Paraconsistent Set Theory

Graham Priest
Departments of Philosophy.
Universities of Melbourne and St Andrews

1 Introduction: For JLB

John Bell played a pivotal role in my intellectual life. I was first shown the beauty of mathematics at school, by a particularly gifted teacher. Inspired by this, and I went up to Cambridge to read the subject. The process was disillusioning. With only few exceptions, the lectures were dull and boring. It was not uncommon for a lecturer to spend an hour writing on the blackboard with his or her back to the audience. Supervisors obviously derived pleasure from solving the problems that I could not solve, but this was with all the emotional engament of cross-word puzzle solving. The sense of intellectual excitement that I had experienced at school evaporated.

By the end of my undergraduate career I had decided that philosophy was probably more interesting than mathematics, and decided that my education might best be served by studying mathematical logic; I went off to London to take an MSc in the subject. It was then my great good fortune to meet John, who was lecturing on model theory in the MSc programme. I am sure that the topic could have been made just as dull and boring as the topics I had studied in Cambridge, but John was inspirational. The sense of intellectual excitement in his lectures was palpable. The lectures were crystal clear, delivered with enormous enthusiasm, and the hand-written lecture notes that he handed out were a model of elegance. Many of us would go off with him for coffee and a cigarette after the lecture; John would hold forth on the things he loved, which of course included mathematics. I was reinspired.

At the end of the MSc, I decided to do a PhD, and asked John to be my supervisor. The area I wanted to work on was somewhat peripheral to John's work, but with magnanimity, he agreed. I was a pretty self-motivated student, and John realised that he could largely leave me to follow my own devices. I would go and see him at the LSE once a month or so. We would have a coffee, and talk for hours; John, it must be said, would do most of the talking. For some reason, I seem to remember, many of the topics started with 'm': music, Mozart, morality, masochism, Marxism, Mao Zedong... And just occasionally mathematics: he would listen to what I was up to, make helpful suggestions, and wonder, no doubt, where it was all going. He was the ideal supervisor for me: enthusiastic, patient, and—though he was only a few years older than I—with a perspective of the richness of life of which I was in awe.

Paraconsistency was only a twinkling in my eye in those days. John persuaded his research students to go to an alternative logic conference in Uldum, Denmark, in 1971—not alternative in the sense of alternative logic, but politically alternative: it was organised as a protest against the acceptance of NATO funding by the organisers of the usual UK logic conference. It was listening to a talk by Moshe Machover on the philosophy of mathematics at that conference that started me thinking about the matter. (Moshe was another of my lecturers for the London MSc, and in his own way, just as inspirational as John.) Virtually nothing about paraconsistency made its way into my thesis.[1] Nonetheless, it was clear that my interests—driven as they were by a mathematical nominalism—were unorthodox even in those days. John's mathematical interests were in fairly mainstream areas of mathematical logic, but he never tried to push me into those. He was always happy to engage with and encourage things less orthodox. I suspect it appealed to the subversive in him. At one stage of my candidature, Imre Lakatos learned that there was someone with interests in the philosophy of mathematics working in the Mathematics Department, and insisted that I immediately transfer to the Philosophy Department. John protected me from the Lakatosian imperialism.

That was all many years ago. After I finished the doctorate, I left London. But John and I have remained good friends over the years, though living on different continents. My interests have since became even more heterodox. Exactly what he thinks of paraconsistency now, I am not entirely sure. I suspect that he views it as bizarre, though always with a deep chuckle in his voice and glint in his eye. Anyway, the rest of this essay is dedicated to John. Serve him right.

What I will discuss here is paraconsistent set theory. Set theory, at least, is a topic close to John's heart... Specifically, I will discuss the shape of an acceptable paraconsistent set-theory. I will review what is currently

[1]Though I remember discussing some of the material that would become 'The Logic of Paradox' (Priest (1979)), with John and my other examiner, Michael Dummett, at my PhD viva in Oxford. Neither, I think, saw much in it.

known about the matter, and suggest some new ideas. There are, it must be confessed, as many questions as answers. At the end of the essay I will apply the discussion to another important issue for paraconsistency: that concerning its metatheory—and especially the model-theoretic definition of validity. The connection is, of course, that such a metatheory is formulated within set theory.[2]

2 Paraconsistent Set Theory: Background

The problem posed by Russell's paradox and its set-theoretic cousins may be thought of as generated by two factors. First: an unrestricted abstraction—or comprehension—principle of set existence, which allows an arbitrary condition to specify a set. Second: various principles of logic which allow certain instances of this (or their conjunction) to entail everything. Since the discovery of these paradoxes, the orthodox reaction has been to maintain the principles of logic in question, but reject the unrestricted comprehension principle. This strategy gives type theory, Zermelo-Fraenkel set theory, and so on.

There is, however, another possible strategy: maintain the comprehension principle and reject, instead, some of the principles of logic in question. There are various ways one may do this, but the one which will be concern us here is the paraconsistent way. Allow for the set theory to entail contradictions, but reject the principle *ex contradictione quodlibet*, or to give it its more colourful name, Explosion, $\{\alpha, \neg\alpha\} \vdash \beta$, and hence obtain a theory that is inconsistent but non-trivial.

How should one do this? Part of the answer is easy. A paraconsistent set-theory can naturally be thought of as a theory that endorses the two axioms (or one axiom and one axiom schema):

$$\forall x(x \in y \leftrightarrow x \in z) \rightarrow y = z \qquad \textit{Ext}$$
$$\exists x \forall y(y \in x \leftrightarrow \alpha) \qquad \textit{Abs}$$

where x does not occur free in α.[3] The rest of the answer is not easy,

[2]What follows is essentially Chapter 18 of the second edition of Priest (1987). I am grateful for the permission of Oxford University Press to republish the material.

