The Theory of Inconsistency

Inconsistent Mathematics and Paraconsistent Logic

C.E.Mortensen

Volume 3

CONTENTS OF VOLUME 3

Part 4: Papers on the Theory of Inconsistency

Six papers addressing philosophical problems of the theory of inconsistency.

Part 5: Papers on the Philosophy of Mathematics and Physics

Ten papers addressing philosophical questions on the reality of mathematical constructions and the interpretation of various theories in physics.

(40) “Change”, *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu, 18.12.2002. (M2002c)

(41) -- and Julianna Csavas, “In the Beginning”, *Erkenntnis*, 59 (2003), 141-156. (M2003c)

(42) "Dharmakirti and Priest on Change", *Philosophy East and West* Vol 54, No 1 (2004), 20-28 (M2004)

Part 6: Computer Studies

Three computer studies

(43) -- & Steve Leishman, *Computing Dual Paraconsistent and Intuitionist Logics, Technical Reports in Automated Reasoning,* Australian National University, TR-ARP-9/89 (1989), 70pp. (M1989c)

(44) -- & Steve Leishman, “Inconsistent Control Systems”, The University of Adelaide, 1997 (M1997d)

(45) --, Peter Quigley, Steve Leishman and John Mercier *Inconsistent Images*, University of Adelaide (2003-5, website), *http://www.arts.adelaide.edu.au/humanities/philosophy/inconsistent_images/index.html*
(M2005b)

PART 4

Papers on The Theory of Inconsistency

Notre Dame Journal of Formal Logic
Volume 24, Number 1, January 1983

The Validity of Disjunctive Syllogism Is Not So Easily Proved

CHRIS MORTENSEN*

1 This note is prompted by John Burgess's "Relevance: A Fallacy?" [2], which offers an argument in favour of the deductive validity of the argument form Disjunctive Syllogism, *DS* (*A*,not-*A* or *B*/∴*B*). The kind of argument he gives is not so unusual, and can be encountered around the literature (e.g., [3], p. 666) and not infrequently in the verbal pronouncements of philosophers. The bones of the reply I will give to Burgess can also be found in a number of places and as long ago as 1972 (e.g., [4]-[6]), though I do not think it has been systematically developed anywhere. Since Burgess's argument is representative of a widespread kind of mistake about relevant logics, it is worthwhile to try to say clearly what is wrong with it.

Burgess disclaims any attempt to discuss the extensive literature on relevant logics other than Anderson and Belnap's 1975 "masterwork" *Entailment*. Hence, his argument is best viewed as a piece of internal criticism of that book. However, he makes several remarks which imply fairly clearly that his sights are on more general targets, and are intended to apply to all "self-styled" relevant logicians. Let me therefore concede straight away that in my view, Anderson and Belnap's discussion of *DS* in *Entailment* is inadequate. It would be rash, however, to draw the conclusion that there is no hope offered within the broad programme loosely classifiable as "relevantist" for shoring up their rejection of *DS*. Indeed, in view of the well-known Lewis proofs of the irrelevant principle of *Ex Falso Quodlibet*, there had better be.

Burgess says that the issue as far as he is concerned is whether relevant logics "are in better agreement with common sense than classical logic"; and

*I wish to thank Bob Meyer, Graham Priest, and Stephen Read for helpful comments.

Received September 25, 1981; revised February 5, 1982

not whether some other intuitively comprehensible modeling for the relevant logics can be found elsewhere. So be it. Beyond the mild caveat that common sense might be strained somewhat by our investigation, let me agree with this demarcation of the area of contention. It is worth noting, though, that while Anderson and Belnap obviously thought their logics were in better agreement with common sense logic (or natural logic, or natural language, or some such thing), that might not be the only reason that could be advanced for adopting a relevant rather than a classical logic. Instead, one might appeal to pragmatic criteria such as overall simplicity of the foundations of mathematics or science as grounds for a *reconstruction* of natural logic along relevant lines.

One final piece of clarification. In all the standard relevant logics, a distinction is made between extensional, truth functional 'or', 'v', and intensional 'or', '+'. The extensional form of *DS* (A, $\sim A \vee B/\therefore B$), is not generally valid in these logics (even in their purely truth functional, "zero degree" fragments), whereas the intensional form (A, $\sim A + B/\therefore B$) is. Burgess wants to show that certain valid natural language examples of *DS* have to be understood as of the extensional kind, so that extensional *DS* must be deductively valid. He represents the relevance position as having to hold that valid natural language examples of *DS* involve appeal to the (valid) intensional form of *DS*, either by virtue of direct translation of 'or' into '+' or because in such cases the crucial premiss using '+' is always available. I think that questions about intensional disjunction cloud the issue here, something for which Anderson and Belnap are at least partly to blame. I will be concerned to show where Burgess goes astray by agreeing that his examples use 'or' extensionally but arguing that they do not show that extensional *DS* is deductively valid; so I want to set aside questions about '+' altogether. Throughout this paper, then, 'or' is taken as 'v'.

2 I begin by offering an explication of the useful intuitive idea of a *deductive situation*. Human beings are often in the position of deducing sentences from other sentences. Disputes as to the validity of a deduction from certain premisses can, I propose, be thought of as disputes as to the exact nature of the deductive situation containing those premisses. To make this more precise, we can introduce the idea of an L-theory (relative to a logic L). Consider a language $\underline{L}$ closed under conjunctions ($\wedge$), disjunctions ($\vee$), implications or entailments ($\rightarrow$), and negations ($\sim$). A *logic* in $\underline{L}$ is a subset L of $\underline{L}$ closed under the rule of uniform substitution. Now we can define the notion of an *L-theory relative to a logic L* (e.g., a *PC*-theory, or an *S5*-theory, or an *E*-theory). An L-theory is a set of sentences containing all the consequences of all the members of the theory which the logic L says are consequences. More formally, a subset X of $\underline{L}$ is an *L-theory* iff: (1) if $A \in X$ and $\vdash_L A \rightarrow B$ then $B \in X$, and (2) if $A \in X$ and $B \in X$ then $A \wedge B \in X$. (The second of these requirements is reasonable, but does not in general follow from the first, so needs independent specification.) I propose to explicate the idea of a deductive situation by identifying it with that of an L-theory where L is "natural" or "common sense" or "correct" logic.

We will say that an L-theory X is a *DS-theory* (written $DS(X)$) iff if $A \in X$ and $\sim A \vee B \in X$ then $B \in X$. Suppose, as Burgess believes, that *DS* is a universally valid principle of common sense or natural logic. Then it must be that every deductive situation is a *DS*-theory: whichever natural logic L is, it must always be that B is in every deductive situation which contains both A and $\sim A \vee B$. Furthermore, if Burgess is wrong and Anderson and Belnap are right and DS is not universally valid, then for some A, B, it must be that B fails to be in some deductive situation containing A and $\sim A \vee B$. (Any reasonable completeness theorem for L will deliver that result.) And that is what any putative counterexample to DS must evidently achieve: to produce a deductive situation in which A and $\sim A \vee B$ hold (belong) but B does not.

3 In the terminology of this paper, Burgess's strategy is to produce examples of deductive situations, claim that the examples are not special in any way which vitiates the generality of his argument, and claim that the situations are DS(-theories). This is to show that the validity of DS is required for common sense thinking. Now it is important for such a strategy that the deductive situation be correctly identified, for it might be that *further* information about the deductive situation is covertly imported which is sufficient to ensure that the deductive situation is DS. We would then be dealing, in effect, with a larger deductive situation and so the demonstration that B holds in it does nothing to show that the universal validity of DS is what is solely responsible for B's holding. It is this error which I claim Burgess has made.

The position I propose is that although DS is not universally valid, it is an acceptable mode of reasoning under certain circumstances. The situation seems to be like this. Many relevance people feel suspicious of DS because it seems to break down in what might be called "abnormal" deductive situations, particularly inconsistent situations. It is not infrequently claimed by relevance logicians that theories such as naïve set theory, classical pre-Cauchy calculus, the Böhr theory of the atom, quantum theory, natural language with its own truth predicate, and Peano arithmetic are or might well be non-DS. If this claim is correct, then some logic for which DS fails is a better model of natural logic than classical logic is. On the other hand, DS does seem to be a natural mode of inference in 'normal' deductive situations, the kind encountered every day. These two intuitions about DS can be reconciled if we can give an account of deductive validity according to which DS holds only in normal situations, and that is what I claim.

Some more definitions. An L-theory X is *consistent* iff for no A are both A and $\sim A$ in X. X is *trivial* iff X is the whole language $\underline{L}$. X is *nonprime with respect to* $A \vee B$ iff $A \vee B \in X$ but $A \notin X$ and $B \notin X$. X is *nonprime* iff X is nonprime with respect to some disjunction. X is *prime* iff it is not nonprime. The failure of primeness is no mystery, even for truth-functional disjunction. Consider Peano arithmetic formulated with a base of classical logic (i.e., classical Peano arithmetic, PA). Let G be its Gödel sentence. Then certainly $\vdash_{PA} G \vee \sim G$. But, by Gödel's first Incompleteness Theorem, if PA is consistent, neither $\vdash_{PA} G$ nor $\vdash_{PA} \sim G$. Hence if PA is consistent it is nonprime (with respect to $G \vee \sim G$).

Now some things are known about conditions under which L-theories are, and fail to be DS:

(1) Any inconsistent but nontrivial L-theory fails to be DS (under very weak assumptions about the logic L). Reason: If X is inconsistent, then for

some A, $A \in X$ and $\sim A \in X$. Since $\sim A \in X$, if $\vdash_L \sim A \rightarrow (\sim A \vee B)$, we have $\sim A \vee B \in X$, for arbitrary B. If *DS* held for X, we could deduce that $B \in X$, for arbitrary B, i.e., X would be trivial. Ex hypothesi, X is not trivial, so *DS* fails for X. In particular, all the usual relevant logics have inconsistent and nontrivial theories.

(2) If X is nonprime with respect to any disjunction $\sim A \vee B$ while $A \in X$, then *DS* fails for X. Reason: While $A \in X$, and $\sim A \vee B \in X$, the failure of primeness ensures that $B \notin X$.

(3) If, on the other hand, X is consistent and prime, then *DS* holds for X. Reason: Let $A \in X$, $\sim A \vee B \in X$. Since X is prime, at least one of $\sim A \in X$ and $B \in X$. But X is consistent, so $\sim A \notin X$. Hence $B \in X$.

(4) For certain choices of logic L, such as classical logic and intuitionism, *DS* holds for all L-theories.

A point to note about the proof under (3) that consistency and primeness implies *DS* is that it seems to appeal to a metalinguistic principle of *DS* as it were. However, it is not being claimed that *DS* is never legitimate. On the contrary, in normal well-behaved situations *DS* is to be expected to hold, and there does not seem to be anything untoward about the metalinguistic situation here. For example, we might formalize the metatheory and prove it to be consistent and prime. The foregoing considerations, then, enable us to conclude that a necessary and sufficient condition for a nontrivial deductive situation to be *DS* is that it be consistent and, for all subsets $\{A, \sim A \vee B\}$, prime with respect to $\sim A \vee B$. A special case sufficient for *DS* is where the deductive situation is consistent and complete (as Kripke's possible worlds are), since it is easy to show that, given De Morgan's Laws, consistency and completeness imply primeness.

4 This preamble enables me to make my main point against Burgess. If the deductive situations he describes patently contain *extra* information sufficient to guarantee that *DS* holds of them no matter what logic L is involved—such as the information that they are consistent and prime—then his argument *cannot* show that the (universal) validity of *DS* is required by those situations. I claim that this is what has happened. To see this, let us look at the examples Burgess gives. I simplify drastically for brevity.

In the first example, we are presented with a deck of cards and the information that a certain card in question is not both card A and card B, and that it is card A. Burgess claims that it is legitimate to conclude that it is not card B. Clearly this can be recast as an example of *DS*. But, unfortunately for relevant logic:

> Had Wyberg been a relevantist, unwilling to make a deductive step not licenced by the Anderson-Belnap systems E and R, he would have been unable to eliminate the queen of clubs from his calculations, and would have lost the game. A relevantist would fare badly in this game and others, and in game-like situations in social life, diplomacy, and other areas—unless, of course, he betrayed in practice the relevantistic principles he espoused in theory. ([2], p. 100)

However, now we are in a position to see that the deductive situation as Burgess presents it—a deck of ordinary playing cards, not card A or not card B, etc.—is *certainly consistent and prime.* It would be quite absurd to say that the situation is one where we have both card A and not card A. Equally, if it is either not card A or not card B, then at least one of those options obtains. So of course *DS* may be legitimately used. Relevant logicians are still worth employing as wargamers. The assumption of consistency and primeness here is so obvious as to be invisible. That is why it must be regarded with suspicion as possibly an operative factor in the situation. If it is, then nothing follows about the validity of *DS*.

