GRAHAM PRIEST'S 'DIALETHEISM' — IS IT ALTHOGETHER TRUE?

Lorenzo Peña

SORITES Nº 7
Novembre 1996
ISSN 1135-1349
pp. 28-56

Contents:

§1.— **True Contradictions**

Graham Priest's book ***In Contradiction*** (Dordrecht: Martinus Nijhoff, 1987) is a bold and well argued-for defense of the existence of true contradictions. Priest's case for true contradictions — or «dialetheias», as he calls them — is by no means the only one in contemporary analytical philosophy, let alone in philosophy *tout court*. In some sense, other defenses of the existence of true contradictions are less philosophically «heterodox» than his is, since, unlike Priest's orientation, other approaches are closer to prevailing ideas in mainstream («Quinean») analytical philosophy, whereas Priest's leanings are strongly anti-realist, and not distant from the logical empiricism of the thirties.

However, such issues seem to me almost immaterial for the chief arguments in Priest's book. Much of what he says can be accepted from a wide variety of philosophical outlooks. And most of it seems to me right and important. Other among his arguments are less cogent, and yet can be modified and thus rendered far more convincing. Even when that is not so, weaker — less sweeping, but more plausible — arguments can be put in their place. And, what is more, such a watering-down does not affect the main conclusion of the book, the existence of true contradictions.

Whoever is prepared to accept thesis **ETC** — namely, **that there are true contradictions** — will meet strong opposition. A number of people will be flabbergasted. «But it is contradictory to say that there are true contradictions!». And? Well, as for arguments there is a deplorable shortage of them against **ETC**. Just claiming that **ETC** is contradictory can hardly be regarded as an argument. Priest has in other places tackled some of such arguments as there exist. Let me begin this critical notice with a short discussion of a couple of those preliminary objections against **ETC**.

The main argument is of course Aristotle's, which has always been taken to be an argument for the principle of noncontradiction, but which can be read as an objection against **ETC**. (Whether an argument against ⌜p⌝ is the same as one for ⌜not-p⌝ is a central issue, which will in due course be dealt with in this paper.) Aristotle's objection amounts to this, according to my exegetical lights (although of course there is an astonishingly wild variety of interpretations): if you are prepared to accept just one true contradiction, you are bound to accept any and every contradiction — and hence every statement —, since the only or at least the strongest reason against accepting a contradiction is just the fact that it is one; if that fact does not debar you from accepting it, you lose any reason for rejecting any other contradictory claim. Not that no other reason can be found, but none will be as strong as the mere fact that the belief — or the statement — is contradictory. If the strongest reason against a claim is not strong enough, neither are weaker reasons.

A possible reply is that for a belief to be contradictory is not bad, and so it is no reason against that belief; hence, it is downright false that contradictoriness is the strongest reason against a belief.

Such a reply does not seem to me to be open to G. Priest's peculiar brand of «dialetheism» — adherence to **ETC**. For he seems to share the view that true contradictions are bitter pills to swallow, which must be done only exceptionally in a few fields. Other things being equal we ought to refrain from accepting contradictions. So contradictoriness is after all a reason against a belief. Yet Priest clearly denies that it is the strongest reason. But no clear indication of what reasons against a belief are good or convincing is provided in Priest's book. (I'll dwell on this issue later on.)

A different objection against **ETC** is that whatever reasons there are for ⌜p⌝ are reasons against ⌜not-p⌝ and the other way round. Since reasons for ⌜p and not-p⌝ are only all reasons for ⌜p⌝ and all reasons for ⌜not-p⌝, it emerges that all reasons for ⌜p and not-p⌝ are also reasons against ⌜p and not-p⌝. Since there are other reasons against ⌜p and not-p⌝ (namely reasons for the universal truth of the principle of noncontradiction), the balance is definitely against ⌜p and not-p⌝.

Priest rejects the equation between **reasons-for-not-p** and **reasons against-p**. I think he is right, but unfortunately his dialetheism seems to me incapable of explaining what a reason against a belief may be. The notions of **rejection** and **being-a-reason-against** have to remain primitive (or explained one through the other), with no sentence being even available in the language by asserting which rejection could be expressed. That is unfortunate. What about the etymology of '**contradiction**' as **saying-against**?

Let me at this stage anticipate the core of may coming criticism. Priest does not accept degrees of «against-ness», degrees of negation, degrees of anything connected with the notions he explores in his book. Which blocks one of the most reasonable and straightforward roads to an accommodation with what the objection we are studying is up to. Should there be degrees of truth and falsity, we could say that a reason [strongly] against ⌜p⌝ is a reason not for ⌜not-p⌝ but for some stronger negation of ⌜p⌝ — namely ⌜not p at all⌝. Thus the objection would be rebutted on count of its failing to distinguish natural or weak from **strong negation** or **overnegation**. No such move is available to G. Priest.

Last, there is the hackneyed and by now discredited claim that from ⌜p and not-p⌝ everything follows. Again, yet, the claim depends on a number of steps, the only one which seems reasonably rejectable being **DS** (**Disjunctive Syllogism**). But rejection of **DS** is not easy. There are clear uses of **DS** which are right. Priest allows **DS** with a contextual proviso

— namely that contradictions in the field within which **DS** is being applied do not arise, or are likely not to arise. I'll discuss his position on that issue a little later.

Alternatively, if we had both weak and strong negation we could say that the only true contradictions which exist involve weak negation, whereas **DS** is valid only for strong negation. Which of course is no reply available to Priest.

Thus I conclude that the preliminary objections against **ETC** fail, and that there are not unreasonable ways of parrying them — even if the ones I prefer are not those which are in agreement with Priest's particular sort of dialetheism.

If the preliminary objections against **ETC** are far from being the knock-down arguments Aristotelians were fond of thinking they were, there are several good arguments for **ETC**. Priest's book displays several of them. Not that they are knock-down either — although Priest seems to me to think they are final. Acceptance of true contradictions is not going to offer a miraculous whole-sale solution to all problems in philosophy, but is going to make things much easier in all fields. After all, many convincing reasonings end in contradictory conclusions some of which can be maintained upon consideration, with a lessening of the strains and the constraints under which we used to find ourselves. At the very least, in many cases an alternative is now open to weighing, which was formerly unthinkable, due to acceptance of Aristotelian logic and its offspring.

---

§2.— **Are true contradictions in between complete truth and complete falsity?**

I find it odd that G. Priest, in his mustering of reasons for **ETC**, does not consider the notion of degrees of truth. Now, that notion is connected with the blossoming of fuzzy logics and set-theories — about which not a word is said in Priest's book — which have been extensively argued for in a variety of papers, and shown to have important applications both in theory and in practice (engineering). Priest cannot be unaware of those developments. He has clearly chosen to ignore them as pointless for his enterprise.

Yet, I feel confident that many readers will share my impression that the only non-classical truth value he accepts is precisely in between the extremes of pure truth and pure falsity; and so an intermediary value, which is less true than the wholly true.

Let me explain. Priest posits three truth-values: {T}, {F} and {T,F}. In order to give a uniform account we can identity T with {T} and {F} with F. Then we can say that the truth values are

(1) T, F

(2) for any two values, X, Z, X∪Z

There is no fourth truth value, since the union of any of the three with either of the other two is one of the three values.

Now, it seems obvious that what we have is a trichotomy of pure truth, pure falsity, and a mixture of them. What has a mixture of a property ϕ and its opposite is less ϕ than what only has the property ϕ and completely lacks its opposite. So the third truth-value is intermediary, a middle-course in between the extremes — in the same way as mixing white and black yields grey, which is less black than the black and less white than the white, but blacker than the white and whiter than the black. Mixing dryness with moistness gives something which is in between, humid.

Now, what is the reason for stopping at that stage rather than proceeding to introduce further intermediary truth values? Well, yes, according to the procedure, any new mixture is going to be just one of the three values — namely the mixed value {T,F}. But why not change the procedure slightly?

For instance, we can think in terms of multi-sets rather than sets — a multi-set being characterized by the fact that an entity may belong to it several times. Or we can directly think in terms of fuzzy sets — but, not to beg the question, I will not avail myself of them at this stage. Or we can take as the operation which generates new values, not union, but the pair-forming operation, { , }. So, in addition to T — which we no longer need to identify with {T} — and F — no longer to be equated with {F} — we have as additional truth values: {T,F}; {T,{T,F}}; {F,{T,F}}; {F,{T,{T,F}}}, etc.

It is not easy to understand what the last one is. We can impose a constraint which makes things clearer, by ruling out such combinations as {X,{Y,Z}} if X≠Y and X≠Z. Then we have infinitely many intermediary values which are clearly degrees of truth and falsity — except the original ones, T or total truth, and F or total falsity. {T,F} or ½ is equidistant between: T and F. {T,{T,F}} is equidistant between T and ½. {{T,F},{T,{T,F}}} is equidistant between ½ and the value which is equidistant between ½ and T. And so on.

