vdash^+ A$  and  $\Vdash^- A$  then  $\Vdash^- \circ A$  and it is not the case that  $\Vdash^+ \circ A$
- otherwise,  $\Vdash^+ \circ A$  and it is not the case that  $\Vdash^- \circ A$

Given the use of  $t$ ,  $f$ ,  $b$ , the behaviour of  $\circ$  may be helpfully depicted by the following truth table:

$A$	$\circ A$
$t$	$t$
$b$	$f$
$f$	$t$

Note that  $\circ A, A! \models B$ . So  $B \vee A!, \circ A \models B$ . Note also that  $\circ$  is not monotonic.

Though consistency operators have long been a feature of some paraconsistent traditions—especially, following da Costa, the Brazilian tradition<sup>43</sup>—they have not been popular with those who want to retain an unrestricted  $T$ -Schema or Abstraction principle. The reason for this is that the consistency operator allows us to define a “strong negation”,  $\dagger A: \neg A \wedge \circ A$ . It is easy enough to check that its truth table is as follows.

$A$	$\dagger A$
$t$	$f$
$b$	$f$
$f$	$t$

Now,  $\dagger$  behaves just like negation in classical propositional logic. In fact, the  $\vee\text{-}\wedge\text{-}\dagger$  fragment of the logic is exactly classical logic.<sup>44</sup> This generates

---

<sup>42</sup>This is not considered by Goodship. I am grateful to Hitoshi Omori for emphasising to me the importance of a consistency operator in the present context. See, e.g., Omori (2015).

<sup>43</sup>See, e.g., Carnielli, Coniglio, and Marcos (2007).

<sup>44</sup>Any classical counter-model is obviously a paraconsistent counter-model. Given an  $LP$  counter-model, define an interpretation which is the same, except that any parameter which takes the value  $b$ , takes the value  $t$ . A simple induction shows that any sentence that has the value  $t$  or  $b$  in the old model has the value  $t$  in the new one; and any sentence that has the value  $f$  in the old model has the value  $f$  in this one. Hence, this is a classical counter-model.

two problems. The first is this. If there is an operator that behaves as does classical negation (“Boolean negation”),  $\dagger$ , one might be tempted to argue that paraconsistent negation is not really negation at all. Real negation is  $\dagger$ ;  $\neg$  is just some other odd functor. In particular, then, there can be no true contradictions. Secondly, and relatedly, there is a liar sentence,  $L^\dagger$ , of the form  $\dagger T \langle L^\dagger \rangle$ . The  $T$ -Schema then gives  $T \langle L^\dagger \rangle \Leftrightarrow \dagger T \langle L^\dagger \rangle$ . If  $\Leftrightarrow$  detaches then, since  $T \langle L^\dagger \rangle \vee \dagger T \langle L^\dagger \rangle$ , we have  $T \langle L^\dagger \rangle \wedge \dagger T \langle L^\dagger \rangle$ ; and since  $T \langle L^\dagger \rangle \wedge \dagger T \langle L^\dagger \rangle \vdash B$ , we have triviality.<sup>45</sup>

Neither of these concerns is a serious worry in the present context, however. The definition of  $\dagger$  employs the consistency operator and its semantics; and  $b$ , after all, means ‘both true and false’. So the meaning of  $\dagger$  is *predicated* on the assumption that some things are dialetheias. Moreover,  $\neg$  toggles between truth and falsity, just as one should require of negation;  $\dagger$  does not.

And concerning the  $\dagger$ Liar, once the  $T$ -scheme is taken as a material biconditional, triviality does not arise. Indeed, exactly the same simple non-triviality proof given in 3.1 works. We just take the extension and anti-extension of  $T$  to contain all (codes of) formulas. The material  $T$ -Schema is then verified. Moreover, one can have restricted forms of the  $T$ -rules.  $A, A \supset T \langle A \rangle \vDash T \langle A \rangle \vee A!$ . Given the logical truth of the material  $T$ -Schema,  $A, \circ A \vDash T \langle A \rangle$ . In the reverse direction,  $T \langle A \rangle, T \langle A \rangle \supset A \vDash A \vee T \langle A \rangle!$ . So, given the logical truth of the material  $T$ -Schema,  $T \langle A \rangle, \circ T \langle A \rangle \vDash A$ . Since the  $T$ -Schema contraposes, and  $\circ B$  is logically equivalent to  $\circ \neg B$  we have, similarly:  $\neg A, \circ A \vDash \neg T \langle A \rangle$  and  $\neg T \langle A \rangle, \circ T \langle A \rangle \vDash \neg A$ .

What one cannot have, if one has a consistency operator, are the positive  $T$ -rules—and *a fortiori*, Transparency. For since  $L^\dagger \not\vdash \dagger T \langle L^\dagger \rangle$ ,  $T \langle L^\dagger \rangle \not\vdash \dagger T \langle L^\dagger \rangle$ ; and since  $T \langle L^\dagger \rangle \vee \dagger T \langle L^\dagger \rangle \vdash T \langle L^\dagger \rangle \wedge \dagger T \langle L^\dagger \rangle$ , we are back with triviality.<sup>46</sup> One can have them in a restricted form, slightly stronger than the above, however. Return to the fixed-point construction of 5.1. We start with the extension and anti-extension of the truth predicate each as the whole domain. We then ascend the ordinals, except this time, at successor stages, the extension of the truth predicate comprises all the  $\circ$ -free sentences that are true at the previous stage *plus* all the sentences containing  $\circ$ ; and the anti-extension of the truth predicate comprises all the  $\circ$ -free sentences that are

<sup>45</sup>For this reason, Priest (2006b), ch. 5, argues that Boolean Negation either does not behave as one might suppose, or else it is a tonk-like operator.

<sup>46</sup>In particular, then, the dual strategy is not open to Field (2008). In the dual case,  $\circ A$  would express the claim that  $A$  is not “gappy” or defective. But the detachability of the  $T$ -Schema rules out the possibility of deploying this construction.

false at the previous stage *plus* all the sentences containing  $\circ$ . As far as  $\circ$ -free formulas go, things are the same as before. In particular, we reach a fixed-point interpretation where these are in the extension/anti-extension of the truth predicate iff they hold in the interpretation. All sentences containing  $\circ$  are, however, in both the extension and the anti-extension of  $T$ . The material  $T$ -Schema is validated at the fixed point, as are the rules:  $A \vdash T \langle A \rangle$ , and  $\neg A \vdash \neg T \langle A \rangle$ . (If  $A$  is  $\circ$ -free, this is because of the fixed point. If  $A$  contains  $\circ$ , the conclusion always holds.) It does not validate the rules in the other direction, though:  $T \langle A \rangle \vdash A$  and  $\neg T \langle A \rangle \vdash \neg A$ . These hold if  $A$  is  $\circ$ -free, but may obviously fail otherwise. (Though of course, we still have  $\circ T \langle A \rangle, T \langle A \rangle \vdash A$  and  $\circ T \langle A \rangle, \neg T \langle A \rangle \vdash \neg A$ .)<sup>47</sup>

## 7.2 Classical Recapture

According to dialetheism, some situations are contradictory; but many, of course, are not. In standard mathematics, for example, we are normally reasoning about consistent situations. A natural thought is that, provided we are reasoning about consistent situations, it is perfectly fine to use classical logic. After all, if we take the space of  $LP$  interpretations, and delete the inconsistent ones, what remains is exactly the space of interpretations for classical logic.

The problem is how, exactly, to make sense of this thought. Provided that we stick to the vocabulary of  $LP$ , there is no set of sentences,  $\Delta$ , that one can add to a bunch of premises to ensure that only the consistent interpretations are to be considered.  $\Delta$  might itself be inconsistent.<sup>48</sup>

Going non-monotonic, as in  $LPm$  is one way around this. But if we have a consistency operator at our disposal, we can frame such a  $\Delta$ . On the propositional level, let  $\Pi$  be the set of propositional parameters that occur in  $\Sigma \cup \{A\}$ . Then we may take  $\Delta$  to be  $\{\circ p : p \in \Pi\}$ . At the first-order level, let  $\Pi$  be the set of predicates that occur in some sentence in  $\Sigma \cup \{A\}$ . We may take  $\Delta$  to be  $\{\forall x_1 \dots \forall x_n \circ P_n x_1 \dots x_n : P_n \text{ is an } n\text{-adic predicate in } \Pi\}$ . In both cases it is routine to prove that if the formulas in  $\Sigma \cup \{A\}$  are expressed in the language of classical logic:

- $\Delta \cup \Sigma \models_{LP} A$  iff  $\Sigma \models_{CL} A$

<sup>47</sup>I note also that results of Batens explained in 6.3 carry over to a language with a consistency operator.

<sup>48</sup>See *IC*, ch. 8.

Classical reasoning can always, then, be interpreted enthymematically.

Another way to see this is as follows. Let  $\circ\Sigma = \{\circ A : A \in \Sigma\}$ . Then if  $\Sigma \models_{CL} A$  then  $\Sigma \cup \circ\Sigma \models_{LP} A$ . For suppose that  $\Sigma \models_{CL} A$ , then by the Compactness Theorem, there are  $B_1, \dots, B_n \in \Sigma$  such that  $\{B_1, \dots, B_n\} \models_{CL} A$ . Let  $B$  be the conjunction of these  $B$ s. Then  $B \models_{CL} A$ , and  $\models_{CL} B \supset A$ . Since  $LP$  has the same logical truths as classical logic, it follows that  $\models_{LP} B \supset A$ . Since  $\Sigma \models_{LP} B$ ,  $\Sigma \models_{LP} B! \vee A$ . But  $\circ\Sigma \models_{LP} \circ B_1 \wedge \dots \wedge \circ B_n$ . So  $\circ\Sigma \models_{LP} \circ B$ . Hence,  $\Sigma \cup \circ\Sigma \models_{LP} A$ .

### 7.3 True Only

There is a certain interesting family of interdefinable notions. This includes being true and true only (*to*), being false and false only (*fo*), being consistent (*con*). Thus, a sentence is *fo* iff its negation is *to*, and it is *con* iff it is *fo* or *to*. And a sentence is *to* if it is *con* and true.

Perhaps the objection made most frequently to a dialethic solution to the semantic paradoxes is that a dialethist cannot express any of the notions in this family—on pain of extended paradoxes.<sup>49</sup> Thus put, the claims is patently false. A dialethist can express the claims in question *in those very words*, e.g.:  $x$  is *fo* iff  $Fx \wedge \neg Tx$ . Nor do extended paradoxes framed in terms of these notions generate triviality (as any of the standard non-triviality proofs show).

The usual reply is that those words do not express the thought that a sentence is *really to*, *fo*, or *con*, since something might be *to* and false as well, or *con* and inconsistent as well. Indeed, the usual liar sentence generates such a situation, as we noted in 4.4.

As a moment's thought will attest, though, this has now changed the objection. It is no longer that *con* and its cognates cannot be expressed; it is that they cannot be expressed *consistently*. Despite the fact that many people would *like* to be able to do this, liking is not a proof of possibility; and it is not clear that being able to do this actually buys anything significant. No one, to my knowledge, has ever made a case for this.

In particular, the ability to express consistency consistently certainly doesn't provide a way of forcing things to be consistent. *Nobody*, not even an explosive logician, can do that. It would have helped Frege not one iota to add to the axioms of his *Grundgesetze* an extra axiom to the effect that his

<sup>49</sup>For references and discussion, see *IC*, 20.4.

theory was consistent. The theory would still have been inconsistent—and now doubly so. What a classical logician can do is ensure that any inconsistency collapses into triviality. (Indeed, given that their logic is explosive, they do not have to work very hard at this!) But given a detachable conditional,  $\rightarrow$ , a paraconsistent logician can do the same, by endorsing the Schema  $A! \rightarrow \perp$ . Indeed, they have a more discriminating possibility. They can express the thought that the inconsistency of a particular sentence,  $A_0$ , will induce triviality, by endorsing the sentence  $A_0! \rightarrow \perp$ .<sup>50</sup>

Given a consistency operator, there is an even simpler way of doing this, with the Schema  $\circ A$ , or an instance,  $\circ A_0$ . Moreover, to whatever extent it is desirable, a consistency operator does allow one to express consistency and it cognates consistently. (' $x$  is *to*' is expressed by  $A \wedge \circ A$ ; and ' $x$  is *fo*' is expressed by  $\neg A \wedge \circ A$ .) Thus,  $\circ A$  expresses the consistency of  $A$  in a way that should keep even the most rabid of classical logicians happy.

## 8 Restricted Quantification

### 8.1 The Problem

The matter of a consistency operator brings us to third important issue: restricted quantification. Much reasoning employs restricted quantifiers of the form: 'Some  $P$ s are  $Q$ s', and 'All  $P$ s are  $Q$ s'. In classical logic these are, of course, represented as:

$$(1) \exists x(Px \wedge Qx)$$

$$(2) \forall x(Px \supset Qx)$$

and these renditions work pretty well. They deliver the validity of the sort of inferences that one might expect, including the connections between the restricted universal and particular quantifiers.