In the second example, we are presented with a hypothetical discovery in number theory of the form $(n)(A(n) \vee B(n))$, and invited to conclude from a proof of $\sim A(1)$, that $B(1)$. (Again, to recast as an instance of *DS*, instantiate and use double negation.) What could be more harmless? Quite a bit: the assumption of consistency and primeness is, again, present. But here it is, instructively, much less obviously true. Suppose that number theory is inconsistent, and in particular that $A(1) \vee B(1)$ holds because $A(1)$ holds. Do we really want to conclude, if we come into possession of a proof of $\sim A(1)$, that $B(1)$? Of course, if it is *already* believed that classical logic is true, so that *DS* holds of number theory, then we will be prepared to conclude that $B(1)$, by the principle that everything can be deduced from a contradiction. But that begs the question. Again, suppose that arithmetic failed to be prime at $A(1) \vee B(1)$. Then from a proof of $\sim A(1)$ it would be quite illegitimate to conclude $B(1)$. But this is not what Burgess is supposing to be the case; in fact he quite explicitly supposes that a proof of $B(1)$ exists. The extra information Burgess needs to make his case for *DS* work is clearly present. But the presence of the extra information destroys his case.

One final quick example Burgess gives is that of someone once told that A or B but cannot remember which. Finally, he establishes $\sim A$, and so concludes that B.

> Such examples . . . show that, as far as negation, conjunction and disjunction are concerned, 'classical' logic . . . is far closer to common sense and accepted mathematical practice than is the 'relevant' logic of Anderson and Belnap. ([2], p. 102)

Certainly here the presupposition of consistency and primeness is less obvious. But it is there all the same, I submit, in virtue of the "presupposition of normality". Consistency and primeness are normal, nice, well-behaved. People are not ordinarily confronted with inconsistent or nonprime situations, so find moves like *DS* natural to make. If the situation were abnormal, say a mathematical one where primeness were in doubt, then it would be a more dubious move to deduce B. Thus, again, the kind of argument one often hears informally: '*DS* must hold. Look, if I know that today is Monday or Tuesday, and I know that it isn't Monday, I must conclude that it is Tuesday'. But what more normal a deductive situation can one imagine?

A final point against Burgess. He accuses the relevance programme of confusing logical implication with reasoning or inferring, a distinction of Harman's. I claim, to the contrary, that he is guilty of precisely that confusion.

In those terms, the issue is whether *DS* captures a logical implication, with Anderson and Belnap denying it and Burgess claiming it. In the light of the previous discussion, however, the examples he provides are nothing that a relevant logician need deny to be *useful reasoning or inferring.* It is quite proper to import extra facts about the deductive situation in order to extract all the useful information out of it. In thinking that he has raised a difficulty for the relevantist position, Burgess shows that he has precisely not appreciated the difference between usefully reasoning and universally valid deduction.

5 A consequence of the position of this paper is that the claim of the relevance programme, that *DS* is not universally valid, entails the claim that not all nontrivial deductive situations are consistent and prime. In order to show that to be incorrect, one must plainly adopt a different strategy from Burgess's. What must be considered, instead, are putative examples of nontrivial inconsistent deductive situations. Clearly these will be decidedly of the unusual type. But if a rule such as *DS* is to be valid, then it needs to hold in all deductive situations, not just normal ones. It is precisely the relevantist claim that abnormal, unusual situations where *DS* fails need to be taken into account. It is pointless to dispute this by concentrating on conditions in which it is known that *DS* holds.

REFERENCES

[1] Anderson, A. R. and N. D. Belnap, Jr., *Entailment: The Logic of Relevance and Necessity,* Princeton University Press, Princeton, New Jersey, 1975.

[2] Burgess, J. P., "Relevance: a fallacy?," *Notre Dame Journal of Formal Logic,* vol. 22 (1981), pp. 97-104.

[3] Hanson, W. H., "First degree entailments and information," *Notre Dame Journal of Formal Logic,* vol. 21 (1980), pp. 659-669.

[4] Routley, R. and V. Routley, "Semantics of first degree entailment," *Nous,* vol. 6 (1972), pp. 335-359.

[5] Routley, R., "Ultralogic as universal," *Relevance Logic Newsletter,* vol. 1 (1977), pp. 50-90, 138-175.

[6] Routley, R., "The choice of logical foundations: non-classical choices and the ultralogical choice," *Studia Logica,* vol. 39 (1980), pp. 77-98.

Department of Philosophy
Research School of Social Sciences
Australian National University
P.O. Box 4
Canberra ACT 2600
Australia

Notre Dame Journal of Formal Logic
Volume 27, Number 2, April 1986

Reply to Burgess and to Read

CHRIS MORTENSEN

1 Introduction Either John is foaming at the mouth or John is biting the carpet. John is not foaming at the mouth. Therefore, John is biting the carpet. Such an instance of Disjunctive Syllogism (*DS*) is undoubtedly intuitive, but a form of inference which is intuitive is not thereby valid. There are (at least) three positions which can be taken concerning the validity of *DS*. *First: DS* is valid, and the "or" in it is the two-valued extensional "or". Thus, the argument form Extensional Disjunctive Syllogism (*EDS*), *i.e.*, $A \vee B$, $\sim A / \therefore B$, is valid. *Second: EDS* is invalid. There is a valid argument form, Intensional Disjunctive Syllogism (*IDS*), namely $A + B$, $\sim A / \therefore B$, where "+" is intensional disjunction. Whenever you have a valid example of *DS*, it is because it is an instance of *IDS*. *Third:* The examples of *DS* which seem intuitive are often instances of *EDS*; but this does not make *EDS* valid, and it is not. Whenever it seems intuitive to infer using *EDS*, it is because there is an extra assumption, that things are "normal", which ensures the truth of the conclusion and which explains the apparent intuitiveness of *EDS*.

Recently (in [8]), I defended the third of these. Read (in [9]) defended the second. In the course of my argument, I made the further claim that there are precise sufficient conditions for when the truth of the premises of *EDS* would ensure the truth of the conclusion and that these conditions obtained whenever there was an intuitive example of *EDS*. Both Read and Burgess ([4], see also his [5] and [6]) understood me to be trying to prove my claim by appeal to the validity of *EDS* in the metatheory, an appeal which they took to be circular. In Section 2 of this note, I will argue that there is no circularity in my position. In Section 3, I will argue that my position is a stable one, in that no collapse into a generally valid *EDS* follows from it. In Section 4, I will briefly respond to some of Burgess's other points from [4].

2 The appeal to normality We need some definitions. A theory for a logic L is a set of sentences closed under the consequence relation $\vdash_L$. It is useful to consider the situation we find ourselves in when deducing according to "natu-

Received November 16, 1984

ral" logic, as a theory closed under the natural consequence relation $\vdash$. This has the virtue that actual deductive behavior in the course of theory construction in the sciences can be seen as data about how propositions are related by $\vdash$. Closely connected with the question of whether $A \vdash B$, is the question of whether B belongs to all natural theories to which A belongs. Read objects to my too-quick identification of these two questions; but I do not rely on it in what follows.

A theory will be said to be consistent iff for every sentence A, not both A is in the theory and $\sim A$ is in the theory (or—what is equivalent given metalinguistic laws of De Morgan, Double Negation and Commutation hereafter assumed—either A is not in the theory or (extensionally) $\sim A$ is not in the theory). A theory will be said to be prime iff for every extensional disjunction $A \vee B$ in the theory, either A is in the theory or (extensionally) B is in the theory.

I claimed that theories which are *intuitively well behaved or normal* are closed under *EDS*, and that counterexamples to *EDS* are to be found in abnormal theories only, though that should hardly daunt the fearless logician. I then claimed that a sufficient condition for a theory to be closed under *EDS* is that it be consistent and prime. In proving this, I appealed to something looking like *EDS* in the metatheory. Both Read and Burgess objected that I had no right to such an appeal.

But this is not so. Let me make clear what my contention is. I claim that (given a normal metatheory which we should be able to ensure), for any consistent prime theory *Th* and for any propositions A, B, from $A \vee B \in Th$ and $\sim A \in Th$ it is deducible that $B \in Th$. My argument is in two stages. The first stage is in the metametatheory.[1]

> From the premises that a theory is normal and that $A \vee B$ and $\sim A$ are in the theory, it is deducible that B is in the theory.
>
> The metatheory (of this paper) is normal.
>
> ∴ The metatheory is such that it is deducible that B is in it from the premise that $A \vee B$ and $\sim A$ are in it.

The premises of this argument were not justified by any appeal to *EDS*, but to the pretheoretic data available to us. There do seem to be intuitive examples of *EDS*, and the metatheory needed to put through the argument to follow is minimal: first-order logic with a single binary relation $\in$ with quite weak properties and a relation $\vdash$ which is also quite weak. No reason for suspicion of abnormality or paradox here. That is, I don't claim to *prove* the truth of the above two premises. Proof will have to stop somewhere, especially in the epistemology of logic. I offer support for them, of a reasonable kind.

Now for the second stage of my argument, "drop down a level" to the metatheory. For suppose that *Th* is consistent and prime; I claim that from this fact together with $A \vee B \in Th$ and $\sim A \in Th$ it is deducible that $B \in Th$. For from $\sim A \in Th$ (*i.e.*, not—not— $\sim A \in Th$) together with the consistency of *Th* (either not $\sim A \in Th$ or not $A \in Th$) it is deducible that not $A \in Th$ (by appeal to the conclusion of the Stage I argument). Then from $A \vee B \in Th$ by primeness, either $A \in Th$ or $B \in Th$; hence it is deducible that $B \in Th$ (by appeal to

Stage I again). It is apparent that no question-begging appeal to the validity of *EDS* has been made here. A version of *EDS* has been used, but *only as a property of a particular theory*, and support has been given for that.

3 Formalizing the argument It is evident that a formal version of the foregoing argument can be written down in a straightforward way, taking a single binary predicate $\in$, the usual extensional connectives, and, wherever "it is deducible that" occurs, $\vdash$. Or, instead of $\vdash$, use a metatheoretic $\rightarrow$. For the conclusion $(Con(Th)\ \&\ Pr(Th)\ \&\ (A \vee B) \in Th\ \&\ {\sim}A \in Th) \rightarrow (B \in Th)$, the extra properties needed for $\rightarrow$ are substitutivity with respect to the equivalences of De Morgan, Double Negation and Commutation, and the two rules (a) Transitivity for $\rightarrow$, and (b) $A \rightarrow B,\ C \rightarrow D /\therefore A\ \&\ C \rightarrow B\ \&\ D$. A special case is where we take Th = The True, so that "$\in Th$" is a truth predicate. I presume that this constitutes an answer to Burgess' demand ([4], pp. 49, 51) for a theorem formalizing the principle (Consistency & Primality & $(A \vee B)$ & ${\sim}A) \rightarrow B$. Of course this is not to say that the $\rightarrow$ in question is entailment, since, for example, an enthymematic $\rightarrow$ would do (e.g., [1], p. 259, or [2]). On the other hand, even if we take the $\rightarrow$ of the metatheory to be entailment, *it does not follow that all theories are classical.* To see this, just note that the logical structure of the object language theories has been left unspecified. Nothing prevents them, therefore, from being theories of any of the usual relevant logics. The first stage of the argument delivers the conclusion that B may be deduced from $A \vee B$ and ${\sim}A$ only for instances of A and B *from the particular metatheory.* Nothing follows about *unrestricted* theoremhood of $((A \vee B)\ \&\ {\sim}A) \rightarrow B$, so it is open to us to invest the $\rightarrow$ with unrestricted substitution instances corresponding to a weaker logic than classical.

This is far from eclecticism. As defined by Burgess, that is the view that relevant logic is only "appropriate for certain extraordinary abnormal situations . . . no logic provides canons of validity that are necessary and sufficient for all situations . . . logics have to be local, . . . different situations have different logics" ([4], p. 50). If this means that there are no logical truths and no valid arguments,[2] then I am certainly not committed to it. The view advocated here is consistent with the position that there are some universally valid argument forms, and some argument forms which in more restricted circumstances take us from truths to truths. It would be confusion to describe this as the thesis that relevant logic is *only* appropriate in abnormal situations. One might hold instead that relevant logic describes the correct universal validities, while classical logic is a special case, holding only over a restricted domain.