[3]One might also want to add an appropriate version of the Axiom of Choice to these. There are, however, ways of obtaining the Axiom from unrestricted comprehension. One way is to use the machinery of Hilbert's ε -calculus. (See, e.g., Leisenring (1969), pp. 105-7.) Another, much more radical, way is to take *Abs* in an absolutely unrestricted form which allows α to contain 'x' free. This delivers the Axioms of Choice (see Routley

however. What is the appropriate underlying logic? In particular, what notion of conditional is being employed in *Ext* and *Abs*?

Paraconsistency gives us several choices in answering this question. In making the appropriate choice, there are two constraints that need to be borne in mind. First, the resulting theory should not allow us to prove too much; second, it should not allow us to prove too little.

For the first: although using a paraconsistent logic allows isolated contradictions to be accepted, we still do not want wholesale contradiction. In particular, if *everything* were provable, the theory would be quite useless. And even though contradictions do not imply everything, there may still be arguments delivering triviality. A notorious one is Curry's paradox. Suppose that the conditional of the logic satisfies both *modus ponens* and Contraction (or Absorption): $\{\alpha \rightarrow (\alpha \rightarrow \beta)\} \vdash \alpha \rightarrow \beta$. Triviality then ensues.[4]

This fact puts fairly severe constraints on an appropriate underlying logic. In fact, it rules out very many paraconsistent logics. For example, it rules out da Costa's well known C systems. It also rules out many of the best known systems of relevant logic, such as R.[5] Not everything is ruled out, though, as we shall see.

But before we turn to this, let us consider the other constraint: not too little. It is easy enough to choose an underlying logic for paraconsistent set theory that does not give triviality. Choose the null logic (in which nothing follows from anything)! This is obviously not very interesting. A minimal condition of adequacy on a paraconsistent set-theory would seem to be that we can get at least a decent part of standard, orthodox, set theory out of it. We might not require everything; we might be prepared to write off various results concerning large cardinality, or peculiar consequences of the Axiom of Choice. But if we lose too much, set-theory is voided of both its use and its interest.

It should be remembered here, that paraconsistency, unlike intuitionism, has never been a consciously revisionist philosophy. The picture has always been that classical mathematics, and the reasoning that this embodies, is perfectly acceptable as long as it does not stray into the transconsistent. It is only there that it goes awry. So the unproblematically consistent bits of orthodox set-theory, at least, ought to be delivered by a paraconsistent set-theory.

The results of this second constraint are in some tension with the results

(1977), p. 924 f. of reprint) whilst, surprisingly enough, maintaining non-triviality (see Brady (1989)).

[4]See Priest (1987), 6.2.

[5]For a survey of paraconsistent logics, see Priest (2002).

of the first. Put crudely, the matter is this. If we weaken our logic in a way that is sufficient to avoid triviality, we weaken it so much that it fails to deliver much set theory that we want to keep. We will see how this tension plays out in the following discussion.

3 The Material Strategy

As we have just seen, an underlying logic for a paraconsistent set-theory must invalidate either *modus ponens* or Contraction. Both are live options. Let us start with the rejection of *modus ponens*. There are various ways that one can arrange for *modus ponens* to fail in a paraconsistent logic, but undoubtedly the most natural is to take the conditionals (and biconditionals) in *Ext* and *Abs* to be material. That is, $\alpha \rightarrow \beta$ is simply *defined* as $\neg\alpha \vee \beta$. ($\alpha \leftrightarrow \beta$ is defined in the usual way as $(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$.) In nearly every paraconsistent logic, material detachment fails: $\{\alpha, \neg\alpha \vee \beta\} \nvdash \beta$. I will call this the *material strategy*. (The strategy does not, of course, mean that the language employed does not contain different kinds of conditional. For example, it may contain a relevant and detachable conditional as well—though it need not.)[6]

A simple and natural choice here is the logic LP.[7] A sound and complete tableau system for this is as follows.[8] Lines are of the form $\alpha, +$ or $\alpha, -$. A tableau for the inference $\{\alpha_1, ..., \alpha_n\} \vdash \beta$ starts with the lines:

$$\begin{array}{c} \alpha_1, + \\ \vdots \\ \alpha_n, + \\ \beta, - \end{array}$$

The rules are as follows:

$$\begin{array}{ccc} \alpha \wedge \beta, + & \alpha \wedge \beta, - & \neg(\alpha \wedge \beta), \pm \\ \downarrow & \swarrow \quad \searrow & \downarrow \\ \alpha, + & \alpha, - \quad \beta, - & \neg\alpha \vee \neg\beta, \pm \\ \beta, + & & \end{array}$$

[6]Adopting the material strategy in some form goes *half* way towards meeting Goodship (1996), who advocates taking the main conditional of both the Comprehension Principle and the T-schema to be material. Would treating the conditionals in the two schemas show the paradoxes of self-reference to be of different kinds? No. They still all fit the Inclosure Schema (Priest (1995), Part 3), and so have the same essential structure.

[7]Priest (1987), ch. 5.