Priest offers no reason against adopting a procedure like that. He does not consider doing so. He probably thinks that doing with just T, F and {T,F} is enough, that nothing new is either needed or desirable or even perhaps possible, since we already have Truth, Falsity and the Mixture of them. Yet we also have Wine, Water and … «the» mixture of them. What mixture? A 50-50 one? Is it the same as the 99-1 mixture? Are we to count all such mixtures as the same, because we do not care how much water there is, a small drop carrying the same weight as a million drops?

In most cases, if two qualities — or masses, or whatever — can be mixed, they can be mixed in many degrees of either. Certainty and doubt, love and dislike (or even hate), joyfulness and sadness, sweetness and bitterness, etc. Each mixture has a dose of either ingredient. Now, perhaps Truth and Falsity can never get mixed, as the classical logician contends. If they can consort with one another, why in such a way that degrees cannot be taken into account?

Since, according to Priest, true contradictions have both Truth and Falsity, they are true and false (although a problem arises here with what Priest calls the principle of exclusion, to which I'll come back later on). But, if such is the case, why cannot that mixture admit of degrees of each of the mixed properties — with the obvious constraint that, the more of the one, the less of the other?

By refusing such a gradualistic approach, Priest adopts a stand only slightly less rigid than the Aristotelian, for whom there are exactly two situations as regards truth: either it is [completely] present or else it is [completely] absent. Priest allows a 3d case where it is both or perhaps where both truth and falsity are present — but agrees with the Aristotelian (or classicist) in rejecting any further complication or any degrees of presence. Is our world-view much improved by accepting just a 3d, internally uniform, degree-less, mixture of heaven and hell, along with the originally given extremes, rather than a full scale of infinitely many gradations?

Against the foregoing considerations it can be argued that, truth and falsity not being mass terms, it does not make sense to speak of mixing, still less to talk of «more truth» as we talk of «more water».

Is that so? Well, I suppose my considerations bear a distinctive Platonistic ring: in some — perhaps non-literal — sense, when a thing is hotter than another one, there is more heat in the former — or maybe there is a greater presence of heat. If a proposition is truer than another one, there is more truth in the former than in the latter.

Well, yes, I know, not everybody is prepared to accept that a proposition can be truer than another: either it is true, *tout court*, or else it isn't. (But, please, notice that I of course accept the principle of excluded middle, and that not only nothing of what I've hitherto said runs against excluded middle, but in fact linking truth-graduality with true contradictions assumes excluded middle — else we could avoid the contradictions by contending that sentences with intermediary truth values can be neither asserted nor denied.)

It seems to me, though, that such a line is not open to G. Priest. He has clearly taken the true and the false to be mixable, his third truth value being an alloy of both, and being meant to be both. He even uses such expressions as «purely true» or the like for characterizing sentences with value {T}, as against sentences which are true but not purely true, namely those with value {T,F}.

Since unalloyed truth, pure truth, completely rules out falsity, whereas the mingle of truth and falsity which is {T,F} does not, we seem allowed to gather than sentences with value {T,F} are less true than sentences with value {T}, the latter alone lacking falseness altogether. Since being fully true is the same as lacking falsity altogether, only sentences with value {T} are entirely or wholly true — and only sentences with value {F} are utterly false.

Such considerations can be countered by insisting that a sentence with value {T,F} may be completely true — and completely false, too? Well, if it comes to a matter of definition, it is hard to find an argument on those matters. Yet, I feel that Priest cannot deny that the most natural reaction to his proposal to the effect that some sentences have as their truth value {T,F} is to view such a situation as a case of those sentences being neither completely (or purely) true nor entirely (purely) false, but in between, having both truth (to some extent) and also falsity (up to a point).

Natural reactions may be quite mistaken. Perhaps it is natural to extrapolate our visual field and conclude that the Earth is flat. But then there are arguments to the effect that the Earth is not flat (at all). Are there arguments which show that {T,F} is not a mixture of truth and falsity?

A different objection against my gradualistic construal of value {T,F} as a blend of truth and falsity is that there is no way to make sense of degrees of presence or any such Platonistic talk, which as such is merely metaphoric. But that is wrong. The set-theoretical approach I have sketched is a way of cashing the metaphor. (Moreover, a Platonist needn't even fall back on such a re-wording.)

Therefore, pending arguments against the gradualistic interpretation, it seems to me we can be confident that what in effect Priest is putting forward is the existence of cases of truth which are not cases of complete truth; cases of something being partly true only. Is it reasonable to assume that all beliefs which are true but not purely true are equally true, none of then being truer than another?

Let us come to specifics. One of the main intended applications of Priest’s defense of **ETC** is semantics. He thinks — rightly, to myself — that the sentence ‘This sentence is false’ is both true and false. If we have an implication, ‘$\rightarrow$’, read as ‘to the extent that…, ---’, or the like, then we can construe the (simple) liar in this way.

Graham Priest emphasizes that in any language rich enough to contain Peano arithmetics and endowed with a truth-predicate ‘T’, we obtain a contradiction (p. 99). For any sentence of the language, $\ulcorner p \urcorner$, let ‘#p’ be the numeral of the Gödel number assigned to sentence $\ulcorner p \urcorner$ (under some definite way of codifying expressions into numbers, be it Gödel’s original one or any other). Let us use the fixed-point theorem and the diagonalization technique in order to construct an open sentence $\ulcorner \phi(x) \urcorner$ which is true **to the extent**, and to the extent only, that (loosely speaking) its diagonalization is not true; less inaccurately, let δ map the Gödel number of an open sentence with one free variable, ‘v’, into that of this diagonalization, i.e. of the result of substituting in it the numeral of its Gödel number for its only free variable. In virtue of the diagonal lemma, if $\ulcorner \alpha \urcorner$ is any formula with one free variable ‘v’, there is a sentence $\ulcorner \beta \urcorner$ such that $\vdash \beta \leftrightarrow \alpha(v/\#\beta)$. Let us take the open sentence $\ulcorner {\sim}Tx \urcorner$, where we have laid down that $\vdash T\#p \leftrightarrow p$ as an axiomatic schema. It follows that there is a sentence $\ulcorner p \urcorner$ such that: $\vdash {\sim}T(x/\#p) \leftrightarrow p$. In virtue of the axiomatic schema $\vdash T(\#p) \leftrightarrow p$, we’ll have:

(1) $\vdash T(\#p) \leftrightarrow {\sim}T(\#p)$ (for some «p»)

What does (1) mean? If we cleave to our proposed reading of ‘$\rightarrow$’ (and hence to that of ‘$\leftrightarrow$’ as ‘to the extent, and to the extent only, that’), it means that the sentence $\ulcorner p \urcorner$ under consideration is true to the extent that it is not true, and conversely; hence as true as not true, neither more nor less. Let us grant all that. But if there is such a sentence, equally true and false, why not a sentence slightly more true than false, and one in between being slightly more true than false and being wholly true, and so on?

What is unique to the liar is that it says of itself that it is not true. Hence, in virtue of $\vdash p \leftrightarrow p$, and substitutions, we have that to the extent it is true, it isn’t, and conversely. Nothing like that is available for sentences which are more true than false, or more false than true. Still, if the liar exists, why not those others?

The issue is not whether we can prove merely by such means that such intermediary cases exist. After all, proving the liar depends on a number of very debatable assumptions — although I am confident Priest has shown that the usual ways-out are not as good as their adherents are keen on thinking. I do not deny the naturalness of proving the truth-and-falsity of the liar. The point is that, once its existence has been granted, the plausibility of further intermediary cases is much enhanced.

§3.— Avoiding ineffability and the need for strong negation

One of the most forceful and recurring arguments throughout Priest's book is that the usual ways-out lead to ineffability. I am not going to repeat Priests' detailed and convincing arguments. In fact such a conclusion should be obvious. If there is no language which is the metalanguage of all languages, in what language can such a non-existence be said? It must be a language wherein we can speak of all languages and their semantic qualities. The «hierarhist» approaches — as G. Priest calls them (p. 24) — cannot even say of themselves that they are true. They have to fall back on unquantified schemata, through «systematically ambiguous» expressions. The problem is similar to that encountered by type distinctions, but not quite the same.