Once we move to a paraconsistent context, matters are somewhat different. The rendition of the restricted particular quantifier as (1) is perfectly fine. But the rendition of the restricted universal quantifier as (2) appears not to be. The reason is pretty obvious: one loses the validity of perhaps the

---

<sup>50</sup>Indeed, we may dispense with the detachable conditional if we just adopt the inferential rules  $A! \vdash \perp$  or  $A_0! \vdash \perp$ , as suggested by Beall (2013a).

most obvious and frequently used inference concerning restricted quantifiers:

$Pa$ ; all  $Ps$  are  $Qs$ . Hence  $Qa$ .

since this is a version of the disjunctive syllogism:  $Pa, \forall x(Px \supset Qx) \vdash Qa$ .

The move that is usually made in this context is to replace  $\supset$  with some detachable conditional,  $\rightarrow$ , often a relevant one.<sup>51</sup> Whilst such a move delivers many of the inferences concerning restricted quantifiers which one might expect, they do not deliver others. In particular, the intensional nature of  $\rightarrow$  is bound to invalidate some of the natural connections between the different quantifiers. Thus take, for example: Everything is  $Q$ ; so all  $Ps$  are  $Qs$ :

- $\forall xQx \vdash \forall x(Px \rightarrow Qx)$

The information delivered by the premise is just too weak to deliver the intensional connection required by the conclusion.

## 8.2 Solution by the Consistency Operator

Perhaps one might try to get around these problems by continuing to use the material conditional for the restricted universal quantifier, and taking the inferences that fail to be valid as default inferences, say, using *LPm*—though one might well have reservations about this: one might expect  $Qa$  to follow from ‘all  $Ps$  are  $Qs$ ’ and  $Pa$ , even if  $\neg Pa$  as well.

A more robust approach to the matter is provided by the use of the consistency operator. We can use this to define a “strong” material conditional,  $A \sqsupset B$ , as  $\dagger A \vee B$ . It is not hard to check that its truth table is as follows:

$\sqsupset$	$t$	$b$	$f$
$t$	$t$	$b$	$f$
$b$	$t$	$b$	$f$
$f$	$t$	$t$	$t$

and since the  $\vee$ - $\wedge$ - $\dagger$  fragment of the language is exactly classical propositional logic, so is the  $\vee$ - $\wedge$ - $\dagger$ - $\sqsupset$  fragment.<sup>52</sup> It is also not difficult to see that when

<sup>51</sup>See, e.g., Beall, Brady, Hazen, Priest, and Restall (2006).

<sup>52</sup>The whole logic is the logic *LFI1* of Carnielli, Coniglio, and de Amo (2000). They use an inconsistency operator,  $\bullet$ , instead of a consistency operator; but these things are interdefinable, since one is the negation ( $\neg$ ) of the other.

the language is extended by quantifiers, the  $\forall\text{-}\wedge\text{-}\neg\text{-}\sqsupset\text{-}\forall\text{-}\exists$  fragment is classical first-order logic.<sup>53</sup>

We now define the restricted quantifiers as follows:

All $As$ are $Bs$	$\forall x(A \sqsupset B)$
Some $As$ are $Bs$	$\exists x(A \wedge B)$
Some $As$ are not $Bs$	$\exists x(A \wedge \neg B)$
No $As$ are $Bs$	$\neg\exists x(A \wedge B)$

The only standard inferences that could break down are those concerning ‘All’. But virtually all of these are valid. For example, it is not difficult to check the validity of the following:

- $\forall x(Px \sqsupset Qx), Pa \vdash Qa$
- $\forall xPx \vdash \forall x(Qx \sqsupset Px)$
- $\forall x(Px \sqsupset Qx), \forall x(Qx \sqsupset Rx) \vdash \forall x(Px \sqsupset Rx)$
- $\neg\forall x(Px \sqsupset Qx) \vdash \exists x(Px \wedge \neg Qx)$
- $\exists x(Px \wedge \neg Qx) \vdash \neg\forall x(Px \sqsupset Qx)$

On the failure side of the ledger,  $\sqsupset$  does not contrapose:

- $P \sqsupset Q \not\vdash \neg Q \sqsupset \neg P$

(assign  $t$  to  $P$ , and  $b$  to  $Q$ ). The main inferences concerning restricted quantification which fail are those that require contraposition essentially, such as:<sup>54</sup>

- $\forall x(Px \sqsupset Qx), \neg Qa \vdash \neg Pa$
- $\forall x\neg Px \vdash \forall x(Px \sqsupset Qx)$

---

<sup>53</sup>Again, any classical counter-model is an *LP* counter-model. Given an *LP* counter-model, define an interpretation which is the same, except that any  $n$ -tuple of objects which is in both the extension and anti-extension of a predicate is just in the extension. A simple induction shows that any sentence that has the value  $t$  or  $b$  in the old model has the value  $t$  in the new one; and any sentence that has the value  $f$  in the old model has the value  $f$  in this one. Hence, this is a classical counter-model.

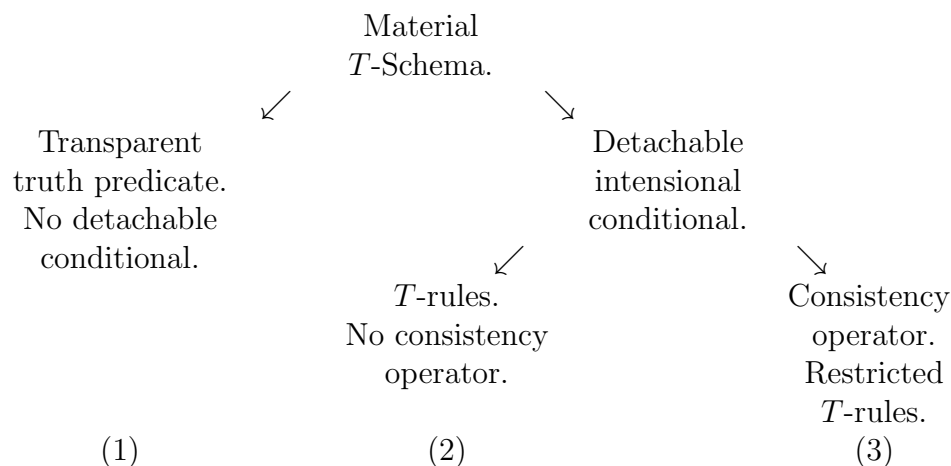
<sup>54</sup>For the second, contraposition is required to take us from  $\neg A \models \neg B \sqsupset \neg A$  to  $\neg A \models A \sqsupset B$ .

But this is entirely to be expected in a paraconsistent context, because of the possible contradictory behaviour of  $P$  and  $Q$ .

I note, finally, that since in positive contexts,  $\sqsupset$  behaves exactly as does  $\supset$ , it is no better as a candidate for a genuine conditional than  $\supset$ , being subject to many of the counter-examples of 6.4.

## 9 Interim Conclusions

The preceding discussion has taken us through many issues. So before moving on to set theory, let me try to draw together some of the threads. According to Strategy 1, the  $T$ -Schema is available in the form of a material, but not a detachable, conditional. This leaves several further options concerning truth, a detachable conditional, and the consistency operator. We may tabulate these as follows:



Let us note the relative advantages and disadvantages of each approach. We set aside Kripke's paradox (4.6), to which considerations of truth are irrelevant.

In option (1), we have a transparent truth predicate, but, on account of this, no detachable conditional. A major benefit of this approach is that it allows indefeasible endorsement.<sup>55</sup> Other plusses are given by the solutions

<sup>55</sup>This is also the only option that is available if one is a deflationist about truth—as is Beall (2008). However, I do not subscribe to such a view. See *IC*, 4.4.



to the problems noted in 4.1 (Curry Paradox) and 4.2 (Validity Curry). The downsides are that it cannot handle the problems of 4.3-4.5 (the Infinitary Curry, Smiley's Paradox, and the Denotation Paradox). It has no plausible account of a genuine conditional (if detachment be part of such an account); and it cannot accept a consistency operator with its attendant virtues.

In option (2), we have the  $T$ -rules, but not a transparent truth predicate. This has all the plusses of approach (1). In addition, it allows a decent account of an intentional and detachable conditional. Its downsides are that it does not solve the problems of 4.3-4.5, and cannot accept a consistency operator either.

In option (3), we have the solutions to all of 4.1-4.5. The option also allows for an intensional and detachable conditional, and admits the use of a consistency operator, with its attendant benefits—including a solution to the problem of restricted quantification which does not depend on defeasibility. On this account, blind endorsement has to be handled by the use of propositional quantification (or maybe defeasibly).

If one takes option (3), then the present considerations put no constraints on truth other than that it satisfy the  $T$ -Schema. This, as I have observed, is a relatively weak constraint. Further constraints may come, not from truth itself, but from the general principle of inconsistency-minimisation: inconsistencies should not be multiplied beyond necessity. The truth predicate should be as consistent as possible. One way of understanding this is that the models of truth should be minimally inconsistent (with respect to truth), in the sense of 6.3.

I would currently rank these three options in reverse order. (So (3) is preferable to (2), which is preferable to (1).) But further investigations could easily change this ranking.

## 10 The Comprehension Schema

Let us now turn to matters set-theoretic. I note that the theory of properties sits midway between the semantic and set-theoretic paradoxes,  $\lambda$ -conversion being structurally isomorphic to both the Satisfaction Schema and the Abstraction Schema. But it is not committed to the corresponding identity conditions: syntactic ones in the case of satisfaction, and Extensionality in the case of sets. Since it is less constrained than either of these notions, anything that I have to say applying to either of these applies to naive property

theory, so I will discuss it no further here.

The paradoxes of set theory are delivered by the naive Abstraction Scheme:<sup>56</sup>

- $y \varepsilon \{x : A\} \Leftrightarrow A_x(y)$

It is often technically easier to deal with a language without set abstracts, though. So let us take this in the form of the Comprehension Schema. Where  $x$  is not free in  $A$ :<sup>57</sup>

- $\exists x \forall y (y \varepsilon x \Leftrightarrow A)$

Strategy 2 is to take  $\Rightarrow$  to be a detachable, non-contracting, intensional conditional. The best work on the second strategy to date is by Zach Weber. In particular, he has shown how to obtain the standard results of the theory of numbers (ordinal and cardinal), in this approach.<sup>58</sup> On the question of the extent to which one can do other important things, such as model-theory, the jury is still out.

The strategy we are exploring here, however, is Strategy 1, where  $\Leftrightarrow$  is  $\equiv$ . This strategy does not determine how to interpret the conditionals in the other important principle of naive set theory, Extensionality:

- $\forall x (x \varepsilon y \Leftrightarrow x \varepsilon z) \Rightarrow y = z$

The obvious candidate for the biconditional is material equivalence. After all, sets are extensional entities *par excellence*. The best candidate for the (other) conditional is not so obvious. The simplest suggestion, however, is to take it to be the material conditional too. That way, the whole theory can be accommodated in the language of *LP*.

Given these understandings, many of the comments concerning the *T*-Scheme in the second part of this essay carry over to the Comprehension Schema in a very simple way. The theory is inconsistent. Just take  $A$  to be  $\neg x \varepsilon x$ . Call the resulting set  $r$ , and instantiate  $x$  with  $r$ , to get  $r \varepsilon r \equiv \neg r \varepsilon r$ . (In what follows, I will often ignore the use/mention distinction harmlessly, in the cause of readability, using a set as its own name.) But it is non-trivial.

---

<sup>56</sup>Well, the universal closure of this. In what follows,  $A$  will always be allowed to contain arbitrary parameters.

<sup>57</sup>There is also the General Comprehension Schema, where this constraint is dropped, which—perhaps surprisingly—makes very little difference to matters of triviality. See Brady (1989).

<sup>58</sup>Weber (2010), (2012).

Take an interpretation with a domain,  $D = \{a, b\}$ , where  $a$  and  $b$  are distinct. Let the extension and anti-extension of  $\varepsilon$  be  $D^2$ , and the interpretation of  $=$  be standard. It is a simple matter to show that this interpretation verifies both the Comprehension Schema and Extensionality. (The antecedent of Extensionality is both true and false.) Because  $=$  is classical, the interpretation is non-trivial. Of course, there are more interesting interpretations, as we shall see in due course.

If the language also contains a detachable conditional,  $\rightarrow$ , we can formulate the Curry Paradox by taking  $A$  to be  $x\varepsilon x \rightarrow \perp$ . Call the resulting set  $c$ , and instantiate  $x$  with  $c$  to get a sentence of the form  $c\varepsilon c \equiv (c\varepsilon c \rightarrow \perp)$ . The Curry triviality argument does not go through, due to the non-detachability of the material conditional. Indeed, non-triviality can be proved when the language contains a consistency operator,  $\circ$  and/or a detachable and contracting intensional  $\rightarrow$ . The argument in the first case is exactly the same as without it. In the second case, simply consider a (constant domain) Routley/Meyer interpretation for the logic  $R$ —so where  $\rightarrow$  is detachable and contracting—and where the base world is of the kind described in the previous paragraph.<sup>59</sup>

Adding set abstracts to the language offers some new options. First, given just the language of  $LP$ , the Abstraction Scheme,  $y\varepsilon\{x : A\} \equiv A_x(y)$  is not trivial. An easy way to see this is to take the two-element model of the Comprehension Schema defined above, extend it to one of the  $\varepsilon$ -calculus, as in 4.5, and define  $\{x : A\}$  as  $\varepsilon x \forall y (y \in x \equiv A)$ . It is easy to see that the model verifies the Abstraction Scheme, even when  $\varepsilon$ -terms are allowed to occur in  $A$ .