I take it that the fact that nonclassical object-language theories are describable by weak metatheory (and any supertheory) in the fashion of this paper demonstrates the logical stability of my position. So one is led to ask what kinds of epistemic considerations Read and Burgess would severally appeal to in support of their own differing positions. I suspect that Read's view brings him dangerously close to logical skepticism.[3] He seems to think that unless some kind of proof of the unrestricted validity of an argument form is forthcoming, then one would never be justified in moving from its premises to its conclusion in a particular case. But if any argument form is valid, then some inference rules are not justified by being proved from others. Burgess, on the other hand, might be making a much stronger demand than I attributed to him at the beginning of this section, namely, the demand to produce a fully developed relevant metalogic, truth theory, model theory, set theory, the lot ([4], p. 51). This ploy is sufficiently common to deserve a name, so let us call it *The Fallacy of the Conservative Theorist*: Unless My Opponents Have a Fully Developed Countertheory, All Their Arguments Against Me are Unsound. But, of course, the above result holds in any supertheory of our metatheory, no matter how much extra baggage it gets.

4 Sundry loose ends This brings me to the question of who has misrepresented whom. I have already argued that Burgess and Read have misunderstood me. Burgess claims that I misrepresented him, and that his intent was all innocence itself: only to show against Anderson and Belnap that common sense employs *EDS*. As I said in [8], Burgess's first paper is best understood as an attack on Anderson and Belnap, but some remarks suggest that his aims are more general. I do not think that anyone could read his paper and not get that impression. Here are just a few points. The aim of the card game example was not just to show that common sense goes his way, but also that "the relevantist" would do "badly" and "in social life, diplomacy, and other areas". I deny this. Notice, too, the inference from "not common sense" to "bad". Again, his arithmetical example insinuates the less-than-innocent conclusion that "the honor of priority goes to Wyberg", the implication being that Wyberg's argument was valid ([6], p. 102). I claim that if arithmetic is inconsistent, then Wyberg's argument is invalid, so the "commonsense" presumption that Wyberg's argument is valid masks the presupposition of consistency. I take it that in disputing the implication of validity, I was meeting Burgess's "challenge" "to explain away some apparent examples of commonsense instances of *DS*" ([4], p. 45). Needless to say, to fail to take up such a challenge is to lose some presumed competition by default. The debate might at this point degenerate into semantic trivialities about how narrow in application were Burgess's phrases like "the relevantist" and "the Anderson-Belnap systems *E* and *R*", and he hastens to tell us how big is the gulf between Anderson and Belnap, and Routley ([4], p. 45). I think, in fact, that Burgess's first paper was written largely in ignorance of what had been published *about E and R* by others, an impression his second paper certainly hastens to counteract. Again, his assurance ([6], p. 104) that he was concerned "solely with the original Anderson-Belnap account of 'relevant' logic and with their claim that their systems *E*, *R*, etc., are in better agreement with common sense than is classical logic", would have helped him better had he not *contrasted* it with "the discovery of serendipitous applications" such as logics of ambiguity; instead of a contrast with, say, Meyer's work.

Burgess uses so many rhetorical devices that his papers read like a list of textbook examples of informal fallacies. I do not propose to catalog all of these, but let me caution readers against fallaciously reasoning on the basis of Burgess' second paper according to the *Fallacy of Divide-and-Conquer*: The Opponents are in Disagreement about Some Issues, Therefore All Their Theses are False. Certainly there is disagreement on some issues, but it is simply distortion to say that "Routleyism and Andersonianobelnapism are so dissimilar that it is misleading to apply a single label 'relevantism' to both" ([4], p. 45). Routley, who describes himself as a relevantist, ([10], pp. 21–23), is hardly a *mere* paraconsistentist: it is not possible to read Routley's published work since [13] without grasping the quite central role played in it by relevance (e.g., [12]). Far from there being huge dissimilarity, the point of the Fine-Meyer-Plumwood-Routley-Urquhart semantics was that it offered an *explanatory* account of the Anderson-Belnap systems, and particularly the prized property of relevance (e.g., [13], or [11], p. 394). For that matter, the generality of that semantics, particularly the move to inconsistent and nonprime or incomplete theories, offered an explanatory account of the intuitions, and the limitations of those intuitions, behind relevantism, classicalism, paraconsistentism, intuitionism, connectivism, and modal logic. The particular application here, that the semantics made it clearer what were the options in dealing with *DS* and that one might propose a semantically based explanation of the illusory intuitiveness of *EDS*, seems to me to represent considerable progress over the original Anderson-Belnap account of *DS*.

The simple point against both Burgess and Read is this. Logic does not operate in a vacuum, but on deductive theories. While all the theories of a logic need to be closed under the deducibility relation of the logic, it is possible for some theories of the logic to be closed under additional rules as well, for instance *EDS*; and it would be surprising if we could not sometimes know this and exploit it. My further point against Burgess still stands: that the conditions under which *EDS* holds might be so normal that there is produced the illusion, even in intelligent and expert deducers, that it is valid. This is not to be disposed of by the methods of medieval Christianity invoked by Burgess in the opening quotation of [4].

NOTES

1. I do not rely on a rigid distinction between object language and metalanguage, which is one of the less satisfactory aspects of the classical paradigm. The distinction is used here only for expository purposes.

2. In point of fact this is a view with which I have recently become more sympathetic [7]. But Burgess was in no position to conclude this on the basis of my paper.

3. Cf. also Belnap and Dunn [3]. In the spirit of Belnap and Dunn, we might object: "But what if your metatheory is abnormal?" But what if? That does not count against the claim that *if* a theory *is* normal, *EDS* holds for it. And I take it that it is not so hard to believe that the present metatheory is normal.

REFERENCES

[1] Anderson, A. R. and N. D. Belnap, Jr., *Entailment*, Princeton University Press, Princeton, 1975.

[2] Anderson, A. R. and N. D. Belnap, "Enthymemes," *The Journal of Philosophy*, vol. 58 (1961), pp. 713–723.

[3] Belnap, N. D. and J. M. Dunn, "Entailment and the disjunctive syllogism," pp. 337–366 in *Contemporary Philosophy: a New Survey*, Nijhoff, The Hague, 1981.

[4] Burgess, J. P., "Common sense and relevance," *Notre Dame Journal of Formal Logic*, vol. 24 (1983), pp. 41–53.

[5] Burgess, J. P., "Read on relevance: A rejoinder," *Notre Dame Journal of Formal Logic*, vol. 25 (1984), pp. 217–223.

[6] Burgess, J. P., "Relevance: A fallacy?," *Notre Dame Journal of Formal Logic*, vol. 22 (1981), pp. 97–104.

[7] Mortensen, C. E., "Anything is possible," to appear.

[8] Mortensen, C., "The validity of disjunctive syllogism is not so easily proved," *Notre Dame Journal of Formal Logic*, vol. 24 (1983). pp. 35–40.

[9] Read, S., "Burgess on relevance: A fallacy indeed," *Notre Dame Journal of Formal Logic*, vol. 24 (1983). pp. 473–481.

[10] Routley, R., *Relevantism and the Problem As To When Material Detachment and the Disjunctive Syllogism Argument Can Be Correctly Used*, Research Papers in Logic No. 12, Logic Group, Research School of Social Sciences, Australian National University, 1983.

[11] Routley, R., "Semantics for connective logic I," *Studia Logica*, vol. 37 (1978), pp. 393–412.

[12] Routley, R. et al., *Relevant Logics and Their Rivals 1*, Ridgeview, Atascadero, California, 1982.

[13] Routley, R. and V. Routley, "The semantics of first degree entailment," *Nôus*, vol. 6 (1972), pp. 335–359.

Department of Philosophy
The University of Adelaide
Adelaide, South Australia 5001

CHRIS MORTENSEN

ANYTHING IS POSSIBLE

ABSTRACT. This paper criticises necessitarianism, the thesis that there is at least one necessary truth; and defends possibilism, the thesis that all propositions are contingent, or that anything is possible. The second section maintains that no good conventionalist account of necessity is available, while the third section criticises model theoretic necessitarianism. The fourth section sketches some recent technical work on nonclassical logic, with the aim of weakening necessitarian intuitions and strengthening possibilist intuitions. The fifth section considers several *a prioristic* attempts at demonstrating that there is at least one necessary proposition and finds them inadequate. The final section emphasises the epistemic aspect of possibilism.

1. POSSIBILISM

I begin with the thesis of *possibilism*, by which I mean the group of theses that all propositions are possible, or possibly true, that all propositions are contingent, that no proposition is necessary. The denial of the latter is the thesis of *necessitarianism*, the thesis that at least one proposition is necessarily true. It will be maintained in this paper that necessitarianism is false and that possibilism is true.

One might postulate a link between possibilism and the idea of *epistemic monism*, that there is only one basic epistemic method for investigating the world. It is invariably difficult to say what is monistic about monisms, since there can be disagreement about what is a 'basic' category. Again, within good scientific method (of which there are many different accounts) one can discern distinct roles for theory and sensory experience; though it is common currency these days to acknowledge a unity within scientific method between adjusting theory to sensory information, and adjusting the interpretation of the senses to well-constructed theory. But philosophers have usually shown a preference for monisms over dualisms, a preference which I share and which I think should be recognized in the theory of method, epistemology, as much as anywhere else. When Popper, for example, strayed from his general fallibilism, it was because he supposed a fundamentally different method for establishing logical truth; and Lakatos' criticism of Popper's conventionalism about both logic and mathematics was in the name of a properly general fallibilism.[1]

Erkenntnis **30**: 319–337, 1989.

The first question to ask is about the sources of necessitarianism, and this will be pursued in the following sections. I postulate two sources of necessitarianism, conventionalist necessitarianism and model theoretic necessitarianism, and in the next two sections these are discussed and rejected. In later sections, I endeavour to persuade the reader to abandon necessitarianism altogether. In the final section, I return to the question of the link between possibilism and epistemology, and maintain that only possibilism is the natural epistemological direction to take. One final point: intuitions about possibility seem to be rather easier to come by than intuitions about necessity. It seems to me that we have a strong sense that *some falsehoods are yet possible*, and that is my starting point.

2. CONVENTIONALIST NECESSITARIANISM

Popper and Lakatos objected to the *strategy* of conventionalism, by which they meant the method of preserving a thesis come what may from empirical refutation or criticism. Thus, *qua* strategy, conventionalism about any proposition violates correct fallibilist practice. But conventionalism, particularly conventionalism about logical truth, lends itself to a group of deeper, semantical theses as well. The distinctive appeal to *convention* is an appeal which following Popper,[2] we can characterise as *controllability in principle by decision*. We need to understand the effect this has on necessitarianism. In maintaining that necessary propositions form a nonnull class, conventionalist necessitarianisms have held that a convention can sometimes suffice to make a proposition true, and sometimes not. This evidently raises the question of how such a thing could be and that is discussed presently. Contrary to Wachbroit,[3] some version of conventionalist necessitarianism seems still to be believed by many philosophers, who were taught it as twentieth century orthodoxy. On the other hand, I would agree with Wachbroit that many philosophers do seem to think that Quine's 'Truth by Convention' is decisive against conventionalist necessitarianism, and that the latter is not obviously so.[4] Susan Haack argues that Quine's changing views on the status of the laws of logic are far from satisfactory;[5] and certainly Quine's strategy in 'Truth by Convention' of requiring conformity to the ordinary language meanings of the connectives is as conservative and conventionalist in strategic effect as his later arguments in *Word and Object*.[6] Conventionalism needs re-killing.

It is important not to confuse conventionalism here with what Grünbaum called *trivial semantic conventionalism*. A simple distinction between sentences and propositions serves to make the point. Evidently the relationship between symbols and their meanings is in some sense controllable by a decision to associate a word with one semantic item rather than another. But this fact is entirely compatible with all propositions enjoying an equal epistemic status, since indeed it is a common feature of all sentences. Such semantically based conventionalism is thus trivial. It does not suffice to sustain a distinction between logical truths and fallible, empirical truths. Trivial semantic conventionalism seems to be behind the view of many philosophers that you can create logical truths simply by stipulations or resolutions that words have one meaning rather than another. But it should be apparent that it precisely does not serve to make a difference between logical and contingent truth. What would be needed for that would be a prior distinction between those kinds of meanings or propositions which, when fixed, *ipso facto* have a truth value; as opposed to those propositions whose truth value also varies with, or is determined by, the world. Grünbaum himself, as is well known, espoused metric conventionalism, which he explicitly denied was trivial. (See [4]). He argued that *even* given satisfactory semantics for metric propositions, a continuous space is *indeterminate* with respect to those metrical features, as demonstrated by the existence of alternative, incompatible metrical descriptions of it. Note the distinctive indeterminateness thesis. If a decision beyond a trivial semantic convention is needed to (or able to) determine truth, then the world must not determine the truth of the proposition in question.[7] We should be clear that it is not simply that there is something wrong with an indeterminateness thesis about a class of propositions, since presumably a thesis such as Grünbaum's is at least intelligible.[8] But it is *the coupling of an indeterminateness thesis with the further claim that such propositions can have their truth determined by decisions*. It is not easy to understand in what the *truth* of a proposition, with respect to which the world played no determining role, could consist. I take it that this objection applies particularly severely to conventionalisms about the propositions of logic and mathematics. After all, how could a decision (more generally, a convention) *conceivably* make something true if the world plays no determinative role? One might feel inclined to wonder whether it was like printing several copies of the morning paper in order to make it true. Hilary Putnam puts this nicely: 'To put it bluntly, you can't make the Principle of Contradiction true by convention unless it's *already true*'.[9] Again, one might ask of any conventionalism including Grünbaum's, what could justify the use of the word 'true' of propositions with respect to which the underlying universe is indeterminate. If such are controllable by decision and escape control by an indeterminate world (while other propositions do not), then it seems odd that the word 'true' would be worth using of them at all. That way lies pragmatism.