[8]See Priest (2001), 8.3.

$$
\begin{array}{ccc}
\alpha \vee \beta, - & \alpha \vee \beta, + & \neg(\alpha \vee \beta), \pm \\
\downarrow & \swarrow \quad \searrow & \downarrow \\
\alpha, - & \alpha, + \quad \beta, + & \neg\alpha \wedge \neg\beta, \pm \\
\beta, - & &
\end{array}
$$

$$
\begin{array}{c}
\neg\neg\alpha, \pm \\
\downarrow \\
\alpha, \pm
\end{array}
$$

$$
\begin{array}{ccc}
\forall x\alpha, + & \forall x\alpha, - & \neg\forall x\alpha, \pm \\
\downarrow & \downarrow & \downarrow \\
\alpha(x/b), + & \alpha(x/a), - & \exists\neg x\alpha, \pm
\end{array}
$$

$$
\begin{array}{ccc}
\exists x\alpha, + & \exists x\alpha, - & \neg\exists x\alpha, \pm \\
\downarrow & \downarrow & \downarrow \\
\alpha(x/a), + & \alpha(x/b), - & \forall\neg x\alpha, \pm
\end{array}
$$

$$
\begin{array}{c}
. \\
\downarrow \\
b = b, +
\end{array}
$$

$$
\begin{array}{c}
b = c, + \\
\alpha(x/b), \pm \\
\downarrow \\
\alpha(x/c), \pm
\end{array}
$$

Here, b and c are any terms on the branch, a is a constant new to the branch, and '$\pm$' can be disambiguated consistently either way.[9] The closure rules for a branch are two:

$$
\begin{array}{cc}
\alpha, + & \alpha, - \\
\alpha, - & \neg\alpha, - \\
\times & \times
\end{array}
$$

(The second of these enshrines the Law of Excluded Middle.)

The paraconsistent set-theory that this logic produces has a number of interesting features. It is provably non-trivial.[10] It validates all those axioms of ZF that are instances of *Abs* (of course). It validates the Axiom of

[9] The rule for $b = b$ means that this can be introduced at any time.

[10] It might be thought that without detachment the axioms cannot be shown to be inconsistent. This is false, though. An instance of *Abs* is $\forall x(x \in r \leftrightarrow \neg x \in x)$. Whence we have $r \in r \leftrightarrow \neg r \in r$; and cashing out the conditional in terms of negation and disjunction gives $r \in r \wedge \neg r \in r$. More generally, whenever α is a classical consequence of Σ, there is a β such that $\alpha \vee (\beta \wedge \neg\beta)$ follows from Σ. (See Priest (1987), 8.6.) Hence, any classically inconsistent theory is inconsistent in this logic also.

Infinity, but not the Axiom of Foundation. It can also (unlike ZF) demonstrate the existence of a universal set.[11] What theorems of ZF—beyond the axioms—it can (or cannot) establish, is as yet a largely unanswered question. But the failure of material detachment means that most of the natural arguments fail. Whilst this does not meant that there are no unnatural arguments for the same conclusions, the prospects look rather bleak. The failure of detachment is a singular handicap. For the same reason, any other way of pursuing the material strategy does not look promising.

A more promising strategy is to look at the consequences of the axioms, not in LP, but in the non-monotonic extension LPm. (See Priest (1987), 2nd ed., ch. 16.) The results of this approach are presently unknown.

4 The Relevant Strategy

A second, and perhaps more plausible, strategy is to use a conditional in a logic which validates *modus ponens*, but not Contraction. The most plausible candidate for this is a relevant logic weaker than R, one of the *depth relevant* logics, as they are sometimes called. The following is a tableau system for such a logic.[12] Lines are now of one of two forms. One is $\alpha, +i$ or $\alpha, -i$, where i is a natural number (thought of as representing a world). Premises and conclusion take the number 0. The other is $rijk$, where i, j and k are natural numbers ('r' representing a ternary accessibility relation, as is standard in the semantics for relevant logics). The rules for LP are all present, except that a natural-number world-parameter, i, is added uniformly. Thus, for example, the rule for $\wedge+$ is:[13]

$$\begin{array}{c} \alpha \wedge \beta, +i \\ \downarrow \\ \alpha, +i \\ \beta, +i \end{array}$$

[11]For details of all this, see Restall (1992). Note that he defines '$x = y$' as '$\forall z(z \in x \leftrightarrow z \in y)$'.

[12]A semantics with respect to which it is sound can be found in Priest (1987), 2nd. ed., 19.8.

[13]The rules for identity are an exception. The rules for this are:

$$\begin{array}{c} . \\ \downarrow \\ b = b, +0 \end{array} \qquad \begin{array}{c} b = c, +i \\ \alpha(x/b), \pm j \\ \downarrow \\ \alpha(x/c), \pm j \end{array}$$

It is easiest to define the conditional, $\rightarrow$, in terms of a non-contraposing conditional, $\Rightarrow$. Thus, $\alpha \rightarrow \beta$ is $(\alpha \Rightarrow \beta) \wedge (\neg\beta \Rightarrow \neg\alpha)$. The rules for $\Rightarrow$ are as follows. When $i > 0$ (i is an impossible world):

$$\begin{array}{ccc} & \alpha \Rightarrow \beta, +i & \\ & rijk & \\ \swarrow & & \searrow \\ \alpha, -j & & \beta, +k \end{array} \qquad \begin{array}{c} \alpha \Rightarrow \beta, -i \\ \downarrow \\ rijk \\ \alpha, +j \\ \beta, -k \end{array}$$

$$\begin{array}{ccc} & \neg(\alpha \Rightarrow \beta), -i & \\ & rijk & \\ \swarrow & & \searrow \\ \alpha, -j & & \neg\beta, -k \end{array} \qquad \begin{array}{c} \neg(\alpha \Rightarrow \beta), +i \\ \downarrow \\ rijk \\ \alpha, +j \\ \neg\beta, +k \end{array}$$

In the left hand rules, j and k are any numbers on the branch. In the right hand rules, j and k are new to the branch.

When $i = 0$ (i is a possible world), the rules simplify to:

$$\begin{array}{ccc} & \alpha \Rightarrow \beta, +0 & \\ \swarrow & & \searrow \\ \alpha, -j & & \beta, +j \end{array} \qquad \begin{array}{c} \alpha \Rightarrow \beta, -0 \\ \downarrow \\ \alpha, +j \\ \beta, -j \end{array}$$

$$\begin{array}{ccc} & \neg(\alpha \Rightarrow \beta), -0 & \\ \swarrow & & \searrow \\ \alpha, -j & & \neg\beta, -j \end{array} \qquad \begin{array}{c} \neg(\alpha \Rightarrow \beta), +0 \\ \downarrow \\ \alpha, +j \\ \neg\beta, +j \end{array}$$

In the left hand rules, j is any number on the branch. In the right hand rules, j is new to the branch.