I find Priest's arguments congenial and very plausible. Yet his way of putting all that in terms of all or nothing seems to me unfortunate. You get the impression that either a sweeping, naive, thoroughgoing approach to semantics is accepted, and then the paradoxes ensue and are also espoused — which of course calls for a paraconsistent logic — or else we have to cling to some variety of the hierarchy, be it Tarski's or Kripke's or whatever. But surely there are hierarchies and hierarchies. Not all of them are as harsh and unpalatable as Tarski's. In fact Tarski's was a crude, extreme reaction, whereas later approaches are milder, more refined, less destructive. Priest himself acknowledges that some of them admit of truth-value gluts rather than gaps (see p. 26, n. 20 on Woodruff's treatment). And of course new approaches can be devised, by refining, or qualifying, those which are available.

Against adherents of truth-value gaps, Priest convincingly argues that they cannot express that a sentence is not true (p. 20). Unfortunately though, he encounters exactly the same situation. For he needs to differentiate between that which is only false and that which is both true and false. He does so by using expressions such as 'only', 'purely', 'plain[ly]' (p. 239) and the like. Now, what is a plainly false sentence? One which is true and which is not false? If we could say that, and by saying so enough information were to be provided, it would be fine. Can we? Not if we accept the **exclusion principle**, namely that, to the extent that a sentence is false, it is not true. The immediate effect of the exclusion principle is that the contradictions spread to the «metalanguage» — speaking in the customary jargon. For let ⌜p⌝ be a sentence both true and false, i.e. with value {T,F}. Then the sentence «This is true: ⌜p⌝» will also be true and false. By saying that ⌜p⌝ is true and not false we say nothing incompatible with ⌜p⌝ having as its truth-value {T,F}, and with so doing Tp.

Dropping the exclusion principle avoids such spreading of contradictions into the meta-language at a high price. Rather than having the T schema for mutual (contraposible) coimplication, '↔', we are supposed to do with some makeshift, to which contraposition does not apply (see pp. 88-91 and 99-100). Priest says (p. 100).

> There seems to be no reason why, *in general*, if α is a dialetheia, Tα is too. If α is a dialetheia, Tα is certainly true, but it might be simply true, and not also false. The truth predicate is therefore a *partial* consistensizer.

Priest contends that the exclusion principle spreads contradictions beyond necessity (p. 90). 'On the basis of this I tentatively reject the exclusion principle' (ibid.). It never emerges whether in the end Priest comes to accept the principle.

It is not just a matter of definition. If the exclusion principle fails, many arguments for true contradictions which hinge on the T schema also fail and have to be reformulated. (The reformulation involves more debatable principles (pp. 162-3), so those who oppose Priest's dialetheistic solutions are provided with a number of possible and plausible retreats.) Moreover, the fundamental idea that the truth of ⌜p⌝ is the fact that p is no longer correct.

Furthermore, Priest's purported account of the T predicate without the exclusion principle in effect restores the principle in other form, since it countenances the **antisatisfaction** principle, namely that a sequence antisatisfies a formula iff it satisfies its negation ('iff' is my own reading of Priest's non-contraposible biconditional '⇔'). What emerges (p. 176, bottom) is that we have two predicates, T and F, such that ⊦p⇔T#p and ⊦~p⇔F#p. The proof that the thus weakened truth theory still contains contradictions involves abduction rules for the noncontraposible conditional '⇒', namely p⇒~p ⊦ ~p and ~p⇒p ⊦ p — although the point is not made quite clear in the text. The whole treatment is somehow marred by the fact that '⇒' is not provided with an English reading; and once contraposition has been junked for it, I surmise that not everybody will accept the abduction rules. For what alone is proved is (p. 177) $\mathcal{S}at_1$(x,#α)⇔α(v_i/x); i.e. x satisfies (in one place) [the Gödel number of] formula 'α' iff α(x); by instantiating 'x' with '#[~$\mathcal{S}at_1$(v_i,v_i]' and 'α' with '~$\mathcal{S}at_1$(v_i,v_i)', we are supposed to obtain a contradiction. Yet the contradiction follows only in virtue of the abduction rules (Clavius). If one relinquishes the naive simplicity underlying the original T schema — namely that ⌜p⌝'s truth is just the fact that p —, I guess some contradiction-averting manoeuvres become less implausible: e.g. refusing to accept the antisatisfaction principle, or one or other of the abduction rules for conditional '⇒'.

Hence, I find waiving the exclusion principle unattractive. Nonetheless, even without the exclusion principle we still need to differentiate between situations which are only false and such as are both true and false. Suppose that ⌜p⌝ is both, i.e. that it has truth value {T,F}, while both T#p and F#p are only or plainly true, with no admixture of falseness in them (!). Let ⌜q⌝ be such that its truth-value is {F}. How can we differentiate their values? Not through F, since both F#q and F#p, but through T. But then we shall have a way of expressing a strong negation. Let us define 'H' in this way: ⌜Hp⌝ abbr ⌜T#p∧~T#(~p)⌝. Notice that 'H' is an operator, not a predicate. Yet, within the whole arithmetic-cum-semantic theory Priest provides, it is definable. 'H' is strong assertion. The rule Hp, ~p ⊦ q is truth-preserving — the premises cannot be both true. Strong negation is definable: ⌜¬p⌝ abbr ⌜H~p⌝. The Cornubia rule for strong negation (viz. p, ¬p ⊦ q) is also truth-preserving (for the same reason, of course): p , ¬p ⊦ q. Those rules can of course be avoided by imposing conditions on the turnstile over and above mere truth-preservation. But I do not see what further requirements Priest imposes.

Now, let us considerer the sentence:

(L) ¬L

(L) says of itself that it is completely non-true, i.e. that it is not at all true that it holds. If all the semantic machinery Priest has developed is still available at this stage (and how could it have broken down by now?), an **overcontradiction** can be proved, namely: L∧¬L. (An overcontradiction is simply a contradiction involving strong negation.) Hence q (every ⌜q⌝). Unless, … Unless we impose further conditions on the turnstile (but is that really a solution or a mere stipulation?) or revert to the initial T schema, with ⌜T#p⌝ having the same truth value as ⌜p⌝. Which would mean that we accept the exclusion principle after all.

With the exclusion principle, however, we can no longer — within Priest's account — differentiate truth *tout court* from plain truth. It does not help to say that ⌜p⌝ is true and not false; that will be the case, too, if its truth value is {T,F}: it will both have and fail to have truth; the latter is true, for ⌜~T#p⌝ is — in accordance with the exclusion principle — implied by ⌜T#~p⌝, which is true if the value for ⌜p⌝ is {T,F}.

A strong negation is needed, one '¬' such that ⌜p⌝ completely rules out ⌜¬p⌝ and the other way round. With strong negation we can explain the difference between being true and being plainly (i.e. completely) true; and between being false and being downright false. Then many things come into place. We'll have a criterion on when **DS** can be relied upon (pp. 137ff): whenever the negation involved is strong, or can be taken to be strong. We know when there is an argument against a claim: whenever there is one for the strong negation of the claim. We know why rejection and acceptance are fully incompatible (pp. 128-32): rejecting ⌜p⌝ means or entails accepting ⌜not p at all⌝, i.e. ⌜¬p⌝. We know why Priest's principle R (see p. 141) holds, namely 'If a disjunction is rationally acceptable and one of the disjuncts is rationally rejectable, then the other is rationally acceptable'; the reason it holds is that **DS** is valid for strong negation.

More than that is gained with adding strong negation. **CL**, all of it — including *Modus Ponens* for '⊃', provided ⌜p⊃q⌝ abbreviates ⌜¬p∨q⌝ — is now incorporated into the paraconsistent system. **CL** is shown to be, not wrong but poor, insufficient. The classicist can be placated; he can be asked only to refrain from reading '¬' as 'not'. The resulting system is more ecumenical.

Yet, of course, there is a price, a high price. We can no longer accept the T predicate (within a language sufficiently powerful) or the comprehension axiom in set theory (on which, see hereinbelow), or Priest's elegant and simple treatment of the Gödel sentence B, namely «This very same sentence cannot be proved». For in each of those cases, putting strong negation in the place of natural negation brings about an overcontradiction — unless the principles in question are weakened or somehow qualified.

Priest seems to me implicitly committed to having strong negation. On p. 146 he introduces a propositional constant 'F' such that for all sentences ⌜p⌝ ⊢F→p. He stresses that if the language contains its own truth predicate the constant 'F' can be defined as '∀xTx'; the characteristic principle is then proved. Such being the case, Priest's system — once it encompasses arithmetics and the truth predicate — does in fact comprise strong negation; for let ⌜¬p⌝ abbr ⌜p→F⌝. Priest is right when he thinks that rejection of contradictions in general is expressed by the schema ⌜p∧~p→F⌝ (which amounts to ⌜¬(p∧~p)⌝; i.e. every contradiction is completely false). Replacing '~' with '¬' yields a formula which expresses rejection of overcontradictions, namely ⌜¬(p∧¬p)⌝ (overnegation of any contradiction involving overnegation). Although Priest's own system shuns ⌜¬(p∧¬p)⌝ as a theorem — by dint of avoiding conjunctive assertion (see below, §5) — it is committed to something close, namely ⌜¬p→.p→F⌝.