Quine's changing views on logical truth mean that one must be careful in specifying just which parts one disagrees with. First, for all his apparent gradualism and fallibilism, Quine usually did give classical first order logic a unique role to play in theory, a role which I think has actually served to stifle inquiry into nonclassical logics in the recent past. This is, simply, conventionalist strategy. In fairness to Quine, his views on the conservation of classical logic were supported by the well-known arguments about radical translatability which I will not discuss here; though to the extent that they amount to a sophisticated version of the philosopher's 'I do not understand', more will be said in the next section. Second, while I do not deny that there can be an intelligible decision *never to give up* a certain class of propositions, that is a very different matter.[10] With this, my quarrel is rather different, namely that it is silly to make such decisions. Unless differences in the world under-determine differences in the truth value of the proposition in question, then to make a decision to believe a proposition in advance of ordinary epistemic investigation is to close one's eyes to the possibility of revision, as Quine the fallibilist insisted on other matters. Third, while Quine shied away from the notions of necessity, possibility and analyticity, he certainly thought early and late that conformity to classical logic and in particular to classical consistency was a rigid constraint on theory. That is *some* kind of impossibility thesis: it is one thing to hold that a proposition is inconsistent but something stronger to assert with Quine that this is invariably sufficient for its prohibition. On the other hand, I do not propose an account here of whatever it is that possibility consists in. The aim is, rather, to remain neutral on various of the going accounts, for instance various realisms, and to argue for lifting bounds on the *extent* of possibility.

The earlier distinction between propositions which escape control by the world and propositions which do not, sat uncomfortably with the positivist-conventionalists, who often tried to deal with it with some version of the thesis that analyticities were not 'cognitively meaningful'. A typical conventionalist manoeuvre here is to deny that analyticities are in any straightforward sense true or false, that necessary truth is a kind of truth.[11] Popper's positivist links show here, for example, with his descriptions of them as 'truisms' and 'empty'; though his most favoured terminology is that also favoured by intuition, that 'All cats are cats' and its ilk are, simply, true. The problem of the nature of necessity arises just because some *truths* have seemed to have a special status.

I conclude that conventionalist necessitarianism is in serious trouble, both epistemically, and with its semantic underpinnings. However, mention of the matter of the literal truth of necessary truths brings me to my second classification of necessitarians. If necessary truth really is truth, then what sort of literal truth is it? One answer which many have been inclined to give, is that it is in virtue of being true in all members of some class of models that propositions are necessary. I call this model-theoretic necessitarianism, and in the next section it will be criticised and rejected.

3. MODEL THEORETIC NECESSITARIANISM

The view considered in this section is that necessity (and, let it be added, validity[12]) arises in virtue of playing a distinctive role in all models or semantical assignments. To avoid conventionalist problems about 'truisms', it is frequently claimed that the existence of models and truth preservation is a perfectly objective phenomenon, prior to mere syntax. Now one quick point to make is that this is frequently in an even worse position than conventionalism over its epistemology. At least conventionalism makes a semblance of squaring itself with its own epistemology, in that *declaring a word to have a meaning* is an event we seem to be able to get into epistemic contact with, even if the latter saddles us with the magical power of a decision to make something true when the world is supposedly unable to. But if necessary truth and validity reside in timeless objective models (for instance the possible worlds of the modal realists), then we should want to know with which faculty we are able to get an inkling into their existence. More will be said on this in the next section. Here I want to put the criticism that model-theoretic necessitarians either take too simplistic a picture of what structures are available to them, or saddle themselves with an unarguable and uncriticisable metaphysics.

The fact is, that modern semantics has gone far beyond the simple models of classical first-order logic. Intuitionist models give up Excluded Middle, Paraconsistentist models give up the law of Non-contradiction, the various models of *Relevant Logics and their Rivals* [25] split apart vast numbers of logical theses which are equivalent in the context of classical logic, showing how to maintain one without the other. Some of these are considered in more detail in the next section. The semantics of classical first order logic is very much a special case of a much wider semantical framework, many of the details of which have only become apparent recently. I do not maintain the thesis that the *mere* existence of alternative models *suffices* to demonstrate that propositions refuted therein are not really necessary truths. I am making the weaker point that a model-theoretic necessitarian is in no position to say that various theses are logically true *solely* on the grounds that they hold in all models. Therefore, the model-theoretic necessitarian is in the position of having to argue that certain models have a *preferred* status over others. But it is not easy to see how this is to be done short of metaphysical dogmatism. The model theory *by itself* does not provide this. As a case in point, consider a dispute between two modal logicians defending, respectively, S4 and the Brouweresche system BR. The former maintains that S4 gets necessity right because it contains exactly the theses valid in all models consisting of possible worlds related by an accessibility relation which is reflexive and transitive. The latter begs to differ, saying that relative possibility is reflexive and symmetric. How is this to be resolved? Don't say: by looking at how the sentences match their ordinary language counterparts. We are considering here a dispute between model-theoretic necessitarians; that is, people who maintain that necessity is to be explained by the models. Needless to say, the attack on model-theoretic necessitarianism should not be construed as an attack on the discipline of model theory as such. If there is a nonnull class of necessary truths, then presumably there will be some class C of models such that A is necessary iff A is true in all members of C; and it is the business of model theory to uncover such facts. My concern has been to cast doubt on the belief that the 'iff' *suffices* to furnish an explanation of the left hand side by the right hand side.

My broad point is recognised in a discipline close to modal logic, namely tense logic. The history of philosophy has seen many marvellous claims as to the necessary structure of time. Arthur Prior's pioneering work on the expressability of alternative classes of temporal structures by different classes of tense-logical theses, led naturally to the thought that none captured a necessary structure for time, so much as presented an alternative way time might be. The analogy isn't perfect; but it suggests not only that model theory does not by itself prove necessitarianism, but also that the existence of alternative models ought to make us less confident that any of them expresses necessary truth. The next section amplifies this last point.

4. UNDERSTANDING AND INTUITIONS

It is important to avoid the bewilderment response here: 'But I simply don't understand what 'true' and 'false' could amount to, if it is being proposed that there are models where one and the same linguistic item can be both, or various other logical laws are violated'. The short response is to urge the reader not to give way to such *semantic dogmatism*. There should of course be no quarrel with honest bewilderment, but that is far from dogmatic. The undogmatic response would be to consider seriously that such semantics amount to metaphysical proposals, of a highly general and theoretical kind, as to how the world might be.[13] The situation in 'logical physics' is simply no different from that in theoretical physics, only newer and more abstract. Uncriticisable dogmatism is as out of place in the former as the latter. Thus Popper: 'An argument that proceeds from inconceivability is, like other self evident arguments, always suspect'.[14]

It is all very well to make this point generally, but it seems to me that it helps noncomprehension to see some recent nonclassical logic. Model theory has a useful epistemic role to play, like that of mathematical physics; or differently, of fiction. It can feed into whatever faculty it is which enables us to conceive of false possibilities and contingencies, to expand those intuitions. So in this section I digress into some recent technical work, and consider how it might aid incomprehension.

The *Law of Excluded Middle* and its relatives such as the Law of Bivalence, have come under a lot of pressure lately. Philosophers have often not found it so difficult to believe that some propositions might be neither true nor false. Brouwer and Heyting felt the same about mathematics, and Dummett has recently advanced a similar view about a broader class of propositions. My aim here is not to defend these views, so much as to invite the reader to ask himself or herself the question *what if they are right*? Good philosophy always attempts to understand one's opponent; and these writers have provided extensive descriptions of how they see the world, how the world *would* be if they *were* right. Needless to say, it is open to believe that the Law of Excluded Middle is *true*, without holding the extra thesis that it is *necessary*. And I suggest that in the presence of descriptions of how intuitionists imagine the world to be, it is difficult to feel confident about the *impossibility* of their view.

Recently, in the work of da Costa, Meyer, Priest, Routley, Rescher and Brandom and others, the truth of the *Law of Noncontradiction* has come under much fire. Like Excluded Middle, the Law of Noncontradiction is a group of theses, a representative being: no proposition is both true and false. Philosophers on the whole seem to find the breakdown of Noncontradiction harder to swallow than the breakdown of Excluded Middle, though it is well-known that the former in the form $\sim(A \& \sim A)$ is equivalent to the latter in the form $A \vee \sim A$, given only the independently plausible laws of Double Negation and de Morgan. A common case for giving up Noncontradiction is that it provides the most natural and unified treatment of the set-theoretic and semantic paradoxes, and logics exist which tolerate and contain the effect of contradictions by giving up the classical law *Ex Contradictione Quodlibet* $(A \& \sim A)/\therefore B$. In the course of the development of such logics it became clear that there is more than one position one could adopt. *Strong paraconsistentism* holds that some contradictions are true. *Weak paraconsistentism* holds that no contradictions are true, but that inconsistent 'worlds' need to be admitted to one's model. A version of weak paraconsistentism holds that inconsistent 'worlds' describe possibilities, so that while no contradictions are true, some contradictions are possible. The proof theory and model theory for large classes of such logics is well developed and well discussed by now.

Now a common response by philosophers to the suggestion that a contradiction might be true, is *I don't understand*. The noncomprehension may be genuine, but it does not justify the often-accompanying rejection of the suggestion. One is reminded of philosophers' noncomprehension of the idea that space might be four-dimensional, or curved. There is, of course, a quick argument against noncomprehenders who also hold that contradictions are necessarily false: if you think they are false, then presumably you understand them. To be fair, the noncomprehension often unpacks as 'I don't know what it would be like for a contradiction to be true'. The remedy might well be to read the literature and to think about the way paraconsistentists imagine the world to be.

Sometimes, noncomprehension is accompanied by a manoeuvre which I call the *Three Monkeys' Decision*: the decision to use negation in such a way that contradictions cannot be true. I do not think that this decision is any more plausibly motivated than its conventionalist cousins. One such motivation is that one can determine meanings by decision, but the meanings of linguistic items are also linked to the role those items play in theory. So, it might be thought, there is a sense in which some propositions are impossible: the sense that if the accompanying theory is false then it needs terms with different meanings to describe that fact.[15] But this odd-sounding incommensurability thesis ought to alert us that something is wrong. For if the accompanying theory is false, then surely it is false as it is, with all its terms having whatever meanings their theoretical roles are capable of giving them. The trouble is, as it so often is, that a distinction has already been imported between the status of putatively necessary propositions and ordinary scientific propositions. Supposing that the meaning of ordinary scientific terms like 'electron' is tied to the role they play in theory, this does not ensure that the theories are true. They might be false, with no change in the meaning of 'electron'; and what then would we make of the 'decision' to use 'electron' such that the electronic theory of matter is true? This analogy has a further use. Let it be conceded that if a certain proposition turns out unexpectedly to be false, the new theory needed to describe the world might employ terms with sufficiently different theoretical roles to make it worth saying that they have different meanings. Commensurability has not been sacrificed, however, since we still assert that the old theory is false. But also if the old theory were not too bad, if it got things fairly right in a limited domain even down to its theoretical concepts, then we *might* find sufficient analogies between corresponding theoretical roles in the old and new theories to be worth using the same term for both. This goes as much for negation as for electrons.