The closure rules are now:

$$\begin{array}{c} \alpha, +i \\ \alpha, -i \\ \times \end{array} \qquad \begin{array}{c} \alpha, -0 \\ \neg\alpha, -0 \\ \times \end{array}$$

(So the Law of Excluded Middle is guaranteed only at the base world.)

Naive set-theories based on relevant logics such as this are known to be inconsistent but non-trivial. Indeed, the logic may be strengthened in various ways, and this is still true—though not, of course, with Contraction.[14]

[14] See Brady (1989) and Priest (2002), section 8.

Thus, this relevant set theory satisfies the first constraint. What of the second?

To answer this question (at least to the extent that the answer is known), it is useful to divide set theory into two parts. The first comprises that basic set-theory which all branches of mathematics use as a tool. The second is the more elaborate development of this, which includes transfinite set theory, as it can be established in ZF, "higher" set theory.

The theory is able to provide for virtually all of basic set theory—Boolean operations on sets, power sets, products, functions, operations on functions, etc. (I will return to the reason for the qualification 'virtually' in a moment.) Thus, for convenience, let the language be augmented with set-abstract terms. We may define the Boolean operators, $x \cap y$, $x \cup y$ and $\bar{x}$ as $\{z : z \in x \wedge z \in y\}$, $\{z : z \in x \vee z \in y\}$ and $\{z : z \notin x\}$, respectively, and $x \subseteq y$ as $\forall z(z \in x \rightarrow z \in y)$. We can then establish the usual facts concerning these notions.[15]

How much of the more elaborate development of set-theory can be proved is not currently known. What can be said is that the *standard* proofs of a number of results break down. One thing we obviously lose is that kind of argument which appeals to vacuous satisfaction. Thus, for example, suppose that we wish to establish $\forall x(x < \xi \rightarrow A(x))$ by transfinite induction on the ordinal ξ. We can no longer argue in the basis case that since $\neg x < 0$, $x < 0 \rightarrow A(x)$; but we can make the zero case explicit, and perform the induction on $\forall x(\xi = 0 \vee (x < \xi \rightarrow A(x)))$. The first disjunct must then always be considered as a special case. Things not so easy to reconstruct are arguments employing *reductio*, such as Cantor's Theorem. Where α is an assumption made for the purpose of *reductio*, we may well be able to establish that $(\alpha \wedge \beta) \rightarrow (\gamma \wedge \neg\gamma)$, for some γ, where β is the conjunction of other facts appealed to in deducing the contradiction (such as instances of *Abs*). But contraposing and detaching will give us only $\neg\alpha \vee \neg\beta$, and we can get no further. Even given β, the failure of the disjunctive syllogism prevents us from obtaining $\neg\alpha$.[16] Much remains to be done in investigating higher set-theory in this context.

Let me now return to the qualification 'virtually'. Problems arise with the empty set. There can be no set, ϕ, such that for every a and b:

(1) $a \cap \bar{a} \subseteq \phi$

[15]Much of this is spelled out in Routley (1977), section 8.

[16]This is not the only sort of problem. Various natural arguments require the use of principles that involve nested $\rightarrow$s, such as Permutation, $\{\alpha \rightarrow (\beta \rightarrow \gamma)\} \vdash \beta \rightarrow (\alpha \rightarrow \gamma)$. The logic just described does not contain this principle. Whether it can be added whilst maintaining non-triviality is not known. There is certainly triviality in the area. See Slaney (1989).

(2) $\phi \subseteq b$

For let a be $\{x : \alpha\}$ and b be $\{x : \beta\}$. Then, (1) and (2) together give us: $(x \in \{x : \alpha\} \wedge x \in \overline{\{x : \alpha\}}) \to x \in \{x : \beta\}$. *Abs* then gives $(\alpha \wedge \neg\alpha) \to \beta$, and the theory is not paraconsistent.

If we define ϕ_1 as $a \cap \overline{a}$, then this clearly satisfies (1), but it does not satisfy (2). Alternatively, if we define ϕ_2 as $\{x : \forall y\ x \in y\}$ then it is easy enough to show that this satisfies (2), but not (1). It is provably the case that both ϕ_1 and ϕ_2 have no members. One cannot, though, show that they are identical. For $\{\neg x \in y \wedge \neg x \in z\} \nvdash x \in y \leftrightarrow x \in z$. Generally speaking, one cannot expect the global structure of the universe of sets to be a Boolean algebra, as it is classically (albeit the case that, classically, the maximum element of the algebra and some set-theoretic complements are proper classes). What one will have, instead, is a De Morgan algebra.[17]

This might, perhaps, be something that can be accepted. Boolean algebras are, after all, just special cases of De Morgan algebras. But we are not finished yet. It is not only the empty set that has multiple dopplegangers; so does the universal set. In fact, all sets do. For let α be an arbitrary truth; then $x \in y \leftrightarrow (x \in y \wedge \alpha)$ is not *relevantly* valid (from left to right). Thus, even though y and $\{x : x \in y \wedge \alpha\}$ have the same members, we will not have $y = \{x : x \in y \wedge \alpha\}$. What has gone wrong at this point is clear. *Ext* notwithstanding, the entities in question are not extensional. Nor is this an accident; the identity conditions of the entities in question are given in terms of $\to$, and this is an intensional functor, more at home in giving the identity conditions for properties than sets.