Thus Priest is in a dilemma. If he accepts constant 'F' and so strong negation (or if it is true that 'F', with its characteristic principle, is already present in his whole system), then something not far distant from **CL** is included, the naive and simple arguments for

ETC in semantics and set theory are no longer available, and anyway semantics and set theory need some further remedies over and above the mere acceptance of contradictions (contradictions involving '~' but not '¬', i.e. not **over**contradictions). If he keeps clear of strong negation (and of constant 'F', which I doubt he can do, and so does he), then ineffability results ensue, which lessens the effect of his criticisms of any hierarchist way-out of the semantical paradoxes; moreover, all the whole issue of **arguing-against** (p. 141), the rationale for principle R, rejection, etc., becomes misty if not enigmatic.

---

§4.— **Set Theory**

Although semantics is the main reason for Priest's dialetheism, set theoretic paradoxes also feature among the grounds for his approach. And rightly so. He discusses the abstraction principle under the form ⌜$\exists y \forall x (x \varepsilon y \leftrightarrow \beta)$⌝ and extensionality. He launches (pp. 37ff) an onslaught on the cumulative conception of sets implemented in ZF. His criticism is, to myself, quite cogent. Then he espouses (pp. 178ff) the abstraction principle and hopes that his formal system Δ can admit it without **deliquescence** (i.e. Post-inconsistency, or triviality, as he says following the current fashion). His reason to hope is that Δ is very close to a relevant logic, DK, which has been shown by Brady to be compatible with the abstraction schema.

Priest's objections against the cumulative hierarchy seem to me so obvious that I find it amazing, not that a number of mathematicisms **use** ZF, but that some philosophers take it as what it was never meant to be (not by Zermelo anyway), namely an «intuitive» conception of what sets are. Priest's objections can be strengthened. If something like the temporal metaphor he rightly denounces (p. 39 bottom) is to be taken ever so little seriously, then the idea is a constructivistic one. But then quantifiers cannot be allowed to range over **all** sets, but only over sets which «already» exist (i.e. a predicative set theory is required, which ZF and the like are not).

A different objection against ZF is that it lacks not only a universal set but complements (except relative complements). Not only it cannot be the set theory used in its own metatheory, but it cannot provide a satisfactory semantics for internal negation — there being no set which comprises only all entities that not p, for any ⌜p⌝.

However, Priest fails to discuss Quine's systems NF and ML for the reason that they are 'widely regarded as little more than… curiosit[ies]' (p. 38). Yet, if **CL** is true, and if some sort of Tarskian semantics is to be possible, the hierarchical cumulative approach is bound to be wrong, and something like NF of ML right. Philosophically those systems are infinitely more appealing than ZF, and deserve to be discussed. That NF entails some oddities concerning the ordinals, is of little or no concern, since the whole subject is anyway riddled with surprising results.

Priest does not discuss non-well-founded set theory or Fitch's combinatory logic, either. His case would be much stronger if he took such alternatives into account. Neither does he discuss other paraconsistent approaches, which accept the abstraction principle with some restrictions. Yet I guess what leads him to ignore such approaches.

His line is straightforward and clear. Take the abstraction principle in its pristine and unpolluted purity; a contradiction ensues. Tampering with Abstraction, as in ZF, causes a lot of trouble and extremely undesirable results. Hence.

But suppose that we have both natural negation '~' and strong negation '¬', as we have seen in the previous section. Then Abstraction has to go. Yet we could have something weaker.

Let us discuss, not abstraction, but Comprehension. Let {…:---} be a vbto (variable-binding term-forming operator), which we take as primitive (we could use λ notation). We need a set theory with these principles, or qualified versions thereof:

(Existence) $\exists z(z=\{x:p\})$

(Comprehension) $\{x:p\}x\leftrightarrow p$

(with 'z' not free in ⌜p⌝ in the former principle), and membership being expressed as 'zx' rather than 'xεz').

With strong negation, catastrophe!

$\{x:\neg(xx)\}\{x:\neg(xx)\}\wedge\neg\{x:\neg(xx)\}\{x:\neg(xx)\}$

Hence ⌜q⌝. Suppose that we qualify Comprehension by requiring that ⌜p⌝ contains neither '→' nor strong assertion 'H' (and hence not '¬' either). But we can add half of it without qualification, namely: ⌜$p\rightarrow\{x:p\}x$⌝.

The system — which can be implemented on the basis of some strengthening of relevant logic E — is reasonably strong. We can prove that the set of such sets as do not comprise themselves both comprises itself and does not. We can also prove that the set of such sets as do not comprise themselves at all comprises itself (although of course we cannot prove that it does not comprise itself).

Let us call 'crowds' such sets as our theory, thus conceived, would account for. Crowds can be thronged or [over]crowded. They can let in entities which utterly fail to comply with the entry-condition — with the characterizing matrix or formula. Such cases can be exceptional. We can add a number of particular cases of (full) Comprehension for which we satisfy ourselves that no overcontradiction is going to arise.

Priest does not discuss any such proposal but it is clear why he finds them distasteful. Gone with them is the straightforward argument from Abstraction to dialetheism. The classicist can retort that if so-called strong negation is debarred from featuring in the two-ways Comprehension principle, why not so-called natural negation, too? Are they not on a par? If from applying comprehension to {x:¬(xx)} everything follows, why not from applying it to {x:~(xx)} — '~' having the property of negation in relevant logic **E**? Are those not exactly the same pattern of inference?

No, they are not. An inference pattern is syntactically characterized, as is every proof-theoretic notion. And the difference between the two negations can render one of the patterns correct while the other is wrong.

— «But you do not intuitively see that one is right and the other wrong. You show such conclusion from the fact that within such a system — a strengthening of E —, enriched with strong negation as you call it, from {p, ¬p} everything follows, whereas nothing of the sort happens with {p, ~p}. That is no longer a matter of intuitive principles, but an artificial contrivance. And then why bother us with nonclassical negation at all? Don't you have the worst of both worlds?»)

Yet Priest's approach, thus mollified (or watered-down), could become more attractive. After all, his current approach has it that R (i.e. {x:~(xx)}) both comprises itself and does not. As much is true — according to him. Necessarily true (I gather). Better to recognize a necessary truth than to say that it is so false that everything follows from it — which is what the Aristotelian does, when he mistakes his '¬' for 'not'. A theory which countenances RR∧~(RR) is better than one which enforces RR∧~(RR) $\vdash$ q. So the considered theory of crowds is not worse off than classical set theories (with the possible exception of Quine's — about which I shall say nothing here, sine Priest does not consider them).

Besides, crowd-theory would accept a universal set. We would be rid of the troubles which arise from ZF's necessarily rejecting the existence of a universal set. Gone is the need for the hierarchy. So there are not one but several respects in which crowd theory would be better than standard set theory based on **CL**.

On the other hand, since we'd have strong negation, such difficulties surrounding Priest's approach as have emerged in the last section would be overcome.

We could conceive of a similar treatment of semantics. Rather than the wholesale unqualified schema T, restricting it to sentences with no occurrence of either '→' or strong negation (or strong affirmation), while keeping one half for all sentences:

p→T#p

Then the simple liar would be as true as false (and hence both) but the strong liar would be only true. (The exclusion principle could — but needn't — be dropped, as Priest advises us to do.) Of course, such a retreat leaves us without the sweeping and direct reasoning from semantics to the existence of true contradictions Priest is keen on. We now need more sophisticated arguments to the effect that natural negation is all right for the T schema, but strong negation can be admitted only into half of it. (Well, a reason is that overcontradictions ensue from admitting strong negation, not natural negation, and that for someone to have a brother who causes trouble and hence is not admitted into the best circles is no sufficient ground for her to be also ostrasized.) We could reinforce predicate 'T' with a number of approximations to the original T schema, short of countenancing it in its pristine generality, — which would thus become a regulative ideal to which we would tend asymptotically.

Furthermore, the resulting approach would keep a lot of what Priest has put forward. It is not as if such an approach would have nothing to do with his philosophical enterprise. While obviously closer to canons dear to the **CL** enthusiast, the approach would also accept qualified versions of Priest's principles and most of his conclusions. Most of the contradictions he is keen on proving in his book would remain welcome. Only, things would be a bit less simple and straightforward. And his *a priori* view of logical and mathematical truth would be jeopardized, which of course he does not want. (For you cannot say that your «intuition» tells you, analytically, that the T schema is right under exactly such qualifications and no others.)