Perhaps the reader is comforted by the thought that there is at least one proposition which is absolutely irrefutable, the Law of Propositional Identity, $A \rightarrow A$. This brings me to the topic of *Martin's Theorem*. Martin's Theorem says that in a certain weak propositional calculus, no instance of $A \rightarrow A$ is provable.[16] The proof of the theorem proceeds by showing for given any instance of $A \rightarrow A$, how to construct a model in which it is false. The models are very abstract structures, as is not uncommon in algebraic and operational semantics, and do not on the face of it look much like the spacetime manifolds some have imagined possible worlds to be. And that is a significant feature of the present position: why couldn't what there is be an abstract structure? Philosophers have not found so much difficulty in supposing that abstract entities exist, though obviously I am not claiming that they *do* exist. Furthermore, some recent theorists have suggested that the existence of universals might be contingent,[17] and have proposed an account of contingent laws of nature on that basis. If you are one of these, then give serious consideration to the possibility that contingent, abstract entities might have been *all* that exist. If a way for $A \rightarrow A$ to be false is that the world be very different from the way it is now, then why not? Why *couldn't* the world have been nothing but (say) some three-valued semantical algebra? This point is reinforced by the observation that it is important not to confuse *generality* with *abstractness*. Our world has very general structural features too, for instance very general aspects of its differential topology. It is possible to present General Relativity, Quantum Mechanics, Gauge Theory, even Newtonian Dynamics in very abstract fashion. Considered in isolation from the concrete universe out of which they arise, it can be difficult to grasp their connection with our world. I suggest that things might well be that way with abstract-looking logical countermodels too. Something might be a universe not-too-dissimilar from our own, yet have structural aspects which render false all manner of propositions such as $\sim(A \,\&\, \sim A)$ or $A \rightarrow A$. There is, I suggest, no reason why such very general or abstract structures should not be realised. And if so, then they represent false-but-possible states of affairs.

The importance of Martin's Theorem is considerable, I think, since it bears on well-entrenched intuitions about analyticity. It is seductively easy to believe that 'If Smith is a bachelor, then Smith is an unmarried man' can be made true by the decision to use 'bachelor' interchangeably with 'unmarried man'. But it would take more than that to make it necessary. It needs 'If Smith is a bachelor then Smith is a bachelor' to be necessary also, and that is a substitution instance of $A \rightarrow A$. Indeed, it seems to me that the intuitive solidity of mathematics rests on the same foundation. Short, quite obvious inferences in mathematics often derive, like the previous bachelor case, from some definitional decision to use terms interchangeably applied to $A \rightarrow A$, (or to $(A \,\&\, B) \rightarrow A$ or $A \rightarrow (A \vee B)$).[18] Mathematical connections established by longer chains of reasonings appealing to more complex deductive principles are to that extent less evidently necessary. I am not suggesting here that it is *easy* to understand how standard mathematics might have been false. But then we should beware of projecting the limitations of our imaginations onto the world. The easiest understanding I am able to offer here is of the order of difficulty of whatever would make $A \rightarrow A$ false; and that, as has already been noted, looks to be pretty strange stuff.

I trust that these examples have been sufficient to shake the reader's confidence that any proposition is impossible. But note that they share the feature of arising at the level of propositional logic. So we should also look inside atomic propositions, to see if necessity might arise from sub-atomic structure. Here, as elsewhere in this paper, the weakness of the imagination limits consideration of every candidate for necessity which might be put up. But I do want to show some possibilities connected with the Law of Identity, everything is self-identical. I report here the work of various theorists, including Brady, da Costa, Meyer and others.

For instance, what would it be like for everything to be self-identical but also some things to be non-self-identical? That would be an inconsistent universe, but that should be the last of our worries by now. We can take set theory as our model for this possibility. Suppose we take identity to be: every member of the one is a member of the other and vice versa; and non-identity to be: $x \neq y_{\mathrm{df}} (\exists x)(z \in x \mathbin{\&} \sim z \in y)$. Now consider an inconsistent set theory[19] containing the Law of Identity $(x)(x = x)$ and containing in addition the Russell set R. It takes only a few principles to show, as Russell did, that $R \in R \mathbin{\&} \sim R \in R$. Whence $(\exists z)(z \in R \mathbin{\&} \sim z \in R)$, i.e. $R \neq R$. There are variants of this for different accounts of identity, which exploit variants of the Russell Paradox, e.g. (i) $x = y =_{\mathrm{df}} (\forall z)(x \in z \equiv y \in z)$ and $x \neq y\, y =_{\mathrm{df}} (\exists z)(x \in z \mathbin{\&} \sim y \in z)$. Again, the Russell Set with $R \in R \mathbin{\&} \sim R \in R$ gives $R \neq R$, (ii) $x = y =_{\mathrm{df}} (\forall F)(Fx \equiv Fy)$ and $x \neq y_{\mathrm{df}} (\exists F)(Fx \mathbin{\&} \sim Fy)$. For F, take '$\in R$', (iii) $x = y =_{\mathrm{df}} (\forall u)$ (x participates in $u \equiv y$ participates in u). Then, supposing it were true (why not!) that $(\exists u)(\forall x)$ (x participates in $u \equiv Fx$), we have $u = u \mathbin{\&} u \neq u$. A somewhat different approach develops inconsistent number theories,[20] according to which $(x)(x = x)$ but also $O \neq n$ for some n. This produces $O \neq O$ and, indeed, $n \neq n$ for every n so that $(x)(x \neq x)$ as well. These structures do not have every proposition true in them, since $O = 1$ is true in none of them. Indeed in some such structures, distinct numbers are distinct so that $n \neq m$ for distinct n, m. In fact it is possible to produce such structures in which the domain is divided into two disjoint classes, those for which $n = n$ but not also $n \neq n$, and those for which $n \neq n$ but not also $n = n$; which gives a universe in which some things are self-identical and some (other) things are non-self-identical. Again, a different approach might take self-identity as the criterion of existence. This is not uncommon; recall the definition of the null set $\Lambda =_{\mathrm{df}} \{x : x \neq x\}$. This gives the possibility of a universe in which nothing is self-identical and everything is non-self-identical. Indeed, this universe seems to be consistent. At least, it is consistent if it is consistent that nothing exists, which has seemed a desirable proposition to some.

The aim in this section has been to outline recent technical experimentation which, at the very least, has to be recognised as exploring the limits of conceivability. In the absence of successful general arguments for setting those limits narrowly, one has to take a tolerant attitude to which abstract and general theories count as possible, a practice which is entirely uniform with theoretical science. Needless to say, it is the aim of this paper to criticise attempts at such general arguments; and in the next section, a group of such arguments are considered.

5. A PRIORI ARGUMENTS FOR NECESSITARIANISM, AND MINIMAL METALINGUISTIC CONSISTENCY

In 'There Is At Least One A Priori Truth', Hilary Putnam tries out, then rejects, an argument to the effect that logical truth is needed to constitute rationality.[21] Since he is attracted by the epistemic link between fallibilism and possibilism, he proposes the view that every statement is revisable in the sense of 'challenging a concept it contains', but that it is not the case that the rational revision of every statement permits denying it. Evidently this is an attempt to save some logic. Putnam notes that it echoes Quine's well-known translatability-of-the-connectives argument, which I take to be a sophisticated version of the 'I don't understand' argument criticised earlier. Putnam is inclined to believe the argument that it would follow that it is a necessary truth that if the relevant concepts are not revised, then *P*, for suitable *P*. The correct strategy against Putnam here would seem to be to deny the premiss of the argument, that the rational revision of some statements does not permit denying them. It has already been argued that such denial can be *intelligible*, and there does not seem to be any obvious methodology to force us to admit that it is *irrational*.

Putnam is also worried by the thought that it looks pretty necessary to assert that *not every proposition is both true and false*. In a similar vein, it used to puzzle me whether the following might be a candidate for necessity: *at least one proposition fails to be true*. But I do not think so. What would the universe be like for that not to hold? That is easy to describe: every proposition would be true. Needless to say, that is not how things *are*. But necessarily so? We can *say* easily enough what it would be for every proposition to be true. It has to be admitted that it is not always easy to understand the claim that *P* is possible, for selected extreme *P*'s, but it is good advice not to trust one's noncomprehensions, especially when there is comprehension here of a sort. Again, consider the proposition that *at least one proposition is true*. Of course that is true, but consider the following argument which might be made out for its necessity. Suppose it were false that at least one proposition is true. Then the proposition that it is false that at least one proposition is true, would be true. Hence at least one proposition is true. Hence, it is necessary. What I think can be said against this argument is that it turns on the principle $\sim A \rightarrow A / \therefore \Box A$. While this would be plausible if the $\rightarrow$ in question were entailment, it should not tempt anyone if the $\rightarrow$ is some ordinary (nonentailmental) 'if . . . then'. It fails, for instance, if the $\rightarrow$ is $\supset$. One would need a further necessitarian premiss if one is not to be accused of simply pulling a necessitarian conclusion out of the hat.

One might wonder whether there is a transcendental argument for some sort of metalinguistic controls which amount to necessity. With any assertion we make, for instance the assertions in this paper, we presumably also do not wish to put forward its denial. This suggests a principle of *minimal metalinguistic consistency*[22] as a constraint on assertibility or perhaps rationality; or at the *very* least minimal metalinguistic *nontriviality*, that is that one could hardly intelligibly entertain all one's assertions and *all* their denials. Now I do not think that this is an easy argument to come to grips with, though it must be agreed that it touches on very deep intuitions about assertibility and intelligibility. The situation is complicated by the fact that, *pace* Tarski, it is notoriously difficult to make out a case for a rigid distinction between object language and metalanguage for useful natural languages, only for artificial languages which fall short of reality. So one might be inclined to conclude that asserting any contradiction violates a metalinguistic constraint on assertibility and so intelligibility. The further conclusion is thus that contradictions are unassertible, thus unintelligible, and so their denials are necessarily true.

I think that the argument here is better off aiming for the weaker conclusion that one's theory should be minimally *nontrivial*, rather than the stronger conclusion that it should be *consistent*. As to the latter, I would agree that in asserting A one does not also typically intend to deny it, but it might be that one's conceptual discoveries can *surprise* one. One might find that there are isolated instances where one is *forced* to contemplate inconsistency: consider recent paraconsistentist work on inconsistent solutions to the Liar Paradox and Russell's Paradox, or on the possibility that the empirical science of motion might need to be inconsistent. Consistency, even metalinguistic consistency, is one among many theoretical desiderata, and fallibilism should recognise that consistency might occasionally best be sacrificed in the interests of overall theoretical economy. That is to say, *our concepts escape our control*: we may seek to quarantine the true from the false, but discover that this is not entirely achieved in the best theory available. To repeat an earlier point, this does not mean that 'true' and 'false' have lost all their meaning, since they may well retain strong theoretical links and parallels with the Tarskian concepts, sufficient for it to be worth viewing as a theoretical discovery about the concepts of truth and falsity. The same observations apply to minimal metalinguistic nontriviality. Remember that it is not being maintained that every proposition and its denial *are* true, only that it might have been so. It would be hard to make out a case for such a universe to be *interesting*, and just as hard to envisage anyone seriously asserting that it is our universe. On the other hand, it would seem to be that in *our* universe the conditions for the minimal intelligibility and assertibility of the proposition of the nontriviality of the world are fulfilled, and thus the conditions for the intelligibility of the denial of that proposition would also seem to be met. The latter does seem to be in accord with intuition.

6. CONCLUSION

We have seen that some, though not all, of the considerations urged against various versions of necessitarianism are epistemological in spirit. Now Susan Haack has pointed out that the doctrines of fallibilism and necessitarianism are formally consistent.[23] Whatever one makes of fallibilism, it is easy to assent to the proposition that good scientific method is consistent with the existence of at least one necessary truth. However, I maintain that one can sidestep these points by holding that a properly general epistemic monist conception of scientific method, fallibilist or not, gives no reason to think that any proposition is necessary and every reason to think not.

After all, consider what theoretical function fallibilist necessitarianism might serve. Why should one believe that at least one proposition is necessary? In the first place, it is hard to make sense of how our *perceptions* of the world might make for a difference between two different sorts of truths, sufficient to dignify one kind as necessary and the other not. We do not perceive, with any sense organ, anything more than that propositions are true, certainly not that they are necessary. There is no epistemic dualist split in perception.

Thus if necessity is to get an epistemic foot in the door at all, it should be conjectural or postulational. One would in one's theory conjecture of certain propositions not merely that they are true, but *also* that they are necessarily true. I suggest that the role of such necessitarian elements in the theory would be, indeed is, to erect a barrier to legitimate criticism of the theory, to avoid having to back up one's position against critical scrutiny.

Ask yourself how often you encounter necessitarian elements in philosophical theory. I think that they are frustratingly ubiquitous. Philosophers are hooked on *a priori* certitudes like logical junkies, unable to discard the necessitarian needle. No doubt this is partly a hangover from the not-so-distant past when philosophy was thought of as discontinuous from science, as second order and essentially *a prioristic*. But such dualism ought to be seen to be as undesirable as it is unnecessary. I submit that *necessitarian explanations are simply never needed in theorising about reality*. Putnam nibbles at this:

> Nor do we really need a proof that a statement is *a priori* in this sense (rationally revisable) very often. If a statement has the property that *we cannot now describe* any circumstances under which it would be rational to give it up, that will surely suffice for most purposes of philosophical argument.[24]

So I suggest that the present position would seem to have many interesting ramifications as epistemology is naturalised throughout philosophy, abolishing the hydra of necessitarianism. If a chain of responses to requests for reasons is stopped with the claim that a particular premiss is necessarily and self-evidently true, then nothing is added and no light is shed. '*P* because *Q*' may be helpful, but '*P* because necessarily *P*' is useless. '*P* because I can prove it' only invites the request to do so. Certainly human debate stops, but Putnam showed the correct stopping point: we stop when imagining an alternative is beyond us.