This suggests changing the biconditional in *Ext*. A natural thought is to replace it with the material biconditional, $\equiv$. Natural as this thought is, the strategy does not work. For $\{\alpha \wedge \neg\alpha\} \vDash \beta \equiv \alpha$. Now let α be any provable contradiction. Then for any z, $x \in z \equiv \alpha$. By *Ext*, it now follows that $z = \{x : x \in \alpha\}$; there is only one set. (Note that this argument does not go through in the material strategy because the material conditional does not detach to give the identity.)

There is another possibility. To see this, consider restricted quantification for a moment. It is natural to express 'all As are Bs' using a conditional, thus: $\forall x(Ax \to Bx)$. If $\to$ is a standard relevant conditional, then the inference:

1. Everything is B; hence all As are Bs

fails, since it depends on the validity of the inference $B(a) \vdash A(a) \to B(a)$. Yet inferences of this form are frequently appealed to when employing

[17] For a more systematic discussion of the issue, see Dunn (1988).

restricted quantifiers of the kind in question. If we interpret $\rightarrow$ as $\supset$, the material conditional, the inference is valid enough. But now the inference:

2. All As are Bs; a is an A; hence a is a B

fails, since it employs the Disjunctive Syllogism, $A(a), \neg A(a) \vee B(a) \vdash B(a)$. This is even worse.

A solution to this problem is to use another sort of conditional. In many formulations of relevant logics, there is a logical constant, t, which may be thought of as the conjunction of all truths.[18] (So t is true at the base world, 0, and any other world at which all the things true at the base world are true.) The appropriate tableau rules for t are:

$$\begin{array}{ccc} . & \quad & \alpha, +0 \\ \downarrow & & t, +i \\ t, +0 & & \alpha, -i \\ & & \times \end{array}$$

It is not difficult to check that these validate the following inferences:

$\vdash t$

$\alpha \vdash t \rightarrow \alpha$

We may now define an enthymematic conditional, $\rightharpoonup$, in terms of t:

$\alpha \rightharpoonup \beta$ is $(\alpha \wedge t) \rightarrow \beta$

and use this as the conditional involved in restricted universal quantification. Thus, 'All As are Bs' is to be understood as $\forall x(A(x) \rightharpoonup B(x))$. We now have:

$$\begin{array}{rcl} B(a) & \vdash & t \rightarrow B(a) \\ & \vdash & A(a) \rightharpoonup B(a) \end{array}$$

Hence $\forall x B(x) \vdash \forall x(A(x) \rightharpoonup B(x))$. And $\forall x(A(x) \rightharpoonup B(x)) \vdash (t \wedge A(a)) \rightarrow B(a)$. Hence $A(a), \forall x(A(x) \rightharpoonup B(x)) \vdash B(a)$. So both the inferences 1 and 2 are valid.[19]

Now return to set-theory. It is natural to hear 'y is a subset of z' as 'all members of y are members of z', that is, on the present account,

[18] See, e.g., Dunn and Restall (2002), p. 10. Sometimes, depending on the context, t gets interpreted as the conjunction of all logical truths.

[19] For a general discussion of restricted quantification in relevant logic, see Beall *et al.* (200+), which suggests the use of a different, but closely related, kind of enthymematic conditional.

$\forall x(x \in y \rightharpoonup x \in z)$. Let us define $y \subseteq z$ in this way. We may now take Ext to be $\forall x(x \in y \rightleftharpoons x \in z) \to y = z$, where $\rightleftharpoons$ is the biconditional corresponding to $\rightharpoonup$. This is equivalent to $(y \subseteq z \wedge z \subseteq y) \to y = z$.

Using $\rightleftharpoons$ instead of $\leftrightarrow$ in Ext overcomes many of the problems we noted. Thus, for example, there is only one set that contains everything. $\forall x\, x \in y \vdash \forall x(x \in z \rightharpoonup x \in y)$. So $\forall x\, x \in y, \forall x\, x \in z \vdash y = z$. Moreover, let α be any truth. Then we have $t \to \alpha$, so $x \in y \rightharpoonup (\alpha \wedge x \in y)$. Since $(\alpha \wedge x \in y) \rightharpoonup x \in y$, we have $y = \{x : x \in y \wedge \alpha\}$. The structure of sets is still not a Boolean algebra since the empty set is still not unique.[20] We do not have $x \notin y \vdash x \in y \rightharpoonup x \in z$. Hence, we do not have $\neg\exists x\, x \in y \vdash y \subseteq z$ or, therefore, $\neg\exists x\, x \in y, \neg\exists x\, x \in z \vdash y = z$. But the empty set is enough of an oddity that this may not matter too much. Reconstructing the reasoning of set-theory using $\rightharpoonup$ in Ext and the definition of $\subseteq$ therefore looks much more promising.

5 The Model-Theoretic Strategy

I have discussed the material strategy and the relevant strategy for naive set theory. These do not exhaust the possibilities. Let us return to the axiomatisation that employs a material conditional uniformly. Call this M. (And suppose that the language contains just the standard extensional connectives and quantifiers, as in the usual formulations of ZF—and no set abstracts.) This time, we will consider, not what is provable in M, but what the models of M are. M has many models, many of which are clearly pathological. For example, there is the model with but a single element, which both is and is not a member of itself. (This verifies the trivial theory.)