§5.— **Conjunctive Assertion**

Although in his chapter on set theory (pp. 178ff) Priest focuses on how the mere acceptance of contradictions — and hence the rejection of the Cornubia rule — makes it possible to have a full-fledged, unqualified comprehension axiom, things are not that simple. As Priest himself has made abundantly clear in a previous chapter (pp. 103ff) the Curry paradox compels any set theory with an unqualified comprehension axiom (and extensionality — or even weakened versions of extensionality) to have as its underlying sentential calculus one without the **contraction** schema $\ulcorner p\rightarrow(p\rightarrow q)\rightarrow.p\rightarrow q\urcorner$. Priest gives us 'the strongest form of it' (p. 103): **conjunctive assertion** — **CA** for short —, namely $\ulcorner p\rightarrow q\wedge p\rightarrow q\urcorner$.

Although **CA** entails **contraction** only once other principles have been countenanced which are not unanimously agreed upon — e.g. any result of adding reduced factor ($\ulcorner p\rightarrow q\rightarrow.p\rightarrow.p\wedge q\urcorner$) to the set of axioms and inference rules of **E** minus **contraction** —, yet, since I accept the implication of **contraction** by **CA**, I agree with Priest that the latter is — for our current purposes anyway — what is really at issue. Moreover, in a rule form we can deduce **contraction** from **CA** under very weak entailment principles (see e.g. Richard Routley, ***Exploring Meinong's Jungle***, p. 917, bottom); i.e. from $\ulcorner p\rightarrow.p\rightarrow q\urcorner$ we infer $\ulcorner p\rightarrow q\urcorner$ with the aid of **CA**, even if we do not have as a theorem schema $\ulcorner p\rightarrow(p\rightarrow q)\rightarrow.p\rightarrow q\urcorner$.

Giving up **CA** seems to me too high a price for anything you can win by such a move. Iff $\ulcorner p\urcorner$ can be **strongly inferred** from set A of premises (i.e. inferred in such a way that the degree of falsity of the conclusion is at most as high as the one of the falsest premise) is it the case that, if $\ulcorner q\urcorner$ is the conjunction of all formulae in A, $\vdash q\rightarrow p$. Since *Modus Ponens* for implication is a strong inference rule in that sense, **CA** is bound to obtain. (My talk of degrees may sound distasteful to relevantists — although Priest is not one of them — but something similar, if worded in somehow different terms, is the idea underlying Anderson & Belnap's famous Entailment thesis.) A proof theory for a logic without **CA** is then bound to be weak or awkward. And, most of all, the plausibility of **CA** much exceeds that of any set-theoretical principle. For, suppose that an instance of **CA** fails (completely — i.e. its strong negation is true). Then although $\ulcorner p\rightarrow q\urcorner$ and $\ulcorner p\urcorner$ are both true (since $\ulcorner p\rightarrow q\wedge p\urcorner$ is true), $\ulcorner q\urcorner$ is either not true at all or else: less true than $\ulcorner p\urcorner$ and less true than $\ulcorner p\rightarrow q\urcorner$. But if $\ulcorner q\urcorner$ is less true than $\ulcorner p\urcorner$, $\ulcorner p\rightarrow q\urcorner$ is utterly false. (Again my talk in terms of degrees is not that essential; Anderson & Belnap in their construction of system **E** would put it in different words, with substantially the same conclusion.)

And, if **CA** fails, so does importation, of course ($\ulcorner p\rightarrow(q\rightarrow r)\rightarrow.p\wedge q\rightarrow r\urcorner$) or else self-implication ($\ulcorner p\rightarrow p\urcorner$). (Perhaps importation is the strongest form of the principle we are now discussing; it is no coincidence that e.g. Łukasiewicz's logics contain neither importation nor **CA** nor **contraction**.)

Thus the price of having an unqualified comprehension principle is much higher than Sect. 10.1 of the book («Naive Set Theory», pp. 178-80) suggests.

Much the same can be said of semantical paradoxes. There is a semantic duplicate of the Curry paradox (initially I think, put forward by P. Geach in ***Logic Matters***, pp. 209-11). Priest gives it in these terms (pp. 103-4): given any arbitrary sentence, β, by diagonalization, self reference or a similar device, we can find a sentence, δ, of the form T#δ→β. The T schema for this sentence yields: T#δ↔.Tδ→β. With contraction we prove ⌜β⌝ (the details are left as an exercise to the reader). Thus the system becomes deliquescent. Everything can be proved.

Thus we ought to choose. Either **CA** or unqualified comprehension and unqualified T schema, not both (not **at all** both). I really feel at a loss as to how to conceive implication without **CA**. Of course, Priest is not the only logician who rejects **CA**. Most characteristically the principle has been fought by Richard Sylvan in his endeavour to develop **deep relevantism**. In ***Exploring Meinong's Jungle*** (authored with his then name, 'Richard Routley'), p. 919, Sylvan argues in this way against **CA**:

> [**CA**] would exclude situations of the type which occur with semantical paradoxes, where A→B and A both hold but B fails to hold, that is, where an implication which holds is also counterexampled.

Let us critically examine the argument. What does 'fail' mean here? If strong negation is let in — which of course runs against relevantism in general and more so against deep relevantism —, then that B fails probably means that ¬B holds — where '¬' is strong negation, 'not… at all'. But then of course, if B fails to hold, it cannot be true **at all** that both A and A→B hold. On the other hand, if 'fails' is taken here in a weak sense, i.e. if for B to fail is just that ~B is true — '~' being natural negation, the mere 'not' —, then what B's failing entails is that either A→B or A fail **in the weak sense**, i.e. that either ~(A→B) or ~A is true. But that one of them is true does not [completely] rule out A and A→B being both true, too.

A further broadside against **CA** is displayed in ***Relevant Logics and their Rivals*** (by R. Routley, V. Plumwood, R.K. Meyer & R.T. Brady, Atascadero: Ridgewiew, 1982, pp. 278ff). As the argument makes it clear, in order for one to junk **CA** it suffices, within a modal modelization, to relinquish reflexivity of the accessibility relation. (And in fact in Priest's own modal semantics for his paraconsistent system Δ — propounded in his book, pp. 106-8 — reflexivity of the accessibility relation R does not hold, even though our world accesses all worlds, and so also accesses itself.) Under relevant modellings — I quote ***RLR***, p. 279, top — the three-place relation has to be reflexive for **CA** to hold generally, i.e. Raaa has to obtain for every [world or setup] a. The argument goes on:

> Such reflexivity requirements in fact impose serious restrictions on the class of situations admitted in semantical evaluation. They have the effect of ruling out various paradoxical situations. The simple relation Raa includes [excludes?] various paradoxical worlds in rather the way that the simple equation a=*a excludes inconsistent worlds.

What is thus claimed is not that **CA** is incompatible in general with negational inconsistency — 'in rather the way that the simple equation a=*a excludes [upon Routley-Meyer semantics with usual constraints] inconsistent worlds'. The context makes it abundantly clear that what is at issue is Curry's paradox. What emerges is that **CA** rules out situations which would render our theories deliquescent should we have **both** unqualified comprehension or the T schema **and CA**. But it is not enough to devise some modelling according to which **CA** fails, since — as ***RLR*** itself puts it a couple of pages later (p. 281, paragraph 2):

> such an argument has its dangerous aspects, especially as we can now countermodel virtually any logical principle. Accordingly a general objection to semantics which falsify entrenched principles like [**CA**] takes the following turn (…)

What follows is a quotation of a previous paper by G. Priest («Sense, Entailment, and *Modus Ponens*», ***Journal of Philosophical Logic*** 9 (1980), pp. 415-35). The gist of the argument is that saying 'We are more sure of the Truth of [**CA**] than of any theoretical account of logical truth' is, in the context of the present discussion, to beg the question.

Is it? Suppose — as Priest claims — that the two incompatible claims — **CA** on the one hand and the T schema and Comprehension on the other — are, as he puts it, obvious. The argument in support of **CA** points to the fact that it is **more** obvious. Again, I feel that, when degrees are ignored or overlooked, things go awry.

In the end, ***RLR*** rounds out the discussion of **CA** with a simple remark, 'observing the damage [**CA**] wreaks in paradoxical situations is enough to shake confidence in it and to begin to shift the onus of argument'. And the authors add that *Modus Ponens* 'no more supports its «normalization» [**CA**] than Material Detachment licenses Disjunctive Syllogism'. Quoting Russell they show that, unlike the rule of **MP**, **CA** merely requires the **hypothesis** that A is true; 'in short, it applies in situations beyond the actual one'.

The idea is that in any framework the world or setup, $\mathfrak{w}$, which there plays the role of «the actual world», has to be closed for **MP** (i.e. if both ⌜p→q⌝ and ⌜p⌝ hold at $\mathfrak{w}$, so does ⌜q⌝), but other worlds may fail to enjoy such closure. Is that idea plausible?