The analogy with truth suggests that we might need something stronger, though, namely the Abstraction Inference:

$$\frac{A_x(y)}{y\varepsilon\{x : A\}}$$

(and maybe its contraposited form too). This delivers the material Comprehension Schema. These rules can be shown to be non-trivial by a fixed-point construction.<sup>60</sup> The construction does not validate Extensionality, however.

<sup>59</sup>One may naturally ask what the mi consequences of Comprehension and Extensionality are. The answer to this is presently unknown, since we have no knowledge of what the mi models are.

<sup>60</sup>See Priest (2002a), 8.5.

I take it that the two together probably are non-trivial. But it is not clear what their models are.<sup>61</sup>

If there is a consistency operator, however, the Abstraction inference does deliver triviality. Where  $\dagger A$  is  $\neg A \wedge \circ A$ , take  $a$  to be  $\{x : \dagger x \varepsilon x\}$ . We then get  $a \varepsilon a \dashv\vdash \dagger a \varepsilon a$ , and triviality follows, as in the case for truth. Moreover, also as in the case for truth, if we have a detachable intentional arrow validating  $A \leftrightarrow A$ , and abstracts are “transparent”, in the sense that  $A_x(y)$  is everywhere intersubstitutable with  $y \varepsilon \{x : A\}$ , then we have Abstraction in a detachable form,  $y \varepsilon \{x : A\} \leftrightarrow A_x(y)$ , and Strategy 1 collapses into Strategy 2.

## 11 Regaining Set Theory

The situation concerning the theory of sets is importantly different from the theory of truth and other semantic notions, however. In the case of set theory, but not semantics, and even before the advent of Zermelo-Fraenkel set theory, there was a substantial development of informal set-theory, to which any regimentation must answer. Any formalisation which could not account for the central parts of this would not get very far. So a pressing question for Strategy 1, when it comes to set theory, is whether it can do this, and if so, how.

As an axiomatic theory, Extensionality and Comprehension would seem to be quite inadequate in this regard. The non-detachability of the material conditional would seem to render it entirely incapable of reasoning in anything like the normal way. The presence of a detachable conditional in the language does not seem to help much either: it does nothing to overcome the fact that one cannot detach in the cases of the central set-theoretic principles.

Adding set-abstracts and especially the Abstraction inference may do something to help—though given the non-detachability of Extensionality, perhaps not enough. At any rate, without an understanding of the models of this theory, it is hard to say anything very substantial about the matter.<sup>62</sup>

It might be thought that adding a classicality operator would help: given that certain sentences may be taken to be consistent, and stated to be so, a number of standard results of set-theory might be forthcoming. Thus, one could postulate that the sets in the cumulative hierarchy behave consistently.

---

<sup>61</sup>And *a fortiori*, what their mi models are.

<sup>62</sup>The same is true of using *LPm* as an inference engine: without an understanding of the mi models, nothing of much substance can be said.

The problem, though, is to prove that they exist in the first place. Another thought is to take the axioms of  $ZF$  to be consistent, so that we can deduce classically therefrom. The problem here is that, once inconsistent sets are part of the picture, there is no reason to believe that the axioms of  $ZF$  are *consistently* true.

Indeed, they are not. Zermelo's Separation Principle (*Aussonderungsaxiom*)—a restricted version of the Comprehension Schema—says that:

$$\bullet \forall z \exists y \forall x (x \varepsilon y \equiv x \varepsilon z \wedge A)$$

where  $y$  is not free in  $A$ . Now, where  $r$  is the Russell set, we have as an instance:

$$\bullet \forall z \exists y \forall x (x \varepsilon y \equiv x \varepsilon z \wedge r \varepsilon r)$$

Suppose that this has the value  $t$ . Then any instantiation of the initial universal quantifier also has the value  $t$ . Choose any  $i$  and  $j$  such that it is true that  $i \varepsilon j$ . Then the following is  $t$ :

$$\bullet \exists y \forall x (x \varepsilon y \equiv x \varepsilon j \wedge r \varepsilon r)$$

So there must be some witness of  $y$  which makes it  $t$ . Choose one such,  $k$ . Then the following is  $t$ :

$$\bullet \forall x (x \varepsilon k \equiv x \varepsilon j \wedge r \varepsilon r)$$

Again, every instantiation of the universal quantifier is  $t$ , so

$$\bullet i \varepsilon k \equiv i \varepsilon j \wedge r \varepsilon r$$

is  $t$ . Now, the right hand side of this is true, since both conjuncts are, and also false, since  $\neg r \varepsilon r$  is true. Hence, whatever value  $i \varepsilon k$  has, the biconditional has the value  $b$ , contrary to assumption.

## 11.1 Type Lifting

### 11.1.1 With a Partition

In virtue of these facts, another strategy recommends itself—a model theoretic strategy.<sup>63</sup> If there are interpretations of naive set theory that verify

---

<sup>63</sup>This is outlined in *IC*, 18.4.

all the theorems of  $ZF$  (including the Axiom of Choice—this will go without saying in what follows), we may take the universe (or universes, if there are more than one<sup>64</sup>) of sets to be like that. Since anything proved classically in  $ZF$  is true in such universes, anything which can be established in  $ZF$  would then be perfectly acceptable in dialethic set theory. That includes, of course, most of standard set-theory, though the truth would go beyond this—for example, in allowing the existence of proper-class sized totalities, which are not legitimate in  $ZF$ .

But are there such interpretations? There are. In what follows, we will look at some techniques for constructing them, starting in this section, with type-lifting. I note that our target theory will, of course, have many models, and some of them will clearly be pathological (such as the one-element interpretation where everything is true). That a theory can have pathological models is not, of course, news; but there is always the important question of which models are pathological and which are not. That is not a question I will address here, though. For the present, it suffices to show that there are models of the appropriate kind.

Let  $\mathfrak{M} = \langle D, \delta \rangle$  be any  $LP$  interpretation,<sup>65</sup> and let  $\pi$  be any partition of  $D$ . That is,  $\pi$  is a set of non-empty subsets of  $D$ , such that every member of  $D$  is in exactly one member of  $\pi$ . If  $d \in D$ , let  $[d]$  be the set of which it is a member. We define a new interpretation,  $\mathfrak{M}_\pi = \langle D_\pi, \delta_\pi \rangle$ , the *collapsed interpretation*, as follows.  $c$  is any constant in the language of  $\mathfrak{M}$ , and  $P_n$  is any  $n$ -place predicate:

- $D_\pi = \pi = \{[d] : d \in D\}$
- $\delta_\pi(c) = [\delta(c)]$
- $\langle e_1, \dots, e_n \rangle \in \delta_\pi^+(P_n)$  iff  $\exists d_1 \in e_1, \dots, d_n \in e_n \langle d_1, \dots, d_n \rangle \in \delta^+(P_n)$
- $\langle e_1, \dots, e_n \rangle \in \delta_\pi^-(P_n)$  iff  $\exists d_1 \in e_1, \dots, d_n \in e_n \langle d_1, \dots, d_n \rangle \in \delta^-(P_n)$

The collapsed model effectively identifies all the members of any set in the partition, so that the set behaves effectively like a “superposition” of its members in the collapsed model.

$\mathfrak{M}_\pi$  is an  $LP$  interpretation. Checking the relevant conditions: For any  $e_1, \dots, e_n \in D_\pi$ , there are  $d_1, \dots, d_n \in D$  such that  $d_1 \in e_1, \dots, d_n \in e_n$ .

<sup>64</sup>The idea of a plurality of universes is defended by Hamkins in (2012) and elsewhere.

<sup>65</sup>I am assuming, as in 2.1, that there are no function symbols in the language.

Since  $\langle d_1, \dots, d_n \rangle \in \delta^+(P_n)$  or  $\langle d_1, \dots, d_n \rangle \in \delta^-(P_n)$ ,  $\langle e_1, \dots, e_n \rangle \in \delta_\pi^+(P_n)$  or  $\langle e_1, \dots, e_n \rangle \in \delta_\pi^-(P_n)$ . And:

$$\begin{aligned} \langle e_1, e_2 \rangle \in \delta_\pi^+(=) & \text{ iff } \exists d_1 \in e_1, d_2 \in e_2 \langle d_1, d_2 \rangle \in \delta^+(=) \\ & \text{ iff } \exists d_1 \in e_1, d_2 \in e_2 d_1 = d_2 \\ & \text{ iff } e_1 = e_2 \end{aligned}$$

The last line holds because  $\pi$  is a partition.

We can now prove the *Collapsing Lemma*. I formulate this in a slightly more complex way than necessary because of what is to come. For simplicity of notation, I will use members of the domain as their own names. (Effectively, then, each  $d \in D$ , when construed as a name, plays the role of the canonical constant  $k_d$  of 2.1.)  $A(d_1, \dots, d_n)$  will mean that the constants in  $A$  occur amongst  $d_1, \dots, d_n$ .

For any formula in the language of  $\mathfrak{M}$ :

- If  $\mathfrak{M} \Vdash^+ A(d_1, \dots, d_n)$  and  $d_1 \in e_1, \dots, d_n \in e_n$ , then  $\mathfrak{M}_\pi \Vdash^+ A(e_1, \dots, e_n)$
- If  $\mathfrak{M} \Vdash^- A(d_1, \dots, d_n)$  and  $d_1 \in e_1, \dots, d_n \in e_n$ , then  $\mathfrak{M}_\pi \Vdash^- A(e_1, \dots, e_n)$

The proof is by a simple joint induction on the structure of  $A$ . Here are the + cases for atoms, conjunction, negation, the particular and universal quantifiers. The other cases are similar. We suppose that  $d_1 \in e_1, \dots, d_n \in e_n$ .

$$\begin{aligned} \mathfrak{M} \Vdash^+ P d_1 \dots d_n & \Rightarrow \langle d_1, \dots, d_1 \rangle \in \delta^+(P) \\ & \Rightarrow \langle e_1, \dots, e_n \rangle \in \delta_\pi^+(P) \\ & \Rightarrow \mathfrak{M}_\pi \Vdash^+ P e_1 \dots e_n \end{aligned}$$

$$\begin{aligned} \mathfrak{M} \Vdash^+ \neg A(d_1, \dots, d_n) & \Rightarrow \mathfrak{M} \Vdash^- A(d_1, \dots, d_n) \\ & \Rightarrow \mathfrak{M}_\pi \Vdash^- A(e_1, \dots, e_n) \quad IH \\ & \Rightarrow \mathfrak{M}_\pi \Vdash^+ \neg A(e_1, \dots, e_n) \end{aligned}$$

$$\begin{aligned} \mathfrak{M} \Vdash^+ (A \wedge B)(d_1, \dots, d_n) & \Rightarrow \mathfrak{M} \Vdash^+ A(d_1, \dots, d_n) \text{ and } \mathfrak{M} \Vdash^+ B(d_1, \dots, d_n) \\ & \Rightarrow \mathfrak{M}_\pi \Vdash^+ A(e_1, \dots, e_n) \text{ and } \mathfrak{M}_\pi \Vdash^+ B(e_1, \dots, e_n) \quad IH \\ & \Rightarrow \mathfrak{M}_\pi \Vdash^+ (A \wedge B)(e_1, \dots, e_n) \end{aligned}$$

$$\begin{aligned} \mathfrak{M} \Vdash^+ \exists x A(d_1, \dots, d_n) & \Rightarrow \exists d \in D, \mathfrak{M} \Vdash^+ A_x(d)(d_1, \dots, d_n) \\ & \Rightarrow \exists e \in D_\pi, \mathfrak{M}_\pi \Vdash^+ A_x(e)(e_1, \dots, e_n) \quad IH \\ & \Rightarrow \mathfrak{M}_\pi \Vdash^+ \exists x A(e_1, \dots, e_n) \end{aligned}$$

The second line holds because every  $d$  is a member of some  $e$ .

$$\begin{aligned} \mathfrak{M} \Vdash^+ \forall x A(d_1, \dots, d_n) &\Rightarrow \forall d \in D, \mathfrak{M} \Vdash^+ A_x(d)(d_1, \dots, d_n) \\ &\Rightarrow \forall e \in D_\pi, \mathfrak{M}_\pi \Vdash^+ A_x(e)(e, \dots, e_n) \quad IH \\ &\Rightarrow \mathfrak{M}_\pi \Vdash^+ \exists x A(e_1, \dots, e_n) \end{aligned}$$

The second line holds because every  $e$  contains some  $d$ .