But M has many other models. We can construct some of these with the Collapsing Lemma. (See Priest (1991) and (1987), 2nd. ed., ch. 16.) Let $\mathcal{M} = \langle D, I \rangle$ be any model of ZF. Let ξ be any ordinal in $\mathcal{M}$, and a be the initial section of the cumulative hierarchy, V_ξ, in $\mathcal{M}$. (That is, the pair $\langle \xi, a \rangle$ satisfies the formula 'x is an ordinal and $y = V_x$' in $\mathcal{M}$.) Define a relation, $\sim$, on D as follows:

$$(x \text{ and } y \text{ are in } a \text{ (in } \mathcal{M}) \text{ and } x = y) \text{ or } (x \text{ and } y \text{ are not in } a \text{ (in } \mathcal{M}))$$

This is obviously an equivalence relation. (Since there are no function symbols, it is vacuously a congruence relation too.) It leaves all the members of V_ξ alone, but identifies all other members of D. Construct the collapsed

[20] Note, in particular, that $\rightharpoonup$ does not contrapose. So from the fact that $x = y$ we cannot infer that $\overline{x} = \overline{y}$.

interpretation, $\mathcal{M}^{\sim} = \langle D^{\sim}, I^{\sim}\rangle$, with respect to this equivalence relation. The Collapsing Lemma tells us that $\mathcal{M}^{\sim}$ is a model of ZF.

But something else also happens. Let me use boldfacing for names. Then '**a**' refers to a in $\mathcal{M}$, and $[a]$ in $\mathcal{M}^{\sim}$. For all $b \in D^{\sim}$, the sentence $\mathbf{b} \in \mathbf{a}$ has is both true and false in $\mathcal{M}^{\sim}$. For if (in $\mathcal{M}$) b is of rank less than ξ, $\mathbf{b} \in \mathbf{a}$ is true in $\mathcal{M}$, and so in $\mathcal{M}^{\sim}$; and if not, there is some x which is also not of rank less than ξ (e.g., $\{b\}$) such that b is in x. (I am not, here, assuming the Axiom of Foundation.) Since x has been identified with a in $\mathcal{M}^{\sim}$, $\mathbf{b} \in \mathbf{a}$ is true in $\mathcal{M}^{\sim}$. Whatever b is, there are elements which do not have rank less than ξ such that b is not a member of them (e.g., $\{c\}$, where c is distinct from b and has rank greater than ξ). Since these have been identified with a in $\mathcal{M}^{\sim}$, $\mathbf{b} \in \mathbf{a}$ is also false in $\mathcal{M}^{\sim}$. Now consider any sentence of the form $\mathbf{b} \in \mathbf{a} \equiv \alpha(x/\mathbf{b})$. The left side is both true and false. Hence the biconditional is true in $\mathcal{M}^{\sim}$ ($\{\beta \wedge \neg\beta\} \vDash \beta \equiv \alpha$). So $\forall x(x \in \mathbf{a} \equiv \alpha)$ is true, as is $\exists x \forall x(x \in y \equiv \alpha)$. So $\mathcal{M}^{\sim}$ is a model of Abs. It is a model of Ext as well, of course, since this is in ZF. Hence, $\mathcal{M}^{\sim}$ is a model of naive set theory (materially construed).[21]

In fact, we can obtain more than this. Suppose that in $\mathcal{M}$ there are inaccessible cardinals. Let ϑ_1 be the least such, and ϑ_2 be a greater one. Take ξ to be ϑ_2. Since the sets of rank less than ϑ_2, and *a fortiori* than ϑ_1, remain unaffected in the collapse, both of these are consistent substructures of $\mathcal{M}^{\sim}$ which are models of ZF. Moreover, any theorem of ZF with its quantifiers relativised to V_{ϑ_1} (so that $\exists x\alpha$ becomes $\exists x \in \mathbf{c}\,\alpha$, where '**c**' refers to V_{ϑ_1}; and similarly for $\forall$) holds consistently in $\mathcal{M}^{\sim}$. (This is not true of V_{ϑ_2}, since this set itself behaves inconsistently.[22]) That is, V_{ϑ_1} is a consistent inner model of ZF (which shows that the theory of $\mathcal{M}^{\sim}$ is highly non-trivial).

To take stock, what we have established is that there are interpretations that:

- are models of *Ext* and *Abs*
- are models of ZF
- contain the cumulative hierarchy (at least up to V_{ϑ_1}) as a consistent inner model.

[21] The fact that $\mathcal{M}^{\sim}$ is a model of *Abs* is a special case of a more general lemma, to be found in Restall (1992).

[22] In fact, V_{ϑ_2} behaves just like the set of all non-well-founded sets, given Mirimanoff's paradox. It is well-founded, but it is also a member of itself, so is not well-founded.)

We may therefore suppose that the true interpretation of the language of set theory has these properties. This is an appealing picture. The cumulative hierarchy (up to ϑ_1) is a perfectly good, consistent, set-theoretic structure; but it does not exhaust the universe of sets. There may be non-well-founded sets (such as the set of all sets) and inconsistent sets, such as the set of all sets that are not members of themselves. The universe of sets is just much richer than orthodox set theory takes it to be.

Of course, the model $\mathcal{M}^{\sim}$ that we actually constructed using the Collapsing Lemma is still pathological from this perspective. It contains only one inconsistent set, $[a]$, which has to do duty for all inconsistent and non-well-founded sets. There are undoubtedly other models (the details of whose natures require further investigation).[23] It should be remembered that, even in classical logic, set-theory—and every other theory with an infinite model, but an "intended interpretation"—has an absolute infinity of pathological models. Specifying the correct interpretation is always a further issue. The model $\mathcal{M}^{\sim}$ at least suffices to demonstrate the possibility of interpretations of naive set-theory which have the above properties.[24]

And to return, at last, to the question of what to make of the theorems of orthodox set theory, ZF, on this approach. The answer is obvious. Since the universe of sets is a model of ZF (as well as naive set theory), these hold in it. We may therefore establish things in ZF in the standard classical way, knowing that they are perfectly acceptable from a paraconsistent perspective.[25] We cannot, of course, require the theorems of ZF to be consistently true in that universe; but if, on an occasion, we do require a consistent interpretation of ZF, we know how to obtain this too. The universe of sets has a consistent substructure that is a model of ZF.