Summing up, nothing ensures that the principles that the formal logician chooses to follow are necessarily true, but then nothing prevents us from error in any case. A properly general and unified theory of the world is to be desired, and the present view seems to be the only candidate around.

NOTES

1 In defence of these claims, see my paper with Tim Burgess [15]; see Popper [20] and Lakatos [8, 9]. On the term 'possibilism' see Naess' excellent [16] and Nerlich's [17].

2 See Popper [20] p. 78, or [15]. Also in accord is Putnam's [22]. The point of the 'in principle' is to allow for the fact that not all conventions originate in actual decisions.

3 See Wachbroit [28], pp. 48–9.

[4] Wachbroit [28], p. 50; Quine [23].
[5] Haack's [5] is a good account of these changing views.
[6] Quine [24]. Wachbroit presses Quine's objection in 'Truth by Convention' that no finite set of conventions can determine an infinite set of logical truths, and argues that the attempt to salvage this by allowing self-referring conventions falls foul of a diagonal argument. But if even one necessary truth is determined by a convention, necessitarianism is established; so that needs prevention, see [28], pp. 51–3.
[7] Popper: 'It is not the properties of the world which determines this construction [of science]; on the contrary it is this construction which determines the properties of an artificial world', [20], p. 79. For a discussion of indeterminateness in conventionalist semantics, see Nerlich [18] p. 100.
[8] Even so, it would be a distinctly odd thesis: in what sense would the world be indeterminate about whether all cats are cats?
[9] Putnam [22], p. 163.
[10] Even here, though there is a difficulty if 'not give up' amounts to 'continue to believe'; for then the idea that beliefs be decision-controlled amounts to the thesis that beliefs are actions, which is controversial.
[11] Cognate views are evidently the indeterminateness thesis discussed earlier, or some 'no proposition expressed' thesis.
[12] The remarks here are intended to apply as much to those necessitarians for whom proof precedes truth, as the other way around.
[13] See e.g. Routley, [26].
[14] [21], p. 207.
[15] Haack [5] argues against the thesis that there may be truths which are 'analytique' in the sense that everyone learns that they are true by learning the words, attributing it to the later Quine.
[16] Proved by Errol Martin in [10]; see also Martin and Meyer [11]. It constituted a positive answer to Belnap's Conjecture, that in the logic S, consisting of rule transitivity and axioms and rules of prefixing and suffixing, no instance of $A \rightarrow A$ is provable. Martin and Meyer have exploited the result to defend the view that Propositional Identity is not a logical law and that validity is captured exactly by the theory of the syllogism, but it seems to me that the latter conclusion is uwarranted.
[17] Armstrong [1], Tooley [27].
[18] It goes without saying that the latter two have been questioned: $A \& B \rightarrow A$ is denied in connexivism on which the literature is considerable, e.g. McCall [12]. Many find $A \rightarrow A \vee B$ of doubtful value, contrary to Haack; see Parry [19].
[19] Brady [2] or da Costa [3].
[20] [13] or [14].
[21] Putnam [22]. Popper runs a similar line, that logic is needed as an 'organon of criticism', and that classical logic is the preferred logic since it alone maximises retransmission of falsity. This argument is discussed and rejected in [15].
[22] I am indebted for raising these points to Dr. Rainer Trapp, and a referee of this journal.
[23] Haack [6] and [7].
[24] [22], p. 170.

REFERENCES

[1] Armstrong, D. M.: 1978, *Universals and Scientific Realism*, Cambridge University Press, Cambridge.

[2] Brady, R.: 1971, 'The Consistency of the Axioms of Abstraction and Extensionality in Three Valued Logic', *Notre Dame Journal of Formal Logic* **12**, 447–53.

[3] Da Costa, N. C. A.: 1974, 'On the Theory of Inconsistent Formal Systems', *Notre Dame Journal of Formal Logic* **15**, 497–510.

[4] Grünbaum, Adolph: 1964, *Philosophical Problems of Space and Time*, Routledge and Kegan Paul, London.

[5] Haack, Susan: 1977, 'Analyticity and Logical Truth, the Roots of Reference', *Theoria* **43**, 129–43.

[6] Haack, Susan: 1978, *Philosophy of Logic*, Cambridge University Press, Cambridge.

[7] Haack, Susan: 1979, 'Fallibilism and Necessity', *Synthese* **41**, 37–63.

[8] Lakatos, Imre: 1976, *Proofs and Refutations*, Cambridge University Press, Cambridge.

[9] Lakatos, Imre: 1978, *Mathematics, Science and Epistemology, Philosophical Papers Vol. 2*, J. Worrall and G. Currie (eds.), Cambridge University Press, Cambridge.

[10] Martin, E. P.: 1978, *The P = W Problem*, Ph.D. dissertation, Australian National University.

[11] Martin, E. P. and R. K. Meyer: 1982, 'Solution to the P-W Problem', *Journal of Symbolic Logic* **47**, 869–87.

[12] McCall, S.: 1966, 'Connexive Implication', *Journal of Symbolic Logic* **31**, 415–33.

[13] Meyer, Robert K. and Mortensen, Chris: 1984, 'Inconsistent Models for Relevant Arithmetics', *Journal of Symbolic Logic* **49**, 917–29.

[14] Mortensen, Chris: 1988, 'Inconsistent Number Systems', *Notre Dame Journal of Formal Logic* **29**, 45–60; and 1987, 'Inconsistent Nonstandard Arithmetic', *Journal of Symbolic Logic* **52**, 512–518.

[15] Mortensen, Chris and Burgess, Tim: 'On Logical Strength and Weakness', *History and Philosophy of Logic*, to appear.

[16] Naess, Arne: 1972, *The Pluralist and Possibilist Aspect of the Scientific Enterprise*, Allen and Unwin, London.

[17] Nerlich, Graham: 'Pragmatically Necessary Statements', *Nous* **VII**, 245–268.

[18] Nerlich, Graham: 1976, *The Shape of Space*, Cambridge University Press, Cambridge.

[19] Parry, W. T.: 1933, 'Ein Axiomsystem für eine neue Art von Implikation (Analytische Implikation)' *Ergebnisse eines Mathematischen Kolloquiums* **4**, 5–6.

[20] Popper, Karl: 1959, *The Logic of Scientific Discovery*, Hutchinson, London.

[21] Popper, Karl: 1963, *Conjectures and Refutations*, Routledge and Kegan Paul, London.

[22] Putnam, Hilary: 1978, 'There Is At Least One A Priori Truth', *Erkenntnis* **13**, 153–170.

[23] Quine, W. V. O.: 1966, 'Truth by Convention', in *The Ways of Paradox*, Random House, New York.

[24] Quine, W. V. O.: 1960, *Word and Object*, MIT Press, Cambridge.

[25] Routley, Richard et al.: 1983, *Relevant Logics and their Rivals*, Ridgeview, Atascadero.

[26] Routley, Richard: 1976, 'The Semantical Metamorphosis of Metaphysics', *The Australasian Journal of Philosophy* **54**, 187–205.
[27] Tooley, M.: 1977, 'The Nature of Laws', *Canadian Journal of Philosophy* **7**, 667–98.
[28] Wachbroit, Robert: 1987, 'Logical Compulsion and Necessity', in *Erkenntnis* **26**, 45–56.

Manuscript submitted 9 September 1986
Final version received 5 November 1987

Department of Philosophy
The University of Adelaide
North Terrace, S.A. 5001
Australia

HISTORY AND PHILOSOPHY OF LOGIC, 10 (1989), 47–51

On Logical Strength and Weakness

CHRIS MORTENSEN
Department of Philosophy,
University of Adelaide, North Terrace, S.A. 5001, Australia

and

TIM BURGESS
Department of Traditional and Modern Philosophy,
University of Sydney, N.S.W. 2007, Australia

Received 4 March 1988

First, we consider an argument due to Popper for maximal strength in choice of logic. We dispute this argument, taking a lead from some remarks by Susan Haack; but we defend a set of contrary considerations for minimal strength in logic. Finally, we consider the objection that Popper presupposes the distinctness of logic from science. We conclude from this that all claims to logical truth may be in equal epistemological trouble.

1. Introduction

Good cautious epistemology ought to ask for reasons why one should believe in one logic rather than another. Much recent theorising has pursued this question in the particular so to speak, by considering individual theses and various arguments for and against their logical truth. No doubt this is symptomatic of the fact that the epistemology of logic seems still to be in a fairly unsettled state. In this paper, we aim to do several things. In the second section, we consider an argument due to Popper for maximal strength in choice of logic. We dispute this argument, taking a lead from some remarks by Susan Haack but adding further considerations. We then defend a set of contrary considerations, for minimal strength in logic. In the third section, we consider the objection that Popper presupposes the distinctness of logic from science, and caution about how easy it is to do the same, arguing that both Haack and Grattan-Guinness can be viewed as doing so. We conclude from this that all claims to logical truth may be in equal epistemological trouble.

2. Logical strength and logical weakness

Popper the fallibilist argues for a particular logic as a deductive tool in empirical science, classical two valued logic. He distinguishes between the use of logic as a proof tool in mathematics, and its use as an organon of criticism in empirical science. Of the former he says that minimal logic is best since it represents a sceptical attitude toward what is to be proved. We concentrate in this paper on the latter. His argument proceeds from a principle of maximum uncharity in criticism of empirical theories. The best kind of criticism is the one which makes it hardest for the defence. The logical organon most suited to this is one which retransmits the most falsity backward along deductive chains, and that one must be maximal in the sense that it makes valid the most deductions A/∴B (Popper *1973*, 304).

Susan Haack properly objects to this argument for two reasons: first, that it is not obvious that the strongest logic is the severest critic, since a criticism which only needed a weaker logic would seem to be more severe; and second, that Popper does not apply the argument to the theories of empirical science themselves, so he must be presupposing a sharp distinction between logic and science which is hard to justify (Haack *1974*, 35–38).[1] Grattan-Guinness similarly objects that logical strength can be 'brute' when applied to science (*1986*, 193). We return to these matters later.

One point to make against Popper's argument is that by itself it would not single out a unique logic, since it is now known that there are many logics maximal in the sense of being *absolutely complete* (where a logic is absolutely complete if the result of adding any nontheorem and closing under modus ponens and uniform substitution is that every wff is a theorem). Indeed, since every logic has at least one absolutely complete extension, the argument would need to maintain that every logic is a sublogic of classical logic. Perhaps Popper thought this, since his chief example of nonclassical logic is intuitionism; but it is not so. One can speculate that an attempt might be made to save the argument with some analysis of 'true' and 'false' which showed that the semantics of these terms require a sublogic of classical logic. But anyone in the business of questioning principles of logic ought not to balk at reanalysing truth and falsity; plenty of the latter has taken place recently too.

Returning to the earlier point of a distinction between logic and science, Popper's argument might also be thought to prove too much, since without a prior distinction between logical truth and nonlogical truth, the principle of strength would argue retaining as many propositions of *any* sort for critical purposes. This gullibilism is at odds with the healthy caution we emphasised earlier. Both Popper and Lakatos also emphasised a counterprinciple to criticisability, namely healthy dogmatism. A healthily dogmatic defence of a thesis is a cautiously sceptical attitude to its criticism. To hold otherwise seems to introduce an asymmetry between criticism and criticism of criticism.

Richard Routley objects to an uncritical assumption of a principle of logical strength within relevant systems, charging Anderson and Belnap with this error (*1982*, 242). His point is that the assumption derives from the mistaken view that what is wrong with irrelevant systems is principally their irrelevance; instead of seeing that relevance derives from deeper semantical desiderata which unify and explain not only relevance but also the fallacy of suppression, and more generally intuitionist, paraconsistentist and even connexivist insights. While Popper could hardly be expected to have been addressing himself to such an issue, or even aware of it, it points to the fact that a brute monolithic principle of logical strength is quite out of step with all the theorising over the last fifteen years or so about particular logical theses. For these various reasons, then, it should be rejected.