Now, fix the language to be that of  $ZF$  (with Foundation).<sup>66</sup> Let  $\mathfrak{M} = \langle D, \delta \rangle$  be a model of  $ZF$  (a classical model if you like, but it can be any  $LP$  interpretation); let  $\alpha$  be some ordinal in  $\mathfrak{M}$ , and let  $i$  be the object which is  $V_\alpha$  in  $\mathfrak{M}$  (the sets of rank  $< \alpha$ ). Let  $\pi$  be the partition of  $D$  which puts every member of rank  $< \alpha$  (in  $\mathfrak{M}$ ) in its own singleton, and puts all the others (including  $i$ ) into a single set,  $a$ .

Construct the collapsed interpretation,  $\mathfrak{M}_\pi$ . By the Collapsing Lemma, it is a model of  $ZF$ , and since Extensionality is an axiom of  $ZF$ , it is a model of that, but it is also a model of Comprehension.

To see the last of these, take any  $e$  in  $D_\pi$ , and let  $d \in e$ . Then  $e \varepsilon a$  is true in  $\mathfrak{M}_\pi$ . For if  $d$  is of rank  $< \alpha$ , it is true in  $\mathfrak{M}$  that  $d \varepsilon i$ , so the result follows by the Collapsing Lemma. But if it is of rank  $\geq \alpha$ , then in  $\mathfrak{M}$  it is a member of, say, its singleton, which is of rank  $\geq \alpha$  and so in  $a$ . So again, the result follows. But  $e \varepsilon a$  is also false in the collapsed model. For if  $d$  is of rank  $\geq \alpha$ ,  $d \varepsilon i$  is false in  $\mathfrak{M}$ , and so the result follows by the Collapsing Lemma. But if in  $\mathfrak{M}$  it is of rank  $< \alpha$  it is not a member of, say, the singleton of  $i$ , which is in  $a$ , so the result again follows. Thus, in the collapsed model,  $e \varepsilon a!$ , and so for any  $A$ ,  $e \varepsilon a \equiv A$  is true. And since this is true for all  $e$  in  $D_\pi$ , so is  $\forall x(x \varepsilon a \equiv A)$ , and  $\exists y \forall x(x \varepsilon y \equiv A)$ .

Clearly, the collapsed model just constructed is not a very interesting one. The much overworked  $a$  is the witness set for every condition. But it suffices to establish that there are interpretations of both  $ZF$  and naive set theory. In what follows, we will see how to construct more discriminating models.

### 11.1.2 With a Covering

One way to do so is to use what I will call the Hamkins type-lift.<sup>67</sup> This is exactly the same as the previous type-lift, except that we take as  $\pi$ , not a partition, but an arbitrary covering. That is,  $\pi$  is now a family of non-empty

<sup>66</sup>What follows is taken from *IC*, 18.4.

<sup>67</sup>Since it was suggested by Joel Hamkins in discussion.



subsets of  $D$  such that every member of  $d$  is in at least one such set. (It may be in more.)

If one does this, then two features of the partition type-lift fail. The first is that in the collapsed model the denotation of ‘=’ may not be the identity relation. However, as far as constructing models of set theory goes, we can ignore this, since we do not need to assume that the language contains the identity predicate. We can just define  $x = y$  as  $\forall z(z \varepsilon x \equiv z \varepsilon y) \wedge \forall z(x \varepsilon z \equiv y \varepsilon z)$ . In  $ZF$  this delivers the substitutivity of identicals. In naive set theory, as we are understanding it here, it does not. (For example, transitivity fails.) So identity will behave in an unusual way in such a theory. But since the name of the game at this point is recapturing the theorems of  $ZF$ , this does not matter. So let us assume that we are dealing with a set theory in which identity is defined.

The second difference is that the denotation function for constants is no longer well-defined, since the sets of a partition may overlap. As far as set theory goes, this does not matter either, since the language of  $ZF$  has no constant symbols. However, as set up, constants were used in giving the truth/falsity conditions of quantifiers. This is no longer possible: we have, instead, to deploy open sentences and talk of truth under assignments of the free variables. To preserve as much of the notation as possible, let us write  $A(x_1, \dots, x_n)$  to mean that the variables of  $A$  occur amongst  $x_1, \dots, x_n$ . And if  $\mathfrak{M} = \langle D, \delta \rangle$  is any interpretation, and  $d_1, \dots, d_n$  are any members of  $D$ , we will write  $\mathfrak{M} \Vdash^\pm A(d_1, \dots, d_n)$  to mean that  $A$  is true/false in  $\mathfrak{M}$  when  $x_1, \dots, x_n$  are assigned  $d_1, \dots, d_n$ , respectively. (The practice of using members of the domain as their own names smooths this transition.)

With these understandings, the rest of the construction (including the Collapsing Lemma) goes through, as may easily be checked.

Now take any model of  $ZF$ , and let  $\pi$  be any covering of the domain. We have a model of  $ZF$  by the Collapsing Lemma. But there is a large difference between the covering construction and the partition construction. If the original model has cardinality  $\kappa$ , the partition construction produces a model of cardinality less than or equal to  $\kappa$ , whereas the covering construction produces a model of cardinality less than or equal to  $2^\kappa$ —a potentially much richer model.

Let us take  $\mathfrak{M}$  to be a classical model, and look at the richest of these, where  $\pi$  contains all the non-empty subsets of  $D$ . If  $d \in D$ , then  $\mathfrak{M}_\pi$  will contain an element  $\{d\}$ , and this will behave exactly as does  $d$  in  $\mathfrak{M}$ . But now  $\mathfrak{M}_\pi$  will contain things answering to proper classes in  $\mathfrak{M}$  too.

In particular, let  $A(x, x_1, \dots, x_n)$  be any formula in the language of  $\mathfrak{M}$ . Take  $e_1, \dots, e_n \in D_\pi$ , and let  $d_1 \in e_1, \dots, d_n \in e_n$ . Let:

- $\Gamma = \{d \in D : \mathfrak{M} \Vdash^+ A(d, d_1, \dots, d_n)\}$

Suppose that  $\Gamma = \emptyset$ . Take any  $e \in D_\pi$  and some  $d \in e$ . Then it is not the case that  $\mathfrak{M} \Vdash^+ A(d, d_1, \dots, d_n)$ , so  $\mathfrak{M} \Vdash^+ \neg A(d, d_1, \dots, d_n)$ . Hence,  $\mathfrak{M}_\pi \Vdash^+ \neg A(e, e_1, \dots, e_n)$ . Now let  $i$  be the object which is the empty set in  $\mathfrak{M}$ . Then  $\mathfrak{M} \Vdash^+ \neg d\varepsilon i$ . Let  $a$  be any member of  $D_\pi$  such that  $i \in a$ . Then  $\mathfrak{M}_\pi \Vdash^+ \neg e\varepsilon a$ . Thus, for any  $e \in D_\pi$ ,  $\mathfrak{M}_\pi \Vdash^+ e\varepsilon a \equiv A(e, e_1, \dots, e_n)$ , and so  $\mathfrak{M}_\pi \Vdash^+ \forall x(x \varepsilon a \equiv A(x, e_1, \dots, e_n))$ .

Suppose, on the other hand, that  $\Gamma \neq \emptyset$ . If  $d \in D$ , let  $d^s$  be its singleton in  $\mathfrak{M}$ , and let  $a \in D_\pi$ , where:

- $a = \{d^s : d \in \Gamma\}$

Take any  $e \in D_\pi$ , and some  $d \in e$ . Now, either  $d \in \Gamma$  or  $d \notin \Gamma$ .

If the first, then  $\mathfrak{M} \Vdash^+ A(d, d_1, \dots, d_n)$ , so  $\mathfrak{M}_\pi \Vdash^+ A(e, e_1, \dots, e_n)$ . But since  $\langle d, d^s \rangle \in \delta^+(\varepsilon)$ ,  $\langle e, a \rangle \in \delta_\pi^+(\varepsilon)$ . That is,  $\mathfrak{M}_\pi \Vdash^+ e\varepsilon a$ . Hence,  $\mathfrak{M}_\pi \Vdash^+ e\varepsilon a \equiv A(e, e_1, \dots, e_n)$ .

If the second, it is not the case that  $\mathfrak{M} \Vdash^+ A(d, d_1, \dots, d_n)$ , so  $\mathfrak{M} \Vdash^+ \neg A(d, d_1, \dots, d_n)$ , and  $\mathfrak{M}_\pi \Vdash^+ \neg A(e, e_1, \dots, e_n)$ . But if  $g \in a$ ,  $\langle d, g \rangle \in \delta^-(\varepsilon)$  (since  $g$  is a singleton of something in  $\Gamma$ —this is where the classicality of  $\mathfrak{M}$  is used). So  $\langle e, a \rangle \in \delta_\pi^-(\varepsilon)$ . That is,  $\mathfrak{M}_\pi \Vdash^+ \neg e\varepsilon a$ . Hence,  $\mathfrak{M}_\pi \Vdash^+ e\varepsilon a \equiv A(e, e_1, \dots, e_n)$ .

So in either case  $\mathfrak{M}_\pi \Vdash^+ \forall x(x\varepsilon a \equiv A(x, e_1, \dots, e_n))$ .

Thus, in all cases, there is  $a \in D_\pi$  such that  $\mathfrak{M}_\pi \Vdash^+ \forall x(x\varepsilon a \equiv A(x, e_1, \dots, e_n))$  and so  $\mathfrak{M}_\pi \Vdash^+ \exists y \forall x(x\varepsilon y \equiv A(y, e_1, \dots, e_n))$ .

Of course  $\pi$  is much larger than necessary to make this construction go through. The constraints on  $D_\pi$  needed to make the construction work are pretty minimal. But in any case, applying the construction delivers different  $a$ s to witness different instances of the Comprehension Schema, making the model much more discriminating than that of 11.1.

## 11.2 Expansions

### 11.2.1 The Basic Construction

In this section, we will look at another construction which can be used to produce models of  $ZF$  and naive set theory. We may call these ‘Expansions’.

Let  $\mathfrak{M} = \langle D, \mu \rangle$  be any model of  $ZF$  (again, not necessarily a classical model, but it can be). We now define another interpretation,  $\mathfrak{N} = \langle D, \nu \rangle$ , obtained from  $\mathfrak{M}$  by enlarging the extension/anti-extension of  $\varepsilon$ . (It leaves the extension and anti-extension of  $=$ , if it is in the language, alone.) By the Monotonicity Lemma, any such extension is also a model of  $ZF$  (and so, Extensionality).

Let  $A$  be any formula in the language of this interpretation with at most one free variable,  $x$ . The cardinality of the set of such formulas is the same as the cardinality of  $\mathfrak{M}$ . So for each such  $A$ , select a distinct object,  $a_A \in D$ . (Again, I will use members of the domain autonomously for ease of reading.)<sup>68</sup>  $\nu$  is defined as follows:

- $\nu^+(\varepsilon) = \mu^+(\varepsilon) \cup \{ \langle b, a_A \rangle : b \in D, \mathfrak{M} \Vdash^+ A_x(b) \}$
- $\nu^-(\varepsilon) = \mu^-(\varepsilon) \cup \{ \langle b, a_A \rangle : b \in D, \mathfrak{M} \Vdash^+ \neg A_x(b) \}$

For future reference, let the extension and anti-extension of  $\varepsilon$  in  $\mathfrak{N}$  be  $E^+$  and  $E^-$ , respectively.

We can now show that:

- $\mathfrak{N} \Vdash^+ \forall x(x \varepsilon a_A \equiv A)$

Let  $b$  be any member of  $D$ . Then  $\mathfrak{M} \Vdash^+ A_x(b)$  or  $\mathfrak{M} \Vdash^+ \neg A_x(b)$ . In the first case,  $\mathfrak{N} \Vdash^+ A_x(b)$ , by Monotonicity, and  $\mathfrak{N} \Vdash^+ b \varepsilon a_A$ , by construction. Hence  $\mathfrak{N} \Vdash^+ b \varepsilon a_A \equiv A_x(b)$ . In the second case,  $\mathfrak{N} \Vdash^+ \neg A_x(b)$ , by Monotonicity, and  $\mathfrak{N} \Vdash^+ \neg b \varepsilon a_A$ , by construction. Hence  $\mathfrak{N} \Vdash^+ b \varepsilon a_A \equiv A_x(b)$ ,  $\mathfrak{N} \Vdash^+ \forall x(x \varepsilon a_A \equiv A)$ , and  $\mathfrak{N} \Vdash^+ \exists y \forall x(x \varepsilon y \equiv A)$ .

In this construction, distinct sets witness different conditions in the comprehension scheme. But there is obviously an element of arbitrariness in the choice of these. This can be reduced in the following way.<sup>69</sup>

We define  $\mathfrak{N}$  as above, but we now make the choice of  $a_A$  depend on an ordinal,  $\alpha$  (in  $\mathfrak{M}$ ), generating a family of interpretations,  $\mathfrak{N}_\alpha$ . At level  $\alpha$ ,  $a_A$  is chosen as follows. If  $A$  defines a set,  $c$ , in  $\mathfrak{M}$  (so that the things that satisfy  $A$  are exactly the members of  $c$ ), take  $a_A$  to be  $c$ . Then the membership and non-membership of  $c$  will be the same in  $\mathfrak{M}$  and  $\mathfrak{N}_\alpha$ . If  $A$  does not define a

<sup>68</sup>In classical models, the construction can be varied by identifying co-extensive (in  $\mathfrak{M}$ ) sets. Specifically, define  $A_1 \sim A_2$  to mean that  $\mathfrak{M} \Vdash^+ \forall x(A_1 \equiv A_2)$ .  $\sim$  is clearly an equivalence relation, and if  $[A]$  is any equivalence class, we select a distinct object,  $a_{[A]}$ .