[23] Some of these can be obtained by other applications of the Collapsing Lemma. Different methods of constucting models of inconsistent set theory, some of which also model ZF, are discussed in Libert (2003).

[24] Criticising the strategy under discussion here, Weir (2004), p. 398, says: 'It will not do to say ... that the models which ... [do not have the desired properties] are "pathological" or "unintended". All the dialetheist's ZFC models are unintended in the sense that they do not capture anything like the full structure of the naive universe of sets. This compares unfavourably with the unintended models of first-order number-theory: they at least contain the "real" structure of numbers.' This is simply question-begging. The thesis is precisely that one of these models does capture the full structure of the universe of sets. (Or, if there are many equally good models, then each captures the structure of an equally good universe.) From the dialetheic perspective, it is precisely the cumulative hierarchy that is an incomplete fragment of the universe of sets. And the models in question do contain the cumulative hierarchy as a fragment (at least up to an inaccessible cardinal).

[25] In particular, the argument constructing the interpretation $\mathcal{M}^{\sim}$ above can be carried out in ZF, and so is perfectly acceptable.

6 Metatheory of Paraconsistent Logic

Let us turn, finally, to the issue of paraconsistent model-theory. If the paraconsistent strategy for set theory is to be anything more than an intellectual exercise, the underlying logic used must, in some sense, be the right one for reasoning about sets. Hence arise familiar debates about which logic is correct, and why. A frequent objection made against paraconsistency in this debate goes as follows. Paraconsistent logics have metatheories. In particular, they have appropriate semantics, proof systems, and corresponding soundness and (hopefully) completeness results. Now the logic in which such proofs are carried out must be classical, non-paraconsistent, logic.[26] This shows that paraconsistent logic cannot be maintained as the correct logic.

The argument is far too swift. For a start, the logic of the metatheory of a theory need not be classical. For example, an intuitionist metatheory for intuitionist logic is well known.[27] Is there a metatheory for paraconsistent logics that is acceptable on paraconsistent terms? The answer to this question is not at all obvious. First, the standard proofs in the metatheories of paraconsistent logics are usually given, as are most mathematical proofs, in an informal way. The question, then, is how to interpret the proofs formally. A normal assumption is that the proofs are carried out using classical logic. And indeed, this would seem to be sufficient for the purpose. This point is not definitive, however. Most paraconsistent logics are generalisations of classical logic in one way or another. In particular, they coincide with classical logic in those cases (models) which are consistent (i.e., in which all formulas behave consistently). Hence, if an informal argument concerns a consistent situation, and can be regimented using classical logic, it is perfectly acceptable for a paraconsistent logician.[28] Can a paraconsistent logician, or at least, one who subscribes to paraconsistent set theory, look at the metatheoretic arguments concerning paraconsistent logic in this way? The answer, unfortunately, is 'no'. For metatheoretic constructions are carried out in set theory; and paraconsistent set theory is not consistent.

In the model-theory of paraconsistent logic, we must therefore use paraconsistent set-theory, however that is best construed. To what extent model-theory can be developed on the relevant strategy for naive set-theory is still an open question. But the model-theoretic strategy for naive set-theory provides a simple way of accommodating paraconsistent model-

[26]Rescher (1969), p. 229, documents this claim, though he does not endorse it.

[27]See, e.g., Dummett (1977), ch. 5, esp. p. 197.

[28]For further discussion, see Priest (1987), ch. 8.

theory. One may think of the metatheory of the logic, including the appropriate soundness and completeness proofs, as being carried out (as we know it can be) in ZF. According to the model-theoretic strategy, the results established in this way can perfectly well be taken to hold of the universe of sets, paraconsistently construed. The paraconsistent logician can, therefore, simply appropriate the results.

It might be thought that this approach to the metatheory of paraconsistent logic suffers from a problem. In the material and model-theoretic strategies for paraconsistent set-theory, the relationship between the premises and the conclusion of a valid inference is expressed by a material conditional. Thus, simplifying to the one-premise case for perspicuity, and writing the relation 'α holds in $\mathcal{I}$' as '$\mathcal{I} \Vdash \alpha$ ', an inference from α to β is valid iff:

Val for all every interpretation, $\mathcal{I}$ $(\mathcal{I} \Vdash \alpha \supset \mathcal{I} \Vdash \beta)$

Now, the material conditional does not support detachment. Hence an inference can be valid, yet this does not licence the detachment of the conclusion from the premise. Surely this deprives the notion of validity of its punch?

No. The disjunctive syllogism is perfectly acceptable provided that the situation is consistent.[29] Provided that we do not have $\mathcal{I} \Vdash \alpha$ and $\mathcal{I} \nVdash \alpha$, we can get from $\mathcal{I} \Vdash \alpha$ to $\mathcal{I} \Vdash \beta$. In particular, then, provided that $\mathcal{I}$ is in part of the universe of sets that is consistent (the cumulative hierarchy, or a sufficiently generous part thereof), we have business as usual. (Note: this does not mean that the set of things made true by $\mathcal{I}$ is consistent. '$\mathcal{I} \Vdash \alpha$ and $\mathcal{I} \nVdash \alpha$' is quite different from '$\mathcal{I} \Vdash \alpha$ and $\mathcal{I} \Vdash \neg\alpha$'.) If $\mathcal{I}$ is a set outside this part of the universe, matters are different. Thus, we may expect that there is an interpretation, $\mathcal{M}$, that is in accord with the actual, in the sense that for any γ, γ iff $\mathcal{M} \Vdash \gamma$. One should not expect this interpretation to be in the hierarchy. Appropriate techniques of diagonalisation will give us sentences, α, such that $\mathcal{M} \Vdash \alpha$ and $\mathcal{M} \nVdash \alpha$. In such cases, even though **Val** holds, the fact that α (i.e., $\mathcal{M} \Vdash \alpha$) will not allow us to detach β (i.e., $\mathcal{M} \Vdash \beta$). However, such αs will be unusual. In standard cases **Val** will provide a licence to get from α to β.