There are several usually accepted principles which taken together rule out the idea:

(**T1**) If an absurdity follows from a hypothesis, the hypothesis cannot be true.

(**T2**) What cannot be true is impossible (i.e. is not possibly true).

(**T3**) What follows from a hypothesis also follows from the hypothesis being possibly true.

(**T4**) ⌜q⌝ follows from ⌜p→q∧p⌝.

(**T5**) It is absurd that ⌜p⌝ holds and that ⌜p⌝ [utterly] fails to hold.

(**T6**) If from a hypothesis several consequences follow, then from the hypothesis it also follows that the conjunction of those consequences follows.

Suppose now (*Hyp*):

(*Hyp*) ⌜p→q∧p⌝ holds while ⌜q⌝ utterly fails to hold.

Obviously from (*Hyp*) it follows that ⌜p→q∧p⌝ holds; whence (by (**T4**)) it follows that ⌜q⌝ holds. Now from (*Hyp*) it follows, too, that ⌜q⌝ utterly fails to hold. The set of both consequences is an absurdity (by (**T5**)), which follows from (*Hyp*) (in virtue of (**T6**)). But in virtue of (**T3**) the same absurdity follows from ◊(*Hyp*). Hence ◊(*Hyp*) cannot be true (by (**T1**)). Therefore (by (**T2**)) it is impossible that (*Hyp*) be possibly true.

It seems to me that among such assumptions what is rejected by deep relevantists is (**T3**). According to them it would indeed be absurd that ⌜p→q⌝ and ⌜p⌝ hold while ⌜q⌝ didn't [at all], but not that a situation should be **possible** wherein ⌜p⌝ and ⌜p→q⌝ hold but ⌜q⌝ doesn't [at all].

My reply is that, if «the actual situation» was a logically privileged one, other situations wouldn't matter. But it isn't. What cannot hold in the actual situation, whatever it may be, cannot hold either in a possible but nonactual situation.

What is right, though, is that the rule of **MP**, $p \rightarrow q, p \vdash q$, allows us to draw $\ulcorner q \urcorner$ from the truthful assertability of $\ulcorner p \rightarrow q \urcorner$ and $\ulcorner p \urcorner$, whereas **CA** does more than that. Thus, if we have a connective 'B' meaning 'It is truthfully assertable that' (with the rule $p \vdash Bp$, even though we do not have $\vdash p \rightarrow Bp$), **MP** only requires a qualified version of **CA**, namely (**T8**):

(**T8**) $\ulcorner Bp \wedge B(p \rightarrow q) \rightarrow Bq \urcorner$

(**T8**) is equivalent to $\ulcorner B(p \rightarrow q \wedge p) \rightarrow Bq \urcorner$; whereas [unqualified] **CA** requires (**T9**), namely:

(**T9**) $\ulcorner B(p \rightarrow q \wedge p \rightarrow q) \urcorner$

(**T8**) is weaker. But as for rendering deliquescent any system with a full-fledged T schema or an unqualified Comprehension principle, (**T8**) would be enough. Thus weakening **CA** in such a way would be of no avail.

The comparison with **DS** seems to me wrong. No relevant similarity. If there are true contradictions — and, as Priest and I contend, some of them exist necessarily — then **DS** is not a correct deductive inference rule, whereas **MP** (for implication '$\rightarrow$') is.

Finally, against both Priest's argument and that of the authors of ***RLR***, it must be emphasized that **CA** is more basic, more general. It is a principle of sentential logic. Admittedly, sentential logic is also liable to feedback from its applications; so there are reasons to qualify some principle of the sentential calculus in virtue of considerations pertaining to its applications in a number of fields. But there must be very very strong reasons for that. And whenever possible a distinction ought to be offered by which the junked principles can be retained under some reading (e.g. when we reject principles involving negation, a distinction between weak and strong negation allows us to keep all classical principles for strong negation). Sacrificing **CA** just for the sake of coping with Curry's paradox seems to me as *ad hoc* as anything can be in logic.

---

§6.— **Motion**

Traditionally, Zeno's paradox of the arrow has been associated with truth at intervals, not instants. Spinoza put it like that: there is no fixed unique position that the travelling body occupies at an interval, for the interval is made up of infinitely many subintervals, and the body is not at the same position at them all.

However, Priest argues for the contradictoriness of motion on the basis of instant-semantics (pp. 221ff). His version of the arrow argument hinges on the **spread hypothesis**, **SH**, (ibid.), namely:

SH A body cannot be localised to a point it is occupying at an instant of time, but only to those points it occupies in a small neighbourhood of that time.

Let us modify the wording by replacing 'point' and 'points' with 'place', or 'region', or the like (a body cannot be contained in a point).

What is the rationale for **SH**? Priest argues that the difference between motion and not-motion is that something like **SH** obtains for the former. A world wherein there would be a sequence of states cinematographically describable as if there were motion in it would all the same be a motionless world. I agree. A body is not moving if it is now here, and only here, then there and only there, never really passing from here to there — never in a situation which can be described as intermediary between being here and being there.

Yet there are problems. Let us suppose b is travelling from its initial position p_1 to its destination p_n. At each instant i it occupies not just a position p_i but also positions which are in a neighbourhood of p_i — and which partly overlap p_i. Yet, since no degrees of truth are taken into account, b is equally at all those places. And since Priest tells us that the spread is bound to be small, for each point outside the series of stretches b occupies at i it is downright and wholly false that such a point lies in one of the position b occupies at i. There is a clearcut, crisp, trenchant boundary separating b's series of positions at i from the positions it does not have at i [at all].

All of that seems to me implausible. The rationale for taking motion to be contradictory was that we cannot ascribe to the travelling body a unique position at an interval. With instants that is not clear — although probably the best way to understand what «happens» at an instant is a derivative one. Since Priest himself allows of a consideration of intervals, let me henceforth regard **SH** with a further modification, reading 'interval' for 'instant'.

Now, b is travelling (with uniform speed) and its travel begins at noon and ends at 13pm. Let b be of length l. At the interval I, 12:25—12:35, it occupies a number of partly overlapping positions, the stretch they form being that between points x and z. Let p_1 be a position of length l such that the distance between x and the first point of p_1 equals the distance between the last point of p_1 and z. Clearly p_1 is central. So at I as a whole b can more properly be said to be at p_1 than at any position lying outside p_1. To see this more clearly compare what happens at I with what happens at intervals 12:23-12:28 and 12:33-12:38. Positions corresponding to only one of those intervals are less typical of b at I than the central position p_1. Now, in accordance with **SH**, at I b has also positions (partly) overlapping the set of positions at some other intervals. It is clear that b is not to the same extent in all of them, but that the lesser the overlapping between position p and position p_1 the less true it is that b is at p at I. Yet, since the travel is (let us assume) uninterrupted, and there is no cut, probably b to some extent occupies at each subinterval of its travelling time each of the positions in the whole span of its trajectory — but to infinitely many degrees. Even should such a conjecture be downright false, there would still be no cut, no interruption, since the set of positions at I would partly overlap with that of positions at a contiguous interval — but unlike what happens with Priest's instant account, there would be differences of degree.

It seems to me that something along those lines would be more attractive than the jumps which after all Priest countenances, for within his original account, when moving, the body does not gradually decrease its presence at some positions, gradually increasing its presence at other positions, but all of a sudden, at each instant, completely loses certain positions and completely acquires new ones, at all of which it lies, at such an instant, to the same extent.

We can put such a consideration in a different way. It was Spinoza too who wondered whether the body both acquires and loses a position at the same time. Motion can be taken to be just that. At each interval during its travel $\mathfrak{b}$ acquires some position. Yet since it does not remain at those positions, at that time interval it is already leaving them, or at least it is beginning to leave them. So it is not entirely at any position. Still, at the considered interval some of them are being acquired rather than left, and conversely. With truth degrees the story becomes smoother and more plausible.

I am not going to discuss some other sides of Priest's treatment of motion, like his remarks on the instant of change (pp. 200ff) — which could not be maintained without modification if strong negation is posited — or his elaboration on Leibniz's continuity principle (pp. 207ff) or his rejection of symmetry — the set of positions at an instant extending only on one side, to those of the past — which seems to me very Bergsonian — in order not to infringe the principle that what happens until a moment is independent of what happens after-ward, a principle which of course Leibnizians and many other philosophers reject.

To sum up. I think Priest is right when he claims that without contradictions there would be no motion (although he only implicitly uses a principle to the effect that in so far as a body has a position, it does not have other positions). But with just three truth-values a contradictorial account seems to me incredible and committed to a leap view of continuous motion.