<sup>69</sup>The suggestion is, again, due to Joel Hamkins.

set, it is satisfied by things of arbitrarily high rank. However, it has a set-sized “approximation” at  $\alpha$ , namely, the things that satisfy it which occur at that level. Let  $a_A$  be the set in  $\mathfrak{M}$  defined by the condition  $A \wedge x \in V_\alpha$ . Then in  $\mathfrak{N}_\alpha$  the things which are members of  $a_A$  will be exactly those things satisfying  $A$ . The things which are not members of  $a_A$  will be all the members of  $V_\alpha$  that satisfy  $\neg A$ , plus all the things of rank  $> \alpha$ .

If  $A$  defines a set in  $\mathfrak{M}$ , the denotation of  $a_A$  is stable and constant as  $\alpha$  increases. But if not, as  $\alpha$  increases,  $a_A$  may change its denotation. Its extension remains stable, but its anti-extension shrinks. So  $a_A$  increasingly approximates the proper class defined by  $A$ . One might take there to be multiple universes of set theory, as given by the  $\mathfrak{N}_\alpha$ s.

Of the ways we have looked at for constructing models of  $ZF$  and naive set theory, the Hamkins type-lift produces perhaps the most natural candidate for a model of the universe of sets. The other constructions deliver a model of the cumulative hierarchy (as defined in  $ZF$ ) with inconsistent sets inside it. The Hamkins construction delivers a consistent model of the cumulative hierarchy with *extra* inconsistent sets.<sup>70</sup>

### 11.2.2 Variations

The Expansion construction can be used to show not only how dialethic set theory can capture the power of  $ZF$ , but how it can go beyond it. In this subsection, we will see how.

First, it can be used to show that there are models which contain a “standard interpretation”; that is, one satisfying the model-theoretic  $T$ -Schema (which is not, of course, possible classically). To see this, we can apply the construction to any model  $\mathfrak{M}$  (classical or non-classical) of  $ZF$  as before, but with a small addition. There are many sets,  $a$ , such that  $\mathfrak{M} \Vdash^+ [a \text{ is an } LP \text{ interpretation}]$ . (I insert square brackets here and in what follows for readability.) Choose one which is distinct from all the  $a_{AS}$ . (We can always arrange for it not to have been used before.) Maybe a natural candidate here for what follows is the section of the cumulative hierarchy,  $V_\vartheta$ , where  $\vartheta$  is some inaccessible ordinal. The relation of being true in that interpretation will be a perfectly good set (in  $\mathfrak{M}$ ),  $\Vdash_a^+$ . Its members are ordered pairs, the first of which is  $a$ , and the second of which is the code of a sentence. (I write the code of  $A$  as  $\#A$ .) Let  $p_A$  be the set in  $\mathfrak{M}$  which is the ordered pair of  $a$

---

<sup>70</sup>As in the picture suggested in *IC*, ch. 2 and 19.2.

and  $\#A$ . Then the extension and anti-extension of  $\varepsilon$  in  $\mathfrak{N}$  are:

- $E^+ \cup \{\langle p_A, \Vdash_a^+ \rangle : \mathfrak{M} \Vdash^+ A\}$
- $E^- \cup \{\langle p_A, \Vdash_a^+ \rangle : \mathfrak{M} \Vdash^+ \neg A\}$

By the Monotonicity Lemma,  $\mathfrak{N} \Vdash^+ [a \text{ is an } LP \text{ interpretation}]$ . But also:

- $\mathfrak{N} \Vdash^+ [A \equiv a \Vdash^+ A]$

For either  $\mathfrak{M} \Vdash^+ A$  or  $\mathfrak{M} \Vdash^+ \neg A$ . In the first case,  $\mathfrak{N} \Vdash^+ A$ , by Monotonicity, and  $\mathfrak{N} \Vdash^+ [a \Vdash^+ A]$ , by construction. Hence  $\mathfrak{N} \Vdash^+ [A \equiv a \Vdash^+ A]$ . In the second case,  $\mathfrak{N} \Vdash^+ \neg A$ , by Monotonicity, and  $\mathfrak{N} \Vdash^+ [\neg(a \Vdash^+ A)]$ , by construction. Hence  $\mathfrak{N} \Vdash^+ [A \equiv a \Vdash^+ A]$ . So:

- $\mathfrak{N} \Vdash^+ [a \text{ is an } LP \text{ interpretation} \wedge (A \equiv a \Vdash^+ A)]$

In other words, in  $\mathfrak{N}$ ,  $a$  is a standard interpretation.

Extending can be used to produce models of many other interesting claims. There is a lot to be said for the thought that the universe of sets is countable. In the present context, this does not rule out the possibility that it is uncountable too. Indeed, given the close conceptual connection between Cantor's theorem and Russell's paradox, the thought that this is a paradoxical situation is a very natural one.<sup>71</sup>

Now, take  $\mathfrak{M}$  to be a countable model of  $ZF$  in which  $\omega$  is the standard natural numbers. Let  $\theta$  enumerate the domain of  $\mathfrak{M}$ . Choose any  $f$  in the domain of  $\mathfrak{M}$  (distinct from all the  $a_{AS}$ ) such that:

- $\mathfrak{M} \Vdash^+ [f \text{ is a function with domain } \omega]$

Let  $t_n$  be the set in  $\mathfrak{M}$  which is the ordered pair whose first member is  $n$ , and whose second member is  $\theta(n)$ . The extension and anti-extension of  $\varepsilon$  in  $\mathfrak{N}$  are, respectively:

- $E^+ \cup \{\langle t_n, f \rangle : n \in \omega\}$
- $E^-$

Now, by Monotonicity:

- $\mathfrak{N} \Vdash^+ [f \text{ is a function with domain } \omega]$

---

<sup>71</sup>See Priest (2012).

Also, by Monotonicity,  $t_n$  is an ordered pair of  $n$  and  $\theta(n)$  in  $\mathfrak{N}$ . Hence, for any  $n$ :

- $\mathfrak{N} \Vdash^+ \langle n, \theta(n) \rangle \varepsilon f$

So:

- $\mathfrak{N} \Vdash^+ \forall y \exists x \langle x, y \rangle \varepsilon f$

and  $\mathfrak{N} \Vdash^+$  [There is a function which enumerates all sets]

Proof-theoretically, the weakness of  $LP$  is a negative feature. Model-theoretically, it is a positive feature, since it provides great scope for the construction of interesting models.

A final comment. One may ask in what set theory the above proofs are being conducted. The answer is: in  $ZF$ . One might take this to be infelicitous if the correct set theory is naive set theory. But it is not: we have just seen how naive set theory can simply appropriate any theorem of  $ZF$ .

### 11.3 Meadows' Objections

I pause here to consider some recent objections by Meadows,<sup>72</sup> who challenges the coherence of the strategy for regaining the riches of informal set theory, which we have been investigating in previous sections. He has two main arguments against it.

In the first (Section 2) he notes that the model-theoretic constructions provided deliver no reason to suppose that they are decent models of the universe of sets. This is correct. But neither were they meant to. My aim was not to give an argument for the models provided, but merely to show that there are structures of the kind in question. It is a proof of concept. All I claim is that we may suppose that the universe (or universes) of sets verifies the Comprehension Schema and the theorems of  $ZF$ .<sup>73</sup> Recall that

---

<sup>72</sup>Meadows (2015). Page and section references in what follows are to this. Meadows targets just the model of 11.1.1, but his considerations apply equally to all the models constructed in the above sections.

<sup>73</sup>In particular, I am not at all attempting to provide 'a compelling illustration of what the universe of sets out to be like' (Claim 11, p. 186). Indeed, Meadows quotes me as saying (p. 183) 'We may therefore *suppose* that the true interpretation of the language of set theory has these properties' (my italics). And as he himself says 'our goal was to be able to take up the classical perspective of set theory in a fashion that is acceptable for the paraconsistent set theorist'. *Acceptable*, not *mandated*.

the point was to show how the Naive Theory of Sets (*NST*) can make sense of standard results in set theory. It can do so if we suppose that the universe of sets is of the kind the models deliver, since they are models of *ZF*. Why should one suppose that the universe of sets is a model of *ZF*? Because doing so allows us to give an account of the Cantorian picture of sets. Unless set theory is to be unacceptably revisionist, any account must do this. Why does Meadows accept *ZF*? Precisely because it does a pretty good job of this. I do it for the same reason.

His second argument is in Section 4.<sup>74</sup> I am showing how someone who accepts *NST* can make use of the results of *ZF*. I use a construction in *ZF* to do this. This is obviously a bootstrapping construction. (Logic is full of such things. Consider, for example, how the argument for the model-theoretic soundness of a form of inference typically uses that very form of inference.) Meadows urges that it is viciously circular. I have to presuppose the adequacy of *ZF* in order to show that *ZF* is acceptable. Now, if I were trying to justify the reasoning of *ZF*, this would indeed be vicious. But as I have already explained, my aim here is not to *justify* anything: I am proposing a mathematical hypothesis which makes sense of certain facts by its own lights. In a similar way, from a classical perspective, within *ZF* one can construct the cumulative hierarchy, and one can then show that the structure verifies the axioms of *ZF*. If this were a justificatory activity, it would obviously be vicious. But *that* is not the point. It is simply a piece of boot-strapping showing a mathematical structure to be legitimate by its own lights.

In sum, then, Meadows' criticisms get off on the wrong foot right at the start. He assumes that what I am after is a justification for using *ZF*. I am not. I am assuming that we need an account which delivers Cantorian set theory. *ZF* does this, more or less, and I am explaining how *ZF* may be seen of as justified from a paraconsistent perspective.

---

<sup>74</sup>In Section 3 of his paper he considers what he takes to be an alternative strategy for achieving the goals at issue. This would be to solve Problem 5: If there is a nontrivial *LP* model of *NST*, then there is a classical model of *ZF*. I have no brief for this strategy. I expect any model of set theory that is anything like the universe of sets to be inconsistent. The remarks in this section are therefore beside the point.

## 12 Metatheory

### 12.1 Model Theoretic Validity

A discussion of set theory naturally takes us into one of metatheory, and specifically of model theory, since this is formulated in terms of sets.

Classical logicians normally take model-theoretic constructions—for both classical and non-classical logics—to be undertaken in  $ZF$ . What we have just seen is that, on the model-theoretic strategy, working within  $ZF$  produces perfectly acceptable dialethic results. In particular, working within  $ZF$ , we can define the notion of an interpretation, truth/falsity in an interpretation, and validity, for the logic  $LP$ . In particular,  $\Sigma \models A$  is defined in the usual way as:<sup>75</sup>

- $\forall \mathfrak{M}(\mathfrak{M} \models^+ \Sigma \supset \mathfrak{M} \models^+ A)$

On this account, the connection between the truth-in-an-interpretation (*tii*) of the premises and the *tii* of the conclusion is a material one. Hence, one can move from the *tii* of the premises of a valid argument to the *tii* of the conclusion only defeasibly. And given that there is a standard interpretation, where truth in an interpretation is truth *simpliciter*, the same goes for truth. Assuming that such inference is governed by something like  $LPm$ , this means that we can move from the premises to the conclusion if, but only if, the antecedent is consistent. In particular, if  $\Sigma = \{B\}$ , we cannot move from  $\mathfrak{M} \models^+ B$  to  $\mathfrak{M} \models^+ A$  if it is also not the case that  $\mathfrak{M} \models^+ B$ , as well—which is, of course, quite different from its being the case that  $\mathfrak{M} \models^+ \neg B$ .

What is the significance of this?

### 12.2 Defeasible Deductive Inference

This depends on how one understands the significance of model theory. A straightforward way to understand the model-theoretic definition of validity is as specifying the *meaning* of ‘valid’. In this case, even valid deductive inferences are, in the last instance, default inferences!

It might be thought odd to have the validity of a deductive inference grounded in a defeasible inference such as the disjunctive syllogism.<sup>76</sup> But

<sup>75</sup>This is a relationship between a set of formulas and a formula. So strictly speaking, it should be written as  $\Sigma \models \langle A \rangle$ , but I omit the quotation marks, as logicians usually do.

<sup>76</sup>The following paragraphs comes from *IC*, 18.5.



a little thought may assuage this worry. The difference between a material  $\mathfrak{M} \Vdash^+ \Sigma \supset \mathfrak{M} \Vdash^+ A$  and a detachable  $\mathfrak{M} \Vdash^+ \Sigma \rightarrow \mathfrak{M} \Vdash^+ A$  is not as great as might be thought. Both are simply true (or false) statements. Inference, by contrast, is an action. Given the premises of an argument, an inference is a jump to a new state. No number of truths is the same thing as a jump. (This is the moral of Lewis Carroll’s celebrated dialogue between Achilles and the Tortoise.<sup>77</sup>) None the less, truths of a certain kind may ground the jump, in the sense of making it a reasonable action. There is no reason why a sentence of the form  $A \supset B$  may not do this, just as much as one of the form  $A \rightarrow B$ . It is just that one of the latter kind always does so, while one of the former kind does so only sometimes (normally).