Indeed, we propose in its place a principle of logical weakness: prefer the weaker logic. In favour of this principle, we offer several related considerations. First, as is well known, a weaker logic has more theories than a stronger logic. Now the importance of the study of the theories of a logic cannot be stressed enough. One can view the theories of a logic as linking logic and science, in that logical theories can be seen as approximations and precisifications of scientific theories which thus can help to study the deductive principles on which scientific arguments proceed. Indeed, having more logical theories gives a greater chance at approximating the raw data of scientific theory. It is always open to recover extra principles if needed, as postulates holding over a less-than-fully-general range of subject matter. Philosophers of science sometimes give the impression that they think scientific theories should be stuffed into

1 The textual evidence in support of the latter, particularly from Popper *1963*, 207–212, as well as his *1973*, is pretty decisive.

the straightjacket of classical logic (see also below on verisimilitude, this section), when in truth one's logic should be informed by scientific theory, as the quantum logicians have long recognised. Needless to say, it is or ought to be a two-way process, with discoveries in logic and logical theories potentially informing scientists of the expanded possibilities available to them.

Thus we are taking sides here on the disputed matter of the empirical status of logic; and, in passing, also on the bearing of quantum theory upon it (though our argument does not depend on this special case). As to the former, we note that Haack's various works mount an admirable defence of the broad principle of the revisability of logic in the light of empirical science (see also Mortensen *1988*). On the latter, we wish only to digress to point out Dummett's error in arguing against quantum logic (in his *1978*). He identifies *realism* with the *primeness* of the world (where a theory is prime iff for every disjunction in it, at least one disjunct is also in it). But a world in which variables take ranges of values without taking any specific value from that range is arguably (given infinite disjunctions) a nonprime world, and yet realism would not obviously be impugned. It would, perhaps, be a fuzzy world (though that term is theory-laden); but it is a world in which the law of Distribution fails, in which moreover the quantum theory might hold, and, indeed, a world not known not to be our own. Also against primeness, see Mortensen and Priest *1981*, where it is argued that the truthteller paradox 'This sentence is true' is best understood as yielding the nonprimeness of the world.

A second consideration in favour of weakness develops Haack's point that a criticism which proceeds from a weaker logical base ought to be accounted as a more severe criticism. Popper notices this point (*1973*, 307), but regards it as 'not very important'. However, a criticism proceeding from weaker premisses leaves fewer ways of repairing the damage, so to speak. A recent argument in Tennant *1985* illustrates the point nicely, by applying it back to Popper's methods themselves. Tennant shows that intuitionist logic is adequate for the Popperian schema of refutation of hypotheses, and indeed that minimal logic suffices if the connective $\supset$ is dropped. Thus if the most general description of Popper's critical epistemology needs only intuitionism, he is hardly in a position to claim that maximal logic is required by his methods. But also, as it were, the 'adequacy' of the weaker logical base implies that criticisms which proceed from an *unnecessarily* strong logic are actually weaker, in that they permit apparent escape routes which are in fact blocked by the same criticism. Only if one thought with Popper that change of logic was not a genuine escape route in empirical science, would one think that this was not very important; but that is precisely our complaint against him.

A third consideration relates the previous two. The existence of a larger number of theories for a weaker logic means that a criticism developed on a weaker logical base leaves the defence without the option of modifying one's theories to match those of an intermediate logic. Here, too, we can draw on a recent example. Many philosophers have expressed the desirability of a theory of nearness to the truth. It is well known, though, that Popper's theory of verisimilitude was proved by Miller and Tichy to be subject to a severe limitation: that no two false theories could stand in the required verisimilitude relation. However, one of us showed that the limitative proofs depended on the assumption of classical logic; and depended essentially so, in the sense that the limitative result fails for the theories of all the usual relevant logics (Mortensen *1978*). Thus it looked for a time as if escape for verisimilitude might be quite easy: simply flee to one of the independently motivated relevant logics. Before general rejoicing breaks out, however, we should warn the reader that in a second paper it was proved that even for quite weak logics (and any stronger logic), the Popperian account for verisimilitude is subject to a slightly weaker, but still intolerably severe, limitation (Mortensen *1983*). The moral to draw here is that there is a general need to examine the extent to which one's philosophical theories are invariant over broad classes of logics weaker than or incomparable with classical logic, and a need to reassess them in the name of overall theoretical neatness if not.

The further moral to draw from these considerations is that one should seriously consider believing in as few necessary truths as one feels one can get away with, or: do not multiply necessity beyond necessity. This is in line with the 'Principle of Conceptual Economy': postulate as few conceptual connections as possible. (Popper himself warned that one should have as little to do with conceptual questions as possible; see *1973*, 310.) Like other principles of economy and simplicity, it is easier to appreciate the above principle as evidently true than it is to justify it. However, in its defence it can be said that it unifies the earlier points we have made. Conceptual connections *bind* in a way that contingent connections do not. They restrict theoretical freedom. Conversely, arguments, especially criticisms, which do not rely on unnecessarily binding principles, are more telling in that they derive from weaker premisses. If one's method does not otherwise require a logical or conceptual connection, then to saddle oneself with it is to allow oneself less freedom than one is entitled to.

We regard the foregoing case for logical weakness as compelling. It rests in part on the perception of logic as strongly continuous with science, particularly in its epistemological aspect. In the remainder of the paper, we wish to caution about how easy it is to slide back into the view that logic is distinct from science.

3. Necessity is More than Truth

As noted earlier, Haack argues that Popper must be presupposing the distinctness of logic from science, else no sense is to be made of his failure to be even-handed with them. But Haack in another place goes to considerable pains to emphasise the compatibility between the thesis that there is at least one necessary truth (call it necessitarianism); and the theses that logic and necessary truth are revisable and/or fallible (*1979*, especially pp. 60–1). Indeed, she makes it clear that she would like to hold both necessitarianism and fallibilism.

Now at this point there would seem to be two ways for a fallibilist–necessitarian to go: either hold that logic is revisable while enjoying a distinctive epistemology in which, for example, *a priori* proof plays a central role; or hold that logic is more literally a 'part of science', in which logic might be revised 'to save a physical theory'. The former of these has the difficulty of explaining in what its truth consists and how it could be of any use to physical theory,[2] but in any case it is also a strong distinctness thesis. Lakatos's *Proofs and Refutations* might (just) be read as consistent with the former. Haack favours the latter way (the quoted words are from her *1979*, 60), as we do.

But now, if logic is revisable in the light of physical theory, presumably the conditions for its rational acceptance are of the same kind, normal scientific theoretical investigation.[3] But then we need to ask how a scientific theory is to issue in

2 On this and other points see Mortensen *1988*.

3 Haack cautions against *easy* revision of logic because of its generality (*1974*, 37). One would have thought that this was *more* reason to be sceptical of it: unless one had in mind that it *must* be replaced with equally general theses, which looks dangerously like failure of even-handedness.

a recommendation to believe not only that a proposition is true, but *that* it is necessary truth? This is where we think Haack's desire to retain necessitarianism issues in a strong distinctness thesis in spite of her. The necessitarian here presumably is saying not only that the proposition in question, say of the form $A \vee \neg A$, is playing a central role in successful theory, but *also* that the proposition $\Box(A \vee \neg A)$ is too. Yet that is a distinct thesis and thus requires a distinct theoretical role. There would then be a niche for logic as the outcome of the study of distinctively logical truths or at least necessary truths, even if what is necessary is to be uncovered by normal scientific methods and revisable in the light of them. But, we suggest, it is a niche into which there would be no theoretical point in placing anything. The use of phrases like 'necessary truth' or 'logical truth' can conceal this distinction, since it might look as if ordinary investigation of the truth of necessary truths can verify them, and thus show them to be necessary. In a similar vein, when Grattan-Guinness speaks of bivalent logic as a 'refutable theory of applied logic', (*1986*, 193) there is a crucial ambiguity: the words 'applied logic' look like they might contrast with 'pure logic' and thereby imply a distinct subject, logic, presumably characterised by a set of propositions of the form $\Box A$ which applied methods can uncover. Of course, the intention here might not to be to make useless conjectures about $\Box A$, but to concentrate on ordinary scientific methods for uncovering the truth or falsity of A; in which case we applaud it.

We do not mean that we think that it is nonsense to speak of necessary truth and logical truth. Nor do we dispute the claim of the consistency between fallibilism and necessitarianism. All we wish to draw attention to is the extra task one would have to accomplish in establishing the necessary truth of any proposition, over and above its mere truth, and to question whether ordinary scientific methods would be equal to the task.

Inevitably the discussion has moved from the question of logical weakness to the question of whether any propositions are necessary at all, since they share many aspects particularly epistemological aspects. We do not wish to take the latter any further, having discussed it elsewhere (Mortensen *1988*). However, we do wish to point out that the difficulty of finding a reason to believe that any candidate for logical truth really is logical as well as true, is shared by all candidates. The moral is, perhaps, that the principle of logical weakness is an epistemological slippery slope leading to the zero option of no logical truths. We do not view this with alarm, but that is another story.

Bibliography

Dummett, M. *1978* 'Is logic empirical?' in *Truth and other enigmas*, London (Duckworth), 269–289.

Grattan-Guinness, I. *1986* 'What do theories talk about? A critique of Popperian fallibilism, with especial reference to ontology', *Fundamenta Scientiae*, **7**, 177–221.

Haack, S. *1974 Deviant logic*, London (Cambridge University Press).

—— *1979* 'Fallibilism and necessity', *Synthese*, **41**, 37–63.

Mortensen, C. E. *1978* 'A theorem on verisimilitude', *Bulletin of the section of logic of the Polish Academy of Science*, **7**, 34–43.

—— *1983* 'Relevance and verisimilitude', *Synthese*, **55**, 353–364.

—— *1988* 'Anything is possible', *Erkenntnis* (to appear).

Mortensen, C. E. and Priest, G. *1981* 'The truth teller paradox', *Logique et analyse*, **95–96**, 381–388.

Popper, K. *1973* 'A realistic view of physics, logic and history', in *Objective knowledge*, Oxford (Clarendon Press), 285–318.

—— *1963 Conjectures and refutations*, London (Routledge and Kegan Paul).

Routley, R. *et alii 1982 Relevant logics and their rivals*, Atascadero (Ridgeview).

Tennant, N. *1985* 'Minimal logic is adequate for Popperian science', *British journal for the philosophy of science*, **36**, 325–329.

PERGAMON

Language & Communication 22 (2002) 301–311

LANGUAGE & COMMUNICATION

www.elsevier.com/locate/langcom

Paradoxes inside and outside language

Chris Mortensen*

Department of Philosophy, The University of Adelaide, Adelaide, SA 5005, Australia

Abstract

We consider the question of whether paradoxes are *essentially verbal.* It is argued that paradoxes have an essentially verbal component, but that there is a special class of paradoxes, such as the "impossible triangle", which are not *wholly* verbal, but partly perceptual as well. Along the way, attention is paid to explicating the idea of non-verbal *content.* © 2002 Published by Elsevier Science Ltd.

Keywords: Paradox; Visual; Language

1. Introduction

A very common theme within analytical philosophy has been that philosophy is essentially second order, i.e. about language. The great Carnap proposed that in order to solve philosophical problems, it was necessary to re-state them as second order problems. The solutions were there to be seen as discoveries about the language we use. In ordinary language philosophy, Wittgenstein and Austin both diagnosed the mis-use of language as the source of philosophical perplexity, and offered the analysis of language as the method of philosophy. Even when this was successfully challenged by Quine and Smart, who emphasised the continuity between philosophy and science, insights about language remained central to their methods, as witnessed by the title of Quine's (1960) masterwork *Word and Object.*

One bona fide traditional area of philosophy has been that of the paradoxes: The Liar, The Sorites, Grelling's, Russell's, the Unexpected Examination and the like. These have generally seemed to arise from language, at least in that the proposed solutions and debates have generally been about the conditions for language, and how it relates to extra-linguistic reality. We consider a couple of these examples below. Graham Priest (1995) has recently proposed a general schema for paradoxes, the Inclosure Schema. Paradoxes are postulated to arise from two opposed tendencies

* Corresponding author. Tel.: +61-8-8303-4455.
E-mail address: chris.mortensen@adelaide.edu.au (C. Mortensen).

0271-5309/02/$ - see front matter
PII: S0271-5309(02)00009-5

at the limits of thought: first, the tendency to describe limits to thought, that is, necessary and sufficient conditions for something to be thinkable; and second, the tendency to go beyond those limits (the failure of the necessary conditions is thinkable). This at least suggests a thesis to the effect that all paradoxes are linguistic in origin. Solutions will thus involve facts about language rather than facts about the extra-linguistic world.

This speculation is re-inforced by the observation that the mark of a contradiction is the assertion of *p* and *not-p* for some *p*. But of course "not" is a familiar word in natural language. Naturally, the *p* and *not-p* need to be *asserted*, not merely present. It is trivial to *write down* "p and not-p". Indeed, mere "assertion" isn't enough either, as we will see below: a paradox arises because there are *plausible arguments* for the *p* and the *not-p*, where it isn't easy to see what has gone wrong.