---

§7.— **Juridical and deontic logic**

Chapter 13 (pp. 227ff) is given over to discussing legal and moral dilemmas and the contradictions they are supposed to yield.

I think that Priest is right when he claims that there exist such dilemmas and they entail the existence of true contradictions. Yet his arguments could be strengthened and improved upon.

One of the reasons his arguments on this subject are somehow weak is that he seems to share the view of such people as oppose the existence of moral dilemmas, namely that an overridden duty is no duty, or only a *prima facie* duty, or the like. Thus, if one law is of higher rank than another, or later, or can be plausibly interpreted as containing exceptive clauses which accommodate the law with which it clashes, then (pp. 233-4) there is only an apparent conflict. What Priest claims is that there is no guarantee that all apparent legal conflicts can be solved in any of those ways. I agree. There is in fact a lot of evidence that they cannot. Legal disputes will spring to the mind which show to what a point a claim of hierarchical precedence is doubtful in a number of cases.

Yet the most important point is not that, but the fact that even the overridden claim or right is all the same a legal claim or right. True, jurists are so accustomed to reason in the classical terms of all or nothing that they fancy that only a complicated casuistry can decide what claims hold in the end, and which ones do not, the latter being then looked upon as no claims at all. They are wrong. Their wrongness has tremendously serious practical consequences, since it further reinforces the erroneous rule of «all or nothing». In many cases, it is not a question of black and white. This is clear in international disputes, e.g. concerning border demarcations. And many legal paradoxes — which give rise to slippery slope arguments — invite a natural and sensible solution by avoiding hard lines, by choosing fringes, transitions, gradations.

Even if, all in all, country A has a stronger claim to this territory than country B, it does not follow that B's claim is [altogether] null and void. Some compromise may be negotiated which somehow reflects the different (grounded) claims instead of giving all to A and nothing to B. Many legal disputes ought to have sensible solutions by adopting scales of graded allocations — of ownership, or guilt, or whatever — rather than the dry all-or-nothing.

Juridic progress since the 18th century goes in that direction. It used to be the case that almost every infringement of the law entailed maximal penalty — the gallows or the galleys —, whereas our more civilized ways do in effect introduce a notion of degrees of guilt.

The reader has sensed what is the gist of my objection. Again. Ignoring degrees as the source of the true contradictions places us in a very difficult situation. If we want to prove that there are legal conflicts and we share the idea that inferior claims are no claims at all, we need to find cases where it can be shown that no claim is higher or of superior ranking. Although I am sure such cases exist, and are frequent, each of them will be contentious. Playing with where is the burden of the proof is not that interesting. The most important thing to say is that even when in fact one of the claims is overridden, it may be a good, *bona fide* claim, giving the claimant a partial right.

This is more evidently so in the case of moral conflicts. Even if, all in all, a course of action, A, is better and more dutiful than B, it does not follow that we are under no obligation **at all** to do B, that refraining from doing B is completely justified just because B clashes with a greater obligation, that of doing A.

The second reason why Priest's line of argument for moral and legal contradictions seems to me to be in need of an overhaul is that his choice of principles for juridical and deontic logic is not the best one. He adheres to two principles (where '*d*' means 'It is a duty to do', '*p*' means that it is permissible)

(*1*) $\mathit{d}\mathrm{p}\wedge\mathit{d}\mathrm{q}\to\mathit{d}(\mathrm{p}\wedge\mathrm{q})$

(*2*) $\mathrm{p}\to\mathrm{q},\mathit{d}\mathrm{p}\ \vdash\mathit{d}\mathrm{q}$

He rejects:

(*3*) $\mathit{d}\mathrm{p}\to\mathit{p}\mathrm{p}$

And even (at least implicitly) weaken forms, like

(*3'*) $\mathit{d}\mathrm{p}\Rightarrow\mathit{p}\mathrm{p}$

I deem such a choice unfortunate. But I hasten to add that deontic logic is tricky and that such principles among those I now reject which are espoused in Priest's approach once seemed to me right.

The problem with (*1*) (**aggregation**) — which Priest discusses on pp. 238-9 — is that it may be to degree d obligatory to do p, obligatory to degree d' to do q, but not obligatory at all to do both, if doing both is altogether impossible (perhaps not metaphysically impossible in the sense of an abstract possibility but concretely impossible). This is clear in the case of moral dilemmas. I am obliged to rescue Julia, also to rescue Mariana, but not both, which may be utterly impossible. Perhaps Julia’s claim is higher, perhaps both claims are equal. Whatever I do — since I have not had moral luck — I fail an obligation. But, please, do not blame me for failing to rescue both Julia and Mariana. Do not blame me for failing to help hungry people in Sudan, and in Ethiopia, and in Peru, and in Mozambique, and in Tchad, and in Haiti, and… Blame me for each of those failings separately, not for failing to do the impossible.

Blame the famous young man of Sartrean memory for failing his patriotic duty if he gives preference to his mother, or his filial duty otherwise. Do not blame him for this: failing to comply both with his patriotic and his filial duties.

As for (*2*), or the closure rule, it gives rise to paradoxes such as the good Samaritan, which Priest thinks can be solved through scope distinctions. I (now) disagree. Many of those distinctions become unbelievable epicycles, while new counterexamples arise which challenge the principle. Yet if we jettison (*2*) we need something in its place. The main idea is that you are not allowed to do A if doing A prevents someone from fulfilling his duty, or from enjoying his rights. So a causal connection seems to me to be involved here, with a resulting principle of non-hindrance — or something like that.

As for (*3*), or at least (*3’*), imagine that they fail — completely, i.e. that, e.g., Jonathan must do A but he is not permitted at all to do A. That means that Jonathan both is obliged to A and is completely obliged not to A (if, as we commonly think, ⌜*p*p⌝ means ⌜~*d*~p⌝; anyway the principle is discussed by Priest in primitive notation — p. 243).

Priest’s argument against (*3*) is that it multiplies contradictions. Since he adheres to the maxim of avoiding contradictions as far as possible (see below, §9), his line here is parallel to his distrust towards the exclusion principle we have gone into above.

If Jonathan is completely obliged not to A and is obliged to A, he faces two conflicting obligations in such a way that one of them is fully binding. Yet if one of them is fully and entirely binding, the other is not just overridden, but totally overridden, and so no obligation at all. Imagine what it would signify both to **completely** forbid a course of action and yet to make it obligatory.

Moreover, without (*3*) there is no good reason for thinking that moral or legal conflicts entail true contradictions. Priest broaches this subject (pp. 230-1) and answers that there may be other reasons, without (*3*), why legal conflicts give rise to contradictions. E.g. if *x* enjoys legal priority — in virtue of some legal disposition — and yet *y* also enjoys priority — in virtue of a different disposition —, we can infer that *y* does not enjoy priority (since *x* does) and likewise *x* does not enjoy priority (since *y* does). Each of them both has and lacks priority.

But why? Doubtless a principle is being assumed, something like this: ‘If somebody else has priority, you do not have priority’. But what is the rationale? And what is its general formulation — applying to other matters, not just to cross-road driving priority? I can only figure out variations of (*3*). The idea is that, since *x* has a right to cross before *y*, *x* is not entirely bound not to cross before *y* does — i.e. *x* is not wolly bound to defer to *y*’s priority; and so, since *y* has [to some extent] priority, i.e. the right to cross before *x* does, and $\mathscr{x}$ must (up to a point) defer to such a right, $\mathscr{x}$ both has and does not have the duty to give way to $\mathscr{y}$.

---

§8.— Other grounds for true contradictions

Although Priest tends to focus on *a priori* grounds for «dialetheias» — except for motion and moral and legal dilemmas — at one point at least he broaches a different reason, namely complying with only some among a variety of criteria for satisfying a predicate. He offers two illustrations, one about temperature, the other on left vs right in politics. The former seems to me extremely dubious but I do not want to discuss it. As for the latter, the idea is that if being left-wing in politics is complying with conditions c_1, …, c_n, and being right-wing is complying with *not-c* $_1$, and, …, and *not-c* $_n$, then a group $\mathscr{G}$ which for some 1<j<n satisfies c_1, …, c_j, *not-c* $_{j+1}$, …, *not-c* $_n$ will be both left and not left.

I guess that only a few classicists will be convinced by such an argument. They are likely to rejoinder that being left and being right are contraries, not contradictories, and that $\mathscr{G}$ is neither. Call such a reply the «neitherist» ploy.