If it is not clear how a defeasible warrant for an action can work, merely consider sentences of the form:

(\*) You promised to do  $x$

The truth of (\*) is normally a ground for doing  $x$ , in the sense of making it reasonable to do so. But, to use a celebrated example, suppose that (\*) is true, where the  $x$  in question is the returning of a weapon to a certain person. And suppose that the person comes requesting the weapon, but you know that they intend to use it to commit murder. Then the truth of (\*) does not, in this context, ground the action. Just as with validity and the material conditional: the truth of a sentence of a certain kind may ground an appropriate action in normal circumstances, but fail to do so in unusual circumstances.

This approach, it seems to me, is perfectly viable, but it certainly ratchets up the considerations about fallibility of 6.4, another notch.

### 12.3 Extensional Equivalence

Another, perhaps less radical, possibility is to take the model-theoretic account as providing, not a definition of validity, but merely an extensionally equivalent characterisation. The meaning of ‘valid’ itself can be characterised in a different way, say proof-theoretically, or simply taken as an undefinable primitive. The model-theoretic account merely gives us a characterisation of what inferences are and are not deductively valid, nothing more. Valid inferences can then simply license detachment of their conclusions, though

---

<sup>77</sup>Carroll (1895).

this aspect of things may not be captured by the characterisation. In a similar way, an inferentialist who takes validity to be defined in terms of the meanings of the logical constants, spelled out in terms of introduction and elimination rules, may yet hold that a model-theoretic definition of validity delivers an extensionally equivalent characterisation (if sound and complete), but that it captures nothing of the semantic features of genuine validity.

One way to develop this idea is to use the entailment connective,  $\succrightarrow$ , of 4.2. As we noted there, the naive *V*-Schema is:  $Val(\langle A \rangle, \langle B \rangle) \equiv A \succrightarrow B$ . And now we may just think of  $Val(\langle A \rangle, \langle B \rangle)$  as  $\forall \mathfrak{M}(\mathfrak{M} \Vdash^+ A \supset \mathfrak{M} \Vdash^+ B)$ . The material biconditional spells out the extensional equivalence between the set theoretic characterisation and entailment itself, in good paraconsistent fashion. Defeasibility is still present, of course, but it now resides in both directions of the move between a model-theoretic statement of validity and entailment itself.

### 13 *In Contradiction Revisited*

Let me finish the whole discussion by reviewing the places in *IC* where detachability for the biconditionals of the naive principles plays an important role. There are four such places.

(1) In Section 1.8, there is a proof of Berry's paradox. This does not use the *LEM*. Indeed, the point of the section was exactly to show that rejecting the *LEM* does not solve all the paradoxes of self-reference. It does use detachment for the *D*-Scheme, however. So if this is not valid, the proof breaks down (though the argument still delivers a contradiction, as noted in fn. 10). In the present context, this is unimportant, however, since, as we saw in 3.1, the Liar paradox itself does not require the *LEM*.

(2) Section 3.5 gives a proof that the Gödel sentence, *G*, which states its own unprovability with respect to our naive notion of proof, is true,  $T \langle G \rangle$ . Now without the detachability of the the *T*-Schema, or the *T*-rules, we cannot infer *G*—or at best, we can infer it only by default.

In fact, there are issues even before one reaches this point.<sup>78</sup> The argument for  $T \langle G \rangle$  requires a proof of the statement of soundness, in the form  $Prov \langle A \rangle \Rightarrow T \langle A \rangle$ .<sup>79</sup> Now, if one of the rules of proof in operation is *modus*

<sup>78</sup>As noted by Field (2008), ch. 26.

<sup>79</sup>Or more generally, that  $\forall x(Prov(x) \Rightarrow Tx)$ , but only the schematic version is needed. I omit the brackets after '*Prov*' when no confusion can arise, in the cause of perspicuity.

*ponens*, the soundness argument requires us to invoke  $(T \langle A \rangle \wedge T \langle A \Rightarrow B \rangle) \Rightarrow T \langle B \rangle$ . Given a transparent truth predicate, we have  $(A \wedge (A \Rightarrow B)) \Rightarrow B$ . And if  $\Rightarrow$  detaches we are back with Curry's paradox. (See 4.2.) Strategy 1 puts this point in a new light. First, given the *T*-Schema as merely a material biconditional,  $(A \wedge (A \Rightarrow B)) \Rightarrow B$  does not follow. Secondly, even if it does, given the *T*-rules, if  $\Rightarrow$  is  $\supset$ , Curry's paradox is avoided anyway. Of course, the fact that  $\supset$  does not detach will certainly throw other spanners into the works.

All of these issues can be finessed, however.<sup>80</sup> The Gödel sentence is a sentence,  $G$ , of the form  $\neg Prov \langle G \rangle$ . Now the following are apparently analytic Schemas about naive provability:

- (1) If  $\vdash A$  then  $\vdash Prov \langle A \rangle$
- (2)  $\vdash Prov \langle A \rangle \supset A$

Given these, we may establish  $G$  very simply. An instance of (2) is  $\neg Prov \langle G \rangle \vee G$ . That is,  $\neg Prov \langle G \rangle \vee \neg \neg Prov \langle G \rangle$ , that is,  $\neg \neg Prov \langle G \rangle$ , that is,  $G$ . For good measure, since we have proved  $G$ , (1) gives us  $Prov \langle G \rangle$ , that is,  $\neg G$ .

It is also worth noting that if we have the other half of the fact that  $Prov$  represents provability:

- (3) If  $\not\vdash A$  then  $\vdash \neg Prov \langle A \rangle$

then we can prove (2). For either  $\vdash A$  or  $\not\vdash A$ . In the first case,  $\vdash Prov \langle A \rangle \supset A$ , and in the second  $\vdash \neg Prov \langle A \rangle$  by (3), and so  $\vdash Prov \langle A \rangle \supset A$ . The proof of  $G$  therefore depends on nothing more than that  $Prov$  really does represent provability.

I note, finally, that if one does have a proof of soundness in the form of the Schema  $Prov \langle A \rangle \supset A$ , then, given  $Prov \langle A \rangle$ , one can infer  $A$  at best by default. (Of course, for most systems of proof, there is no reason to suppose that provability is dialetheic; but the naive notion of proof certainly is, as we have just seen.) Moreover, if one has a statement of soundness in the form  $Prov \langle A \rangle \supset T \langle A \rangle$ ,<sup>81</sup> or more generally,  $\forall x (Prov(x) \Rightarrow Tx)$ , one needs one default inference to get to  $T \langle A \rangle$ , and—without the *T*-inference—another to get to  $A$ . Of course, using propositional quantification gives us a way around these issues too, as we have already, in effect, noted. Soundness can be

<sup>80</sup>As pointed out in *IC*, 17.5.

<sup>81</sup>Which should not be confused with  $Prov \langle A \rangle \supset \mathfrak{M} \Vdash A$ , for some interpretation,  $\mathfrak{M}$ .

expressed by  $\forall x \forall p ((Prov(x) \wedge S(x, p)) \Rightarrow p)$ , where  $S(x, p)$  can be read as ‘ $x$  says that  $p$ ’.  $A$  can then be inferred from  $Prov \langle A \rangle$  and  $S(\langle A \rangle, A)$ .

(3) Section 4.2 uses the practice of endorsement as an argument for the  $T$ -Schema. This presupposes the correctness of the  $T$ -Scheme in a detachable form.

Suppose that we continue to use the truth predicate as a way of endorsing. Then, as we noted, the  $T$ -inferences (which deliver the  $T$ -Scheme in detachable form) would serve the purpose of endorsement just as well. So we might equally see the practice as an argument for these. However, if these rules are invalid, according to Strategy 1 endorsement works only as a default inference. We can still hold that the practice presupposes these inferences. It is just that the inferences actually work only defeasibly. This does not undercut the argument. We have a practice of believing what people sincerely assert. This presupposes that people do say what they take to be so; and this is generally so—though not, of course, invariably. It remains the case that the practice presupposes this fallible assumption about people. In the same way, the practice of endorsement using the truth predicate presupposes this fallible assumption of the detachability of the  $T$ -Schema. It remains the case that the practice can misfire sometimes, as can the practice of believing what one is told.

Alternatively, if we use propositional quantification for the purpose of endorsement, the phenomenon becomes irrelevant to the defence of the  $T$ -Schema. However, the plausibility of the Schema hardly rests on this argument.

(4) In Chapter 9, it is shown how to construct a semantically closed theory of truth for a language, and prove the  $T$ -Schema recursively. The proof given is formalised with a detachable  $\rightarrow$  in the semantic principles. However, the proof actually relies on very minimal logical resources. Crucial are the substitutivity of identicals, and the substitutivity of whatever sort of equivalents the left and right side of instances of the  $T$ -Scheme (actually, the  $S$ -Scheme) have. Strategy 1 does nothing to affect the former, and the latter holds for material equivalents, as we noted in 2.3. The construction, then, still goes through.<sup>82</sup>

---

<sup>82</sup>Chapter 9 discusses the use of a theory of truth as a theory of meaning (as does 4.3). Note that, for such a use, the instance of the  $T$ -Scheme for  $A$  delivered by the canonical proof procedure can still be taken as giving the meaning of  $A$ , even though it has a material biconditional. As in standard Davidsonian meaning-theory, it is not simply the truth of a biconditional of the form  $T \langle A \rangle \equiv B$  which tells you that  $B$  gives the meaning of  $A$ . It is

## 14 And to Conclude...

Which of Strategy 1 and Strategy 2 do I think a better approach to the paradoxes of self-reference? When it comes to set theory, I definitely think that Strategy 1 is better: on present evidence, the prospects for retaining most of standard set-theoretic reasoning in any other way seem remote. The case is less clear with respect to the semantic paradoxes.<sup>83</sup> Each strategy has things it handles very naturally, and things that require careful handling. At present, I see no decisive reasons to go either way.

Is the fact that there are multiple options for formulating a dialethic picture with respect to the paradoxes of self-reference a problem? On the contrary. It provides us with great flexibility for articulating the core idea of a dialethic approach to the paradoxes of self-reference in the light of future developments.<sup>84</sup>

## 15 Appendix: Curry's Paradox

In this appendix I take up the question, flagged in 4.1, of whether Curry's paradox is of a piece with the Liar paradox. The most sustained argument to the effect that the two are of the same kind has recently been mounted by Beall.<sup>85</sup> His main concern there is to argue against a dialethic solution to the sorites paradox; but in the process, he argues that the Curry paradox, or at least some instances of it, are inclosure paradoxes, and so of a kind with the Liar. In what follows, I will discuss these arguments. Let us start with some background considerations.

---

the fact that it is obtained in a certain way.

<sup>83</sup>If it turns out that Strategy 2 is the appropriate strategy here, does this imply that the two kinds of paradoxes are of different kinds? Not at all, since they are all inclosure paradoxes. See Priest (2002b), Part 3.

<sup>84</sup>I have given talks based on drafts of parts of this paper at various places in the last few years. I am grateful to those present who made helpful comments and suggestions. Thanks, too, go to two anonymous referees of this journal.

<sup>85</sup>Beall (2014b), which is a reply to Weber *et al* (2014), which is in turn a reply to Beall (2014a). References in what follows are to Beall (2014b), unless specified otherwise.

## 15.1 The Inclosure Scheme

The Inclosure Schema concerns monadic predicates,  $\varphi$ , and  $\theta$ , and a monadic function,  $\delta$ , which appear to satisfy the following conditions, where  $\Omega = \{y : \varphi(y)\}$ :<sup>86</sup>

- *Existence*:  $\Omega$  exists, and  $\theta(\Omega)$
- *Transcendence*: if  $x \subseteq \Omega$  and  $\theta(x)$ ,  $\delta(x) \notin x$
- *Closure*: if  $x \subseteq \Omega$  and  $\theta(x)$ ,  $\delta(x) \in \Omega$

Applying  $\delta$  to  $\Omega$ , the Closure and Transcendence conditions give:  $\delta(\Omega) \in \Omega \wedge \delta(\Omega) \notin \Omega$ .

The standard set theoretic and semantic paradoxes fit the scheme. Thus, in the Burali Forti paradox,  $\varphi(y)$  is:  $y$  is an ordinal;  $\theta(x)$  is the vacuous condition,  $x = x$ ; and  $\delta(x)$  is the least ordinal greater than all those in  $x$ . Transcendence and Closure are secured by definition.

In the Liar paradox,  $\varphi(y)$  is  $Ty$  ( $y$  is true);  $\theta(x)$  is the condition that  $x$  has a name, and  $\delta(x)$  is a sentence,  $s$ , of the form  $\langle s \notin \dot{x} \rangle$ , where angle brackets are a name-forming device, and  $\dot{x}$  is a name for  $x$ . Suppose  $x \subseteq \Omega$ . Then  $s \in \dot{x}$ , so  $s$  is true, and so  $s \notin \dot{x}$ . By Excluded Middle, we have Transcendence. But this establishes that  $s$  is true, so we have Closure.