Not that non-verbal means of representation could not figure in an argument. Barwise and Allwein (1996) have convincingly demonstrated the presence of reasoning which employs geometrical transformations directly without translation into any other format such as natural language. This should come as no surprise to anyone save the die-hard verbalist about philosophy. Reasoning is not *wholly reducible* to verbal reasoning, such as is represented by natural language or first order logic.

This observation suggests the further speculation that while paradox needs words for its demonstration, there may be paradoxes where other modes of representation and reasoning are employed essentially. Solutions will thus have to take into account non-verbal representation and contents too. That will be the theme of this paper. It is proposed to discuss a number of paradoxes involving sensory modalities. It will be argued that once we accept the need for non-verbal contents, then the way is clear to see that some paradoxes force us to reach beyond the analysis of natural language for their statement. This is not to say that there are ready solutions. If one thing remains, it is the highly puzzling nature of these paradoxes. What one can say about them falls far short of being satisfying. But at least we have a framework for a kind of solution, whatever would ultimately count as a "solution" here.

2. Paradoxes in language

There have been many curiosities which have been dubbed paradoxes. Two of the most important for the philosophy of language are both attributed to the ancient Greek genius Eubulides.

First there is The Liar paradox. It seems that Eubulides described Epimenides the Cretan, who asserted that all Cretans are liars. Its modern stripped-back version invites us to consider the sentence "This sentence is false". If it is true then in light of what it says about itself, it is false. Hence it is false. But then, again in light of what it says about itself, it is true as well. What makes it paradoxical is that on the face of it, it proves a true contradiction, that the above sentence is both true and false. It we alter the paradox to "This sentence is not true" it appears to yield the conclusion that the latter sentence is both true and not true. I say "appears" though in fact it is my inclination to say that it is both true and untrue. But there is a considerable dialectical gap between the appearance and the reality, and many have thought a different solution to be preferable. The important point here is that the difficulty of determining just what is wrong with the argument for the contradictory conclusion is a necessary part of what makes it worth describing as paradoxical. It isn't enough to have a paradox that we have an argument to a contradictory conclusion, even a valid argument for that conclusion. Such arguments are simple to construct. To be a paradox, contradiction must be *threatened*: it must be that different suggested solutions to the paradox aren't easy to choose between, they all have drawbacks and advantages in one way or another.

I want to describe The Liar as a pure paradox of language. It is here *expressed in* the English Language, using just the 26 letters of the English alphabet plus punctuation marks. But more than that, its *content* is exclusively *about* language: apart from the syncategorematic words "this" and "is", its words are only "sentence" and "false" or "not true", which respectively denote a syntactic item of language and semantic properties of it. It needs nothing extra-linguistic to get going. Mind you, as Kripke (1975) pointed out, we should beware of attempting to avoid the paradox by reference to semantics, such as by banning self-reference, if only because there exist closely related versions of The Liar which need contingent extra-linguistic facts before being paradoxical. Thus if I write on one side of a piece of paper "The sentence on the other side of this paper is true", then it is accidental whether the sentence on the other side of the paper happens to be "The sentence on the other side of this paper is false." Notice too, that this needs to actually be performed before the paradox arises. If no-one ever follows my instructions then there is no sentence or pair of sentences to threaten to be true and false. After all, unlike the purer forms of The Liar, I didn't actually construct the paradoxical sentence in describing it. On the other hand they are clearly variants of the same paradox, both in some sense involving circular reference leading to ungrounded proof of both truth values.

Then there is the Sorites paradox. A typical version of The Sorites argues from two propositions: (1) Anyone having 100,000 hairs on the head is hairy, and (2) (For all n)(if having n hairs on the head is hairy then having $n-1$ hairs on the head is hairy) to the conclusion that anyone having 0 hairs on the head is hairy (or if an exception is made at the last step, replace 0 by 1). The conclusion is false: by observation (or maybe linguistic convention) having 0 or 1 hairs on the head is not hairy but bald. What makes it a paradox is that it *isn't easy to see what is wrong with the argument* for that conclusion. The premises look true, and the reasoning looks valid. So difficult does it prove to say just what is wrong with the premises, that it becomes tempting to try to revise the logic. That alone would qualify it for philosophical importance. But logical revision is no easier, it appears. We can re-represent The Sorites as an argument for a contradictory conclusion: The Sorites seems to show that bald is hairy, but by observation (or linguistic convention or whatever) bald is not hairy. Put like this, it still isn't easy to see what to say about The Sorites, except that for sure we know which side of the contradictory pair of statements is true: bald isn't hairy.

There are of course many alternative versions. All versions, however, involve somehow the consequences of having *vague language* describe a *precise world*. So

The Sorites isn't quite a pure paradox of language. Matching and mismatching words with the world involves the world. Even more, its content is not exclusively *about* language, unlike The Liar. Still, it is a paradox of language in the sense that, like The Liar, it is *stated entirely by means of words*. Many other traditional paradoxes, including Zeno's, the Unexpected Examination and Grelling's, share this feature. But not all, I will argue.

Of course, paradoxes could hardly avoid essential use of language*in part*. After all, demonstrating a paradox cannot but involve use of both p and not-p, for some p. The word "not" is a natural language word, it is a sentence operator. Thus, if the mark of any paradox is the demonstration of a contradiction, then the statement of a paradox must be in language. Here, however, the word "language" needs extending somewhat, because "not" is a word of other languages than natural languages. It is a word in first order logic, for example. Thus it would not be so surprising if paradoxes could be demonstrated in various formal languages. And so it proves to be, not only formal versions of semantic paradoxes such as The Liar and Grelling's, but also set-theoretic paradoxes such as Russell's and Curry's.

Even so, there are various paradoxes which I argue need resources beyond the linguistic for their statement. As a preliminary, we need to ask how this could even be. How could it be that a paradox, *qua* collection of propositions or contents, could even *possibly* require for its statement resources beyond those of the typical natural languages, augmented with the languages of logic and mathematics? We turn to this next.

3. Non-verbal representation

It is hardly contentious these days that the human cognitive apparatus employs more than one representation system. I don't here mean the difference between *spoken* natural language and *written* natural language. These are alternative sign systems for expressing the same propositions (though of course we should leave open the possibility that there be *some* propositional differences between spoken and written English). Rather, I mean to allude to the difference between a natural language description of a person, and a police identikit picture or even better a photograph. The latter are obviously far more useful to us in identifying people. Again, there is all the difference in the world between a musical score of the first four notes on Beethoven's Fifth, and those notes played. Information is stored in us in multiple modules, with only limited cross-modular communication.

Elsewhere (Mortensen, 1989) I called these beliefs, desires, intentions and the like that utilise modes of representation beyond those of the words of natural language, "non-verbal propositional attitudes". Note that the contrast with verbal beliefs and desires is in the first instance in the *mode of representation*, rather than in the *content* of the attitude. This thus leaves open that there can be cross-modular equivalents in contents. The example of the word "red" versus a red colour chip (used to represent the colour of paints) indicates that there are *some* cross-modular semantic connections. But even here, something stronger is suggested: that the *content* of the belief that her face looked like *that*, is *radically different* from that of a verbal description. There is *so much more* information in the former, my behaviour is *so much more fine grained* if the information is stored in me in the former way. Rough equivalents in content perhaps, but identity of contents no.

Beyond verbal contents, then, we have the various non-verbal contents. In this classification scheme, some linguistic information counts as non-verbal, the best examples being formal logic and more generally mathematical text. Note in passing the predominantly *written* character of mathematical text: a *wholly written* mathematical lecture is a commonplace (any journal article), whereas a *wholly spoken* mathematical lecture is virtually unthinkable. There is a significant difference here with philosophy, where both wholly written and wholly spoken presentations of the same lecture are commonplace.

There is more to non-verbal content, however. Specifically, beyond natural or symbolic language there is the non-linguistic. The shape of the face is *geometrical*, for example. That is why Euclid had to draw pictures or diagrams: because the *subject matter* of geometry is space or shape. Imagine Euclid's *Elements* without the pictures. The meanings of the propositions and the proofs all of a sudden become highly non-obvious, to say the least. The point is lost, as it were. We must resist here the temptation to commit the error that 3-D geometry is about R3, the set of triples of real numbers discovered for us by Descartes. Descartes certainly discovered a way to describe geometry by algebra, which greatly extended the proofs available. But the two aren't the same thing. In geometrical figure-drawing we represent by *exemplification*, to use Nelson Goodman's (1981) word. Words (rarely) exemplify what they represent, pictures always do, at some level of generality at least. In exemplifying, pictures show without saying, as Wittgenstein might put it. This further strengthens the link between the mode of representation and the content: if a geometrical shape is part of the content of a belief, then mere words do not have the same contents; they represent without exemplifying. That words represent without exemplifying is an important representational breakthrough of course: it is what makes possible cross-modular unification of information. I can use words to represent both the redness of the cloth and the shrillness of the sounds, but I cannot use an example of red to represent shrillness by exemplification, nor vice versa.

One more point is that exemplification doesn't have to be *perfect* to represent successfully. Exemplification comes in degrees of fit, and perfect exemplification is an ideal which actual representations might accidently achieve. Pieces of space, among which are drawings, certainly have their own exact shapes. But identikit pictures are not exact likenesses, nor do the printed triangles in Euclid have exactly straight sides. Still, they are not too far off; closer than more distorted pictures for example. Very wiggly lines are generally much less satisfactory to represent the sides of a triangle. But there is still exemplification here: the printed straight line exemplifies the *vague* concept of *straight or near enough to* (or not quite straight, as the artist Jeffrey Smart, 1996 would say). And there are three of these nearly straight lines in the representation of a triangle, not 4 or 44. On the other hand, some modes of exemplification are more exact than others: it is hard to see a crimson patch on a colour card as anything but red. More on colours later.

One point to make clear here is that it is hardly being claimed that pictorial representation is somehow *word free*. I'm not even sure that there's a coherent thesis here, but it is certainly not being claimed that shapes represent without being embedded in a linguistic context. Dennett (1969, pp. 132–141) wanted to argue that a shape *by itself* didn't represent another shape, it needed some representing conventions. Euclid did not draw *only* pictures, nor could he have. Euclid's *propositions and proofs* used words. When I believe that her face looks like *that*, words are used to express *part* of that, words like "her", "face" etc.

Once we have contents, linguistic or not, we have the capacity for accuracy and error, truth values in short. I can believe falsely about her face, or lie with a drawing or a photograph. We thus need a truth predicate, governed by an analog of the T-schema. At one level, this is trivial: "The shape of the object is: ◀" is true iff the shape of the object is: ◀. More importantly, we also have the capacity for non-linguistic processing. Barwise and Allwein (1996) has convincingly demonstrated problem solving which essentially uses geometrical transformations. Hence we have the capacity for a semantics which reflects the non-verbal nature of various contents as well as the connections between different modes of representation. For example, we can say that the word "triangle" can refer to a drawn figure, just as we can say that the drawn figure can represent any triangle by exemplification. This enables us to explain the natural conclusion that since the drawn figure represents by exemplification, the word "triangle" can represent any triangle. Truth conditions also emerge: "The shape of the figure is a triangle" is true if the shape of the figure is thus: ◀. The "only if" clause of this is more problematic, but even the "if" clause confers the power of non-verbal modes of representation to confirm linguistic *Protokolsatze*. An observation of a figure is sufficient to confirm the truth of a piece of language: that the shape of an object's being: ◀ entails that the shape of the object is a triangle and so that "the shape of the object is a triangle" is true.

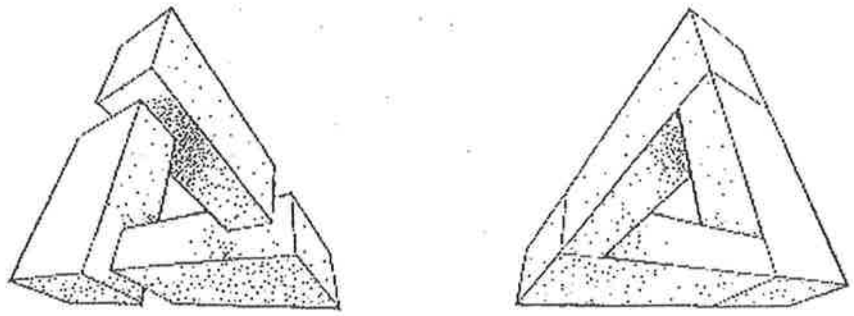
It is useful to mark this distinction with a definition. We can call the non-linguistic content of a proposition a *percept*. Specifically, percepts are contents which arise from the perceptual apparatus. It is not decided that all percepts represent by exemplification, but some do. In consequence, percepts are part of the cognitive apparatus which utilises representation by means of exemplification. Percepts can figure in deductions, at least geometrical transformations. They also figure in the justification of linguistic *Protokolsatze*.

With this somewhat lengthy preamble, we turn back to the theme of paradoxes.

4. Visual paradoxes