It might still be thought odd to have the validity of a deductive inference grounded in a defeasible inference such as the disjunctive syllogism. But a little thought should assuage this worry. The difference between a material $\mathcal{I} \Vdash \alpha \supset \mathcal{I} \Vdash \beta$ and a relevant $\mathcal{I} \Vdash \alpha \rightarrow \mathcal{I} \Vdash \beta$ is not as great as might be thought in this context. Both are simply true (or false) *statements.*

[29]See, again, Priest (1987), ch. 8.

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.[30]) 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 $\gamma \supset \delta$ may not do this just as much as one of the form $\gamma \rightarrow \delta$. It is just that one of the latter kind always does, whilst one of the former kind does so only sometimes.

If it is still not clear how a sentence can function in this way, consider sentences of the form:

(*) You promised to do x

The truth of (*) normally grounds doing x, in the sense of making it reasonable to do it. 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 that person comes requesting the weapon, but you know that they intend to use it to commit suicide. 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 objection dealt with, there would seem nothing to prevent the paraconsistent logician from simply appropriating all the classical metatheoretic results in the way explained. The appropriation might be thought to have all the charms of theft over honest toil (as Russell said in another context); on the other hand, why reinvent the wheel?

7 Conclusion: The Shock of the New

At various times in its history, mathematics has been shocked by the discovery of new kind of entity: irrational numbers, infinitesimals, transfinite sets, and so on. The reception by the mathematical community of these entities has often been controversial and contentious; and the discovery has always been followed by a process of rethinking mathematical reasoning in the light of these entities and their properties. The discovery of inconsistent objects, such as the Russell set—of all those sets that do not contain themselves—is the most recent, and perhaps the most contentious, episode

[30]Carroll (1895).

of this kind; and we are still in the process of thinking through its ramification for mathematical reasoning. In mathematical revolutions of this kind, it is always important to preserve the central parts of previous mathematical thought. What I have been engaged in here is a contribution to this project.

References

[1] Beall, JC, Brady, R., Hazen, A., Priest, G., and Restall, G. (200+), 'Restricted Quantification in Relevant Logics', *Journal of Philosophical Logic*, forthcoming.

[2] Brady, R. (1989), 'The Non-Triviality of Dialectical Set-Theory, ch. 14 of Priest, Routley, and Norman (1989).

[3] Carroll, L. (1895), 'What the Tortoise Said to Achilles', *Mind* 4, 278-280.

[4] Dummett, M. (1977), *Elements of Intuitonism*, Oxford: Oxford University Press.

[5] Dunn, J. M. (1988), 'The Impossibility of Certain Second-Order Non-Classical Logics with Extensionality', pp. 261-79 of D. F. Austin (ed.), *Philosophical Analysis*, Dordrecht: Kluwer Academic Publishers.

[6] Dunn, J. M., and Restall, G. (2002), 'Relevance Logic', pp. 1-128, of Gabbay and Guenthner (2002).

[7] Gabby, D., and Guenthner, F. (eds.) (2002), *Handbook of Philosophical Logic*, second edition, Vol. 6, Dordrecht: Kluwer Academic Publishers.

[8] Goodship, L. (1996), 'On Dialethism', *Australasian Journal of Philosophy* 74, 153-61.

[9] Leisenring, A. (1969), *Mathematical Logic and Hilbert's* $\in$*-Symbol,* New York, NY: Gordon and Breach.

[10] Libert, T. (2003), 'ZF and the Axiom of Choice in Some Paraconsistent Set Theories', *Logic and Logical Philosophy* 11, 91-114.

[11] Priest, G. (1979), 'Logic of Paradox', *Journal of Philosophical Logic* 8, 219-41.

[12] Priest, G. (1987), *In Contradiction*, Dordrecht: Martinus Nijhoff; second, extended, edition, Oxford: Oxford University Press, 2006.

[13] Priest, G. (1991), 'Minimally Inconsistent *LP*', *Studia Logica* 50, 321-31.

[14] Priest, G. (1995), *Beyond the Limits of Thought*, Cambridge University Press; second, extended, edition, Oxford University Press, 2002.

[15] Priest, G. (2001), *Introduction to Non-Classical Logic*, Cambridge: Cambridge University Press.

[16] Priest, G. (2002), 'Paraconsistent Logic', pp. 287-93, of Gabbay and Guenthner (2002).

[17] Priest, G., Routley, R., and Norman, J. (eds.), *Paraconsistent Logic: Essays on the Inconsistent*, Munich: Philosophia Verlag.

[18] Rescher, N. (1969), *Many-Valued Logic*, New York, NY: McGraw-Hill.

[19] Restall, G. (1992), 'A Note on Naive Set Theory in *LP*', *Notre Dame Journal of Formal Logic* 33, 422-32.

[20] Routley, R. (1977), 'Ultralogic as Universal', *Relevant Logic Newsletter*, 50-89 and 138-75; reprinted as an appendix in *Exploring Meinong's Jungle and Beyond*, Canberra: Research School of Social Sciences, 1980.

[21] Slaney, J. (1989), '*RWX* is not Curry Paraconsistent', ch. 17 of Priest, Routley, and Norman (1989).

[22] Weir, A. (2004), 'There are No True Contradictions', ch. 22 of G. Priest, JC Beall, and B. Armour-Garb (eds.), *The Law of Non-Contradiction: New Philosophical Essays*, Oxford: Oxford University Press 2004.