The Schema shows that all these paradoxes are of the same kind. So by the Principle of Uniform Solution (*Same Kind of Paradox, Same Kind of Solution*), they should all have the same kind of solution.<sup>87</sup>

## 15.2 What is a Paradox?

Another part of the background discussion requires us to look at what Beall says on the nature of paradox. He asks (3.5) whether a paradox is an individual sentence (such as a sentence like  $T \langle C \rangle \Rightarrow \perp$ ) or a sentential schema (such as the schematic  $T \langle C \rangle \Rightarrow A$ ). The answer is: neither. It is an argument.<sup>88</sup>

<sup>86</sup>Beall, in fact, gets the status of these conditions wrong (3.1). The conditions do not need actually to be satisfied, as he says, but only *prima facie* so. (Priest (2002), 17.2.) Recall that the origin of the schema is in Russell, who certainly did not take the conditions to be satisfied.

<sup>87</sup>All this is argued in Priest (2002), esp. Part 3, and 17.2.

<sup>88</sup>So I do not think that the sentence referred to by  $g$  is an inclosure paradox, contra his fn. 19.

It would be wrong to identify a paradox with the sentence at the end of the argument: it would be silly, for example, to take Zeno's Arrow Paradox to be that the arrow does not move—or even that it does and does not move. What is important about the paradox is the argument for its conclusion.

One might yet ask whether a paradox is an individual argument or an argument schema. If I had to answer this question I would say that it is a schema. But the answer would be very misleading. Let me explain.

Paradoxes are arguments. So to individuate a paradox, one has to individuate arguments. What are the criteria of individuation? I have no good answer to this question, and neither, as far as I know, does anyone else.

To see the difficulty, let us focus on the notion of a mathematical proof—a particularly well-defined kind of argument. Many mathematical results have a number of different proofs. Thus, take Gödel's incompleteness theorem, say in the form that Peano Arithmetic (PA) is not complete. One proof of this goes via the construction of a sentence which “says of itself that it is not provable in PA”. This is, of course, the proof that was given in Gödel's original paper. It can also be found in text books such as Boolos and Jeffrey's *Computability and Logic* and Meldelsohn's *Mathematical Logic*.

Compare this with the following (sketch of) a proof. First we prove the Halting Theorem, as did Turing: whether a certain program halts is not decidable. We then observe that the statement of halting can be expressed by a sentence of first-order arithmetic. Now, if PA were complete, there would be a decision procedure. We simply enumerate the theorems until we get to either a proof or a refutation. So PA is not complete. This is clearly a different proof—or form of proof: one can naturally say either. That is obvious, even without a precise criterion of individuation. But what should that criterion be?

Another thing is also obvious. There are *some* details of a proof that are “parametric”. They can be filled in in a number of different ways. And how one does so is irrelevant to what is going on. Thus, in the case of the first proof of Gödel's theorem, the precise details of the axiomatization of PA, or of how the gödel coding is done, are irrelevant to the proof.

So it is with arguments in general. Let me illustrate the point again with another example. Consider the form of argument which we might call ‘appeal to the Bible’. John says ‘The Bible says that one should take an eye for an eye and a tooth for a tooth, so one should take an eye for an eye and a tooth for a tooth’. Mary says ‘The Bible says that you should turn the other cheek, so you should turn the other cheek’. They are both using the same form of

argument: ‘The Bible says that  $A$ ; ergo  $A$ ’. The  $A$  here is a parameter.

### 15.3 The Liar and Curry

The third piece of background concerns the Liar and Curry themselves. The Liar involves a sentence,  $L$ , of the form  $\neg T \langle L \rangle$ . Beall takes the truth predicate to be transparent, and so to satisfy the positive  $T$ -rules:

$$\frac{A}{\overline{T \langle A \rangle}}$$

We may follow him in this regard.<sup>89</sup> We now have:

$$\frac{\frac{T \langle L \rangle}{L}}{\neg T \langle L \rangle}$$

Excluded Middle then gives us  $\neg T \langle L \rangle$ , which is  $L$ , and so the  $T$ -Schema gives us  $T \langle L \rangle$  as well.

Various features of such an argument are parametric. It matters not one iota whether the self-reference is obtained by a demonstrative, ‘*this* sentence is not true’; by a definite description, ‘the sentence written on the blackboard is not true’ (where this is the only sentence written on the blackboard); or by some form of gödel coding, ‘the sentence with gödel code  $n$  is not true’—where this sentence has code  $n$ . These details are parametric, and irrelevant.

Now, consider the Curry paradox. This forms a sentence,  $C$ , of the form:  $T \langle C \rangle \rightarrow A$ , where  $A$  is arbitrary.<sup>90</sup> (Let us call  $A$  the *arbitrary sentence*.) On the basis of this,  $A$  is then established. We first establish a conditional.

$$\frac{\frac{\overline{T \langle C \rangle}}{C}}{\overline{T \langle C \rangle} \quad \overline{T \langle C \rangle \rightarrow A}}{\overline{A}} \quad \overline{T \langle C \rangle \rightarrow A}$$

<sup>89</sup>As we noted in 4.1, if one formulates the  $T$ -schema merely as a material biconditional instead, the solutions to the Liar and Curry are the same—the failure of a material detachment, due to a dialethic sentence. One may, then, accept that the two paradoxes are of a piece. The issue becomes vexed only when this is not the case.

<sup>90</sup>I use  $\rightarrow$  for this conditional, since it is necessary to apply *modus ponens* to it.



The last step discharges the two occurrences of the assumption  $T \langle C \rangle$ , and so deploys Contraction. But the conditional delivers its own antecedent.

$$\frac{\frac{\overline{T \langle C \rangle \rightarrow A}}{C}}{\overline{T \langle C \rangle}}$$

*Modus ponens* does the rest.

This is a quite different form of argument from the Liar. The conclusion need not even be a contradiction. Moreover, and importantly,  $A$  is *completely irrelevant* to the Curry argument. It is a simple parameter (not present in the Liar paradox at all). Something—the same thing—has gone wrong with the form of argument, whether or not  $A$  is true or false. One should not be able to establish that God exists or that God does not exist or that God both exists and does not, by such irrelevant considerations.

And since the paradoxical conclusion does not have to be a contradiction, it is entirely obvious that the Curry paradox does not fit the Inclosure Schema.

One can try to force it into the Schema. We take the components of the scheme to be as follows.  $\varphi(y)$  is  $Ty$ ;  $\theta(x)$  is the condition that  $x$  has a name, and  $\delta(x)$  is a sentence,  $c$ , of the form  $\langle c \in \dot{x} \rightarrow A \rangle$ . Suppose  $x \subseteq \Omega$ . The argument for Closure tracks the above reasoning. Suppose that  $c \in \dot{x}$ . Then  $c$  is true. So,  $c \in \dot{x} \rightarrow A$ , and since  $c \in \dot{x}$ ,  $A$  follows. By  $\rightarrow$ -introduction,  $c \in \dot{x} \rightarrow A$ , that is,  $c$  is true. But here we grind to a halt. There is no argument for Transcendence.

## 15.4 Neo-Curry Paradoxes

We are now in a position to address Beall's arguments. The major one is this. One may agree that the Curry with arbitrary  $A$  does not fit the schema. But consider the paradox where the  $A$  in question is false (or maybe logically false). We might call such paradoxes *falsidical* Curries; Beall calls them *Neo-Curry* paradoxes. The argument for Transcendence now goes through. Closure has delivered the truth of  $c \in \dot{x} \rightarrow A$ . Given that  $\neg A$ , we can contrapose to infer that  $c \notin \dot{x}$ . That is, Transcendence. So, it seems, we have an inclosure paradox.

Does this show that Curry's paradox is an inclosure paradox? No. The fact that the argument has special cases that fit the schema does not show

that the general form does. The fact that some cubic equations,  $ax^3 + bx^2 + cx + d = 0$ , are quadratic (when  $a = 0$ ) obviously does not tell you that cubic equations are quadratic.

However, Beall suggest that we should simply take falsidical Curries to be *sui generis*. Not a good move. Such paradoxes are *exactly the same* as the ones with true  $A$ . The heart of the Curry paradox is establishing a conclusion which is entirely arbitrary. Whether or not it is true, is *completely irrelevant*. If one does not understand this, one has not really understood Curry paradoxes.

The point of the Inclosure Schema is an explanatory one: to explain how it is that paradox arises. It needs to locate the essence of the phenomenon, so to say. The instances of the Curry argument constitute, as it were, a natural kind. The falsidical ones do not. So much the worse for thinking that you can individuate arguments in such a simplistic fashion. The following is a valid argument (form):  $A \wedge B \vdash A$ . Suppose we restrict ourselves to those cases where  $B$  is false. Do we have a different kind of argument? Obviously not.

Another way of seeing the point is this. The Neo-Curry paradoxes effectively add a premise of the form  $\neg A$  to the Curry paradox. But the Curry argument to an arbitrary conclusion, *even with false  $A$* , needs no such premise. Thus, if the Neo-Curry paradox, can be massaged to fit the inclosure schema, it is because one has added an irrelevant premise. Then saying that it is an inclosure paradox is like saying that  $\wedge$ -elimination depends on the the Law of Excluded Middle, because  $\{A \vee \neg A, A \wedge B\} \vdash A$ .

But what of the case when  $A$  is just  $\perp$ ? Can one not define  $\neg B$  as  $B \rightarrow \perp$ ? And can we not reformulate the Curry reasoning so that its conclusions are  $T \langle C \rangle$  and  $T \langle C \rangle \rightarrow \perp$ . In that case, how can there be a difference between the Liar paradox and this instance of the Curry paradox? It is not at all clear that  $\neg B$  and  $B \rightarrow \perp$  mean the same: one concerns conditionality; the other does not. But even if they do mean the same, the thought simply confuses a paradox with its conclusion. The fact that a sentence occurs at the end of two different arguments, a Liar and a Curry, does not make the arguments the same. (As I noted, there can be quite different arguments for Gödel's Theorem.)

## 15.5 The Material Conditional

In the course of his discussion, Beall raises the question of what to say about the Curry paradox formulated with a material conditional. Let us look at this.

Consider a sentence  $M$ , of the form  $T \langle M \rangle \supset A$ , where the conditional is material. That is,  $M$  is  $\neg T \langle M \rangle \vee A$ . We may now reason as follows.

$$\frac{\frac{T \langle M \rangle}{M}}{\neg T \langle M \rangle \vee A}$$

Of course, assuming  $\neg T \langle M \rangle$  gives us  $\neg T \langle M \rangle \vee A$  as well. So by Excluded Middle, we have  $\neg T \langle M \rangle \vee A$ . But this is just  $M$ . The  $T$ -Schema and the Disjunctive Syllogism then deliver  $A$ .

To fit the argument into the Inclosure Schema, we would have to proceed as follows.  $\varphi(y)$  is  $Ty$ ;  $\theta(x)$  is the condition that  $x$  has a name, and  $\delta(x)$  is a sentence,  $m$ , of the form  $\langle m \notin \dot{x} \vee A \rangle$ . Let  $x \subseteq \Omega$ . Suppose that  $m \in \dot{x}$ . Then  $m$  is true. So,  $m \notin \dot{x} \vee A$ . But if  $m \notin \dot{x}$ , then  $m \notin \dot{x} \vee A$  as well. So by Excluded Middle, we have established that  $m$  is true. There is no argument for Transcendence, however. As with the usual Curry, if we add an extra (and irrelevant) premise, that  $\neg A$ , we can, though, obtain an argument for Transcendence with the Disjunctive Syllogism.

The situations, then, concerning the material and immaterial Curries are exactly the same. And all my comments carry over to the material Curry wholesale. In particular, it is not an inclosure paradox; and considering falsidical versions of it changes nothing.

Beall finds this incredible. Of the special case of the material Curry where  $A$  is  $\perp$ , he says (p. 843):

Consider, for example,

$$g = \langle \neg Tg \vee \perp \rangle$$

which is no longer an inclosure paradox!

An exclamation mark substitutes for an argument. We are back, however, with Beall's false dichotomy. A paradox is not a sentence: it is an argument. And the same sentence can occur at the end of a (material) Curry-like argument and at the end of a Liar-like argument. Indeed, in this case we need

have no qualms about the relationship between negation and the conditional.  $\neg Tg \vee \perp$  is just a simple logical variant of  $\neg Tg$ .<sup>91</sup>

Of course, there is a close connection between the material Curry and the Liar *qua* arguments, and a little thought shows exactly what this is. We can obtain an argument for an arbitrary conclusion from the Liar, just by extending the argument with an instance of Explosion.<sup>92</sup>

$$\frac{T \langle L \rangle \wedge \neg T \langle L \rangle}{A}$$

The material Curry uses the Disjunctive Syllogism; and as is frequently noted, Explosion and the Disjunctive Syllogism come to essentially the same thing. For  $A, \neg A \vee B$  give us  $(A \wedge \neg A) \vee B$ , and Explosion delivers  $B$ . Conversely,  $A \wedge \neg A$  delivers  $A$  and  $\neg A \vee B$ , and the Disjunctive Syllogism gives us  $B$ . So it is the Liar *plus* Explosion which is essentially the same as the material Curry. The Liar paradox is a part of this. When one senses a connection (perhaps Beall's exclamation mark?), that is what one is sensing.