The trouble with such a reply is that it waives the principle that, in so far as a political group is not right-wing, it is left-wing. If whatever is neither belongs to the «centre», then surely it is going to be very hard to find political groups outside the «centre». In fact, the set of requirements on criteria is open-ended. Very often left-wingers have conservative leanings as regards sexual morality, and sometimes family relations, etc. On economical issues, some right-wingers may fail to support purely private free market economies, while preferring some sort of regulated market, or the like. In fact there are not many people who qualify as either left-wingers or right-wingers on all and every demarcation criterion. Thus, the neitherist ploy, together with the facts of the matter, entails that there is almost no right and no left. And about the same could be said for many a similar classification (rich/poor, pleasant/unpleasant, literate/illiterate, rural/urban, ancient/modern, and so on).

The problem with Priest's treatment is again that it does not allow of degrees. But suppose that $\mathscr{G}_0$ satisfies all criteria on left-wing-ness; $\mathscr{G}_1$ complies with all except two (it, e.g., opposes abortion and is hard on immigration), whereas $\mathscr{G}_2$ fails those two criteria and moreover a 3d one (it does not favour free compulsory education for all, or only with a number of restrictions); $\mathscr{G}_3$ shares all the positions of $\mathscr{G}_2$ except that it advises private ownership of means of production within strict limits; $\mathscr{G}_4$ has the same position as $\mathscr{G}_3$ except that the limits for private ownership are higher and the regulations looser, and so on; at the other end, group $\mathscr{G}_n$ satisfies all conditions for right-wing-ness, but $\mathscr{G}_{n-1}$ favours free immigration. The natural thing to say is that $\mathscr{G}_n$ is wholly right-wing, $\mathscr{G}_{n-1}$ less right-wing than $\mathscr{G}_n$ is, … $\mathscr{G}_2$ less right-wing than $\mathscr{G}_3$ but more so than $\mathscr{G}_1$, … and $\mathscr{G}_0$ completely left-wing. (Of course almost everybody is in between the extremes.)

Moreover, it is not a matter of either entirely satisfying a criterion or else completely failing to satisfy it. There are degrees. But even if the criteria were so crisp that they did not admit of degrees there would be an «all-in-all» consideration. Of course it may happen that in some respects $\mathscr{G}$ is more right-wing than $\mathscr{G}'$; in some other respects $\mathscr{G}$ is less right-wing than $\mathscr{G}'$. (If truth-values are scalar, rather than tensorial, such a possibility is hard to be accounted for.) But in a number of cases some «all-in-all» consideration is in order and plausible. (Otherwise we could say nothing about whether a colleague is a good academic, a student is promising, a radio broadcast interesting, a scientific theory innovative, a software program useful, and so on, except that «in some respects it is, in other it isn't [at all?]». In many such cases an «all things considered» viewpoint is permissible and based on what things are, and what emerges is that, all in all, this colleague is a better academic than that one, this software programme is better than that one, and so on. Not an undifferentiated magma of «good-and-not-good». Almost everybody is good and not good, but doubtless some people are worse than others.)

Doing with just the trichotomic classification Priest offers us is not going to successfully cope with anything of the sort. We'd be bound to regard all groups in between $\mathcal{G}_0$ and $\mathcal{G}_n$ as equally left-and-not-left, none being more right-wing than any other. I do not deny that such a trichotomy is an improvement over the classical all-or-nothing approach. Yet, is the improvement really a great one? I do not think so. In some sense, it still is an all-or-nothing approach: if two groups fail to comply with all conditions for being right-wing and also both fail to comply with all conditions for being left-wing then they are on the same footing, nothing can be said of the one which cannot be said of the other. ('Less… than' is not an expression of which Priest's logical theory takes any notice.)

Adding **grey** to **black** and **white** is fine. But of course it is not enough. There are degrees. Some grey things are blacker than others. And 'less… than' is a matter of logic, since there are logically valid inferences essentially involving comparatives.

---

§9.— The Confinement Policy and Using DS

There are two ways of drawing a line between the contradictions we want to uphold and those we do not want. One of them is to rule that such contradictions as involve strong negation — i.e. **over**contradictions — are beyond the pale, and deserve rejection. The other ones are logically unobjectionable. (But notice that a formula of the form ⌜p∧~p⌝ may be an overcontradiction, even if '~' is natural negation — namely, if ⌜p⌝ is of the form «To some extent, q».) That line is not taken by Priest, since he tries to avoid strong negation — although, as I argued before, he seems to be committed to accepting strong negation, which is definable in his whole system, once Peano arithmetics and semantical predicates are introduced. Thus he opts for the alternative policy, namely to avoid contradictions as far as possible. That is the purpose of his Methodological Maxim (**M**) (p. 145), viz.:

> (**M**) Unless we have specific grounds for believing that the crucial contradictions in a piece of quasi-valid reasoning are dialetheias, we may accept the reasoning.

I do not want to tarry on the details here. The idea is as follows. Suppose from set A of premises a conclusion can be validly deduced to the effect that ⌜p∨.q∧~q⌝, with some constraints being imposed which avoid that irrelevant contradictions ⌜q∧~q⌝ creep in (p. 150). When such constraints are complied with, the contradictions involved are crucial. Then the quasi-valid reasoning allows us to draw conclusion ⌜p⌝ forthright from set A. And the rationale is that more often than not true contradictions fail to arise. As Priest puts it (p. 144):

> The reason is a simple one: the statistical frequency of dialetheias in normal discourse is low. Dialetheias appear to occur in a quite limited number of domains: certain logico-mathematical contexts, certain legal and dialectical contexts, and may be a few others.

My first comment on such an argument is that, if the kind of examples of true contradictions which were provided above, in §8, is right, then true contradictions pop up in all domains. Not «a few others». No domain exists in which true contradictions do not arise.

My second comment is that, if contradictions are admitted in the field in which they were supposed to be least likely to arise and most damaging — mathematics — there is no good reason for us to be coy about them in other domains. Of course, it just might be the case that contradictions were true only where they were expected the least to arise. Some highly improbable things happen sometimes. Yet very often when situations of some kind emerge even where they were the least expected, they are likely to occur elsewhere, in many other domains. E.g. actual infinities were supposed to be ruled out in virtue of Euclides' principle that the whole is greater than any part thereof. Once they arose in mathematics, with the calculus, they had to be admitted in all domains — although nowadays some people think that quantum mechanics has again dislodged them from physics, it remains to be seen how long people are going think that.

My 3d comment is that, if contradictions are not bad — in general — there is no reason to be afraid of them. If they arise in the *sanctum sanctorum*, mathematics, surely it is not irrational to have a contradictory belief. Then when people say that it rains and it does not rain, that the man yonder is and is not bald, that this paper is and is not white, that such a course of action is and is not dangerous, etc., why are we bound to construe what they say in devious ways or to scorn them for their purported irrationality?

My 4th comment is that empirical evidence shows an enormous amount of utterances which are literally contradictory. Philosophers used to allege that, duly (charitably) construed, they were not. But such an exegetical approach was enforced by the view of contradictions as horrible, awful, irrational, utterly rejectable. If there are true contradictions — and in mathematics! — surely it is not irrational to have contradictory beliefs. Thus no need for charity here. There may be special and cogent reasons why some apparently contradictory utterances are, upon consideration, taken not to be genuinely contradictory. But the huge amount of literal contradictions people utter in everyday communication seems to render implausible the idea that all or most of them deserve to be paraphrased away in a charitable manner. Priest advises us — ibid — to 'consider a random sample of the assertions [we have] met in the last few days and see what percentage might reasonably be thought to be dialetheic'. My own — fallible — assessment yields a high percentage — esp. on radio interviews, very often when academics have to answer questions. (Ill-formulated questions, which debar the interviewed person from a simple 'Yes' or 'No' answer? Maybe. Yet, …)

In addition to the just considered argument that contradictions are infrequent, Priest offers a second argument (p. 144-5) to the effect that they are unlikely and so (**M**) is broadly reliable: the sheer fact that people reason using **DS**. If quasi-valid reasonings were widely unreliable, a lot of damage would surely follow.

Nevertheless, the argument assumes that uses of **DS** involve natural or weak negation rather than strong negation. Now, in spoken language, 'not' — with perhaps a prosodic or suprasegmental signal, which stands for the modifier 'at all' — may also express strong negation. It depends on the context. Written language lacks prosodic resources — only a pale reflection of them is available, but with nothing like the richness of speech. Perhaps this is one of the reasons which have led to overlooking the difference between strong and natural negation, and so that between overcontradictions and simple contradictions.

§10.— **Conclusion**

The reader has rightly realized that, despite my objections, I share most of G. Priest's views and many of his arguments — with mitigations. It is very probable that the reader disagrees with us both, finding all contradictions inadmissible. What cannot be said, though, is that the subjects G. Priest studies in his book are of no importance for contemporary philosophy. In fact, I cannot see any subject more important than the question of whether or not there are true contradictions. Priest's book can be ignored by no one. Fervent adherents of Aristotelian logic are invited to take it seriously, and to discuss it rationally — not to rend their garments.