So the material Curry and the Liar are different. But does not a dialetheist want to solve the material Curry by appealing to the fact that we may have a dialetheia present, and so faulting the Disjunctive Syllogism (and Explosion)? Indeed they do. But no one ever said that dialetheism can solve only one form of paradox, as I have already discussed in 4.1. Moreover, and crucially, the dialethic solution to the Liar provides the basis of the solution to the material Curry. It is the fact that the Liar paradox establishes a dialetheia which renders a subsequent or entangled use of Explosion or the Disjunctive Syllogism inappropriate.

Finally, are the standard Curry and the material Curry of the same kind? That will depend on one's view of the conditional. If one takes it that the conditional is the material conditional, they are, of course, exactly the same. If one, more wisely, does not, then they use quite different forms of argument. Consider an argument for some conclusion which uses classical *reductio*, and

<sup>91</sup>Similar considerations apply to the other paradox which Beall notes in this context, which concerns a sentence,  $r$ , of the form  $\langle \neg Tr \wedge \top \rangle$ .

<sup>92</sup>A referee commented that it is this which should be thought of as the Liar paradox—otherwise it is not clear why it really is a paradox. I find this a very strange suggestion. The Liar paradox was recognised as a paradox in Ancient Greek and Medieval logic without any reference to Explosion. Indeed, the Ancient Greeks knew nothing of that inference.

so is intuitionistically invalid. Consider, also, an intuitionistically valid argument for the same conclusion. Most would agree that these are different arguments, since they use different forms of inference. Similarly for the immaterial and material Curries. One uses contraction and *modus ponens* where the other uses  $\vee$ -introduction and the disjunctive syllogism—different forms of inferences.

## 15.6 Sorites

We have now dealt with Beall’s main argument. But he has another, *ad hominem*, argument. I (and others) have argued that the Sorites paradox fits the Inclosure Schema.<sup>93</sup> Beall does not himself subscribe to this view, but he claims that if the Sorites is an inclosure paradox, so is the Curry paradox.

He argues as follows. Just as we may use the Curry to establish both true and false *As*, we can use a sorites argument to establish true and false *As*. And ‘this undercuts the inclosure argument for the sorites at exactly the same point [as one] arrests the corresponding argument for the Curry paradox: namely contraposition’ (p. 847).

Let us look at the matter carefully. How, exactly, does the sorites argument fit the Inclosure Schema?<sup>94</sup>

Let  $P$  be a tolerant predicate. (This is a parameter.)  $\varphi(y)$  is  $Py$ , so  $\Omega = \{y : Py\}$ , and  $a_0, \dots, a_n$  is a soritical sequence that takes us from inside  $\Omega$  to outside. If  $x \subseteq \Omega$  then there must be a first member of the sequence not in it. Let this be  $\delta(x)$ . By definition,  $\delta(x) \notin x$ . So we have Transcendence. But since all the members of  $x$  are  $P$ , and  $\delta(x)$  is next to one of them,  $P\delta(x)$ , by tolerance.<sup>95</sup> The inclosure contradiction is of the form  $\delta(\Omega) \notin \Omega \wedge \delta(\Omega) \in \Omega$ : the first thing in the sequence that is not  $P$  is  $P$ .

Note that an arbitrary  $A$  is playing no role in this argument at all. In particular, *nowhere has contraposition been used*. Transcendence holds by definition. Closure holds by tolerance. We have not had to make a stipulation about an arbitrary  $A$  to ensure either. That we can use the sorites to establish some things that are true and some things that are false is completely irrelevant.

<sup>93</sup>See Priest (2010) and Weber, Ripley, Priest, Hyde and Colyvan (2014).

<sup>94</sup>See Priest (2010).

<sup>95</sup>This assumes that  $x$  is non-empty. If it is empty  $P\delta(x)$  *ex hypothesi*.

## 15.7 The Bottom Line

What the preceding considerations appear to establish is that the Liar and Curry are not of the same kind, though the Liar and the Sorites are. A liar-like sentence can, however, occur at the end of both a Liar-style argument and a Curry style argument. Moreover, a material Curry is also not of the same kind as the Liar—though it is essentially related to it in a way that the Curry is not.

There may well be more to be said about all of these issues. Indeed, one thing we have learned from recent discussions is how subtle and unobvious these matters are. Beall somewhat incautiously entitled his paper ‘The End of Inclosure’. The Schema might well reply, as Mark Twain is often reported as having said: ‘News of my death has been greatly exaggerated’.

## References

- [1] Batens, D. (2000), ‘Minimally Abnormal Models in Some Adaptive Logics’, *Synthese* 125: 5-18.
- [2] Beall, J. (2009), *Spandrels of Truth*, Oxford: Oxford University Press.
- [3] Beall, J. (2013a), ‘Shrieking Against Gluts’, *Analysis* 73: 438-445.
- [4] Beall, J. (2011), ‘Multiple-Conclusion *LP* and Default Classicality’, *Review of Symbolic Logic*, 4: 326–336.
- [5] Beall, J. (2013b), ‘Free of Detachment: Logic, Rationality, and Gluts’, *Noûs* 2: 410-23.
- [6] Beall, J. (2014a), ‘Finding Tolerance without Gluts’, *Mind* 123: 791-811.
- [7] Beall, J. (2014b), ‘The End of Inclosure’, *Mind* 123: 829-49.
- [8] Beall, J., Brady, R., Hazen, A., Priest, G., and Restall, G. (2006), ‘Restricted Quantification in Relevant Logic’, *Journal of Philosophical Logic* 35: 587-98.
- [9] Beall, J., and Murzi, J., (2013): ‘Two Flavors of Curry’s Paradox’, *The Journal of Philosophy* 110: 143-65.

- [10] Brady, R. (1989) ‘The Non-Triviality of Dialectical Set Theory’, ch. 16 of Priest, G., Routley, R., and Norman, J. (eds.) *Paraconsistent Logic: Essay on the Inconsistent*, Munich: Philosophia Verlag.
- [11] Brady, R. (2006), *Universal Logic*, Stanford: CSLI Publications.
- [12] Carnielli, W., Coniglio, M., and de Amo, S. (2000), ‘Formal Inconsistency and Evolutionary Databases’, *Logic and Logical Philosophy* 8: 115-52.
- [13] Carnielli, W., Coniglio, M., and Marcos, J. (2007), ‘Logics of Formal Inconsistency’, pp. 1-93, Vol. 14, 2nd edn. of D. Gabbay and F. Guenther (eds.), *Handbook of Philosophical Logic*, Dordrecht: Kluwer Academic Publishers.
- [14] Carroll, L. (1895), ‘What the Tortoise Said to Achilles’, *Mind* 4: 278–80.
- [15] Chalmers, A. (1999), *What is this Thing Called Science?*, 3rd edn., Indianapolis, IN: Hackett.
- [16] Crabbé, M. (2011), ‘Reassurance for the Logic of Paradox’, *Review of Symbolic Logic* 4: 479-85.
- [17] Crabbé, M. (2012), ‘Reassurance via Translation’, *Logique et Analyse* 218: 281-293.
- [18] Field, H. (2008), *Saving Truth from Paradox*, Oxford: Oxford University Press.
- [19] Goodship, L. (1996), ‘On Dialethism’, *Australasian Journal of Philosophy* 74: 153-61.
- [20] Hamkins, J. D. (2012), ‘The Set-Theoretic Multiverse’, *Review of Symbolic Logic* 5: 416–449.
- [21] Kripke, S. (1975), ‘Outline of a Theory of Truth’, *Journal of Philosophy* 72, 690-716.
- [22] Kripke, S. (2011), ‘A Puzzle about Time and Thought’, ch. 13 of *Philosophical Troubles*, Oxford: Oxford University Press.

- [23] Leisenring, A. (1969) *Mathematical Logic and Hilbert's  $\epsilon$ -Symbol*, London: Macdonald Technical & Scientific.
- [24] Meadows, T. (2015), 'Unpicking Priest's Bootstraps', *Thought* 4: 181-8.
- [25] Meyer, R. K. (1973), 'Intuitionism, Entailment, and Negation', pp. 168-98 of H. Leblanc (ed.), *Truth, Syntax and Modality*, Amsterdam: North Holland.
- [26] Omori, H. (2015), 'Remarks on Naive Set Theory Based on *LP*', *Review of Symbolic Logic* 8: 279-95.
- [27] Priest, G. (1979), 'The Logic of Paradox', *Journal of Philosophical Logic* 8: 219-81.
- [28] Priest, G. (1987), *In Contradiction*, Dordrecht: Martinus Nijhoff. 2nd edn., Oxford: Oxford University Press, 2006.
- [29] Priest, G. (1991), 'Minimally Inconsistent LP', *Studia Logica* 50: 321-31.
- [30] Priest, G. (1993), 'Can Contradictions be True, II?', *Proceedings of the Aristotelian Society, Supplementary Volume* 67: 35-54.
- [31] Priest, G. (2002a), 'Paraconsistent Logic', pp. 287-393, Vol. 6, 2nd edn. of D. Gabbay and F. Guenther (eds.), *Handbook of Philosophical Logic*, Dordrecht: Kluwer Academic Publishers.
- [32] Priest, G. (2002b), *Beyond the Limits of Thought*, 2nd edn., Oxford: Oxford University Press.
- [33] Priest, G. (2005), *Towards Non-Being*, Oxford: Oxford University Press. 2nd edn., 2016.
- [34] Priest, G. (2006a), 'The Paradoxes of Denotation', ch. 7 of T. Bolander, V. F. Hendricks and S. A. Pedersen (eds.), *Self-Reference*, Stanford, CA: CLSI Lecture Notes.
- [35] Priest, G. (2006b), *Doubt Truth to be a Liar*, Oxford: Oxford University Press.



- [36] Priest, G. (2008), *Introduction to Non-Classical Logic: From If to Is*, Cambridge: Cambridge University Press.
- [37] Priest, G. (2009a), ‘Conditionals: a Debate with Jackson’, ch. 13 of I. Ravenscroft (ed.), *Minds, Worlds and Conditionals: Themes from the Philosophy of Frank Jackson*, Oxford: Oxford University Press.
- [38] Priest, G. (2009b), ‘Neighbourhood Semantics for Intentional Operators’, *Review of Symbolic Logic* 2: 360-373.
- [39] Priest, G. (2010), ‘Inclosures, Vagueness, and Self-Reference’, *Notre Dame Journal of Formal Logic* 51: 69-84.
- [40] Priest, G. (2012), ‘A Note on the Axiom of Countability’, *Al-Mukhatabat* 1: 23-32, <http://mukhatabat.unblog.fr/>
- [41] Priest, G. (2014), ‘Revising Logic’, ch. 12 of P. Rush (ed.), *The Metaphysics of Logic*, Cambridge: Cambridge University Press.
- [42] Priest, G. (2015a), ‘Fusion and Confusion’, *Topoi* 34: 55-61.
- [43] Priest, G. (2015b), ‘Kripke’s Thought-Paradox and the 5th Antinomy’, ch. 24 of T. Achourioti, H. Galinon, J. Fernandez, and K. Fujimoto (eds.), *Unifying the Philosophy of Truth*, Berlin: Springer.
- [44] Priest, G. (2016), ‘Logical Disputes and the a Priori’, *Logique et Analyse* 59: 347-366 and *Principios: Rivista de Filosofia* 23: 29-57.
- [45] Priest, G. (201+), ‘Some New Thoughts on Conditionals’, *Topoi*, to appear.
- [46] Restall, G. (2007). ‘Curry’s Revenge: the Costs of Non-Classical Solutions to the Paradoxes of Self-Reference’, pp. 262-271 of J. Beall (ed.), *Revenge of the Liar: New Essays on the Paradox*, Oxford: Oxford University Press.
- [47] Routley, R. (1977) ‘Ultralogic as Universal’, *Relevant Logic Newsletter* 2: 50–89 and 138–75. Reprinted as an appendix to *Exploring Meinong’s Jungle and Beyond*, Canberra: Research School of Social Sciences, Australian National University, 1980.

- [48] Routley, R., Plumwood, V., Meyer, R., and Brady R. (1982), *Relevant Logics and their Rivals*, Atascadero, CA: Ridgeview.
- [49] Smiley, T. (1993), ‘Can Contradictions be True, I?’, *Proceedings of the Aristotelian Society, Supplementary Volume* 67: 17-33.
- [50] Weber, Z. (2010) ‘Transfinite Numbers in Paraconsistent Set Theory’, *Review of Symbolic Logic* 3: 71-92.
- [51] Weber, Z. (2012), ‘Transfinite Cardinals in Paraconsistent Set Theory’, *Review of Symbolic Logic* 5: 269-293.
- [52] Weber, Z., Ripley, D., Priest, G., Hyde, D., and Colyvan, M. (2014), ‘Tolerating Gluts’, *Mind* 123: 813-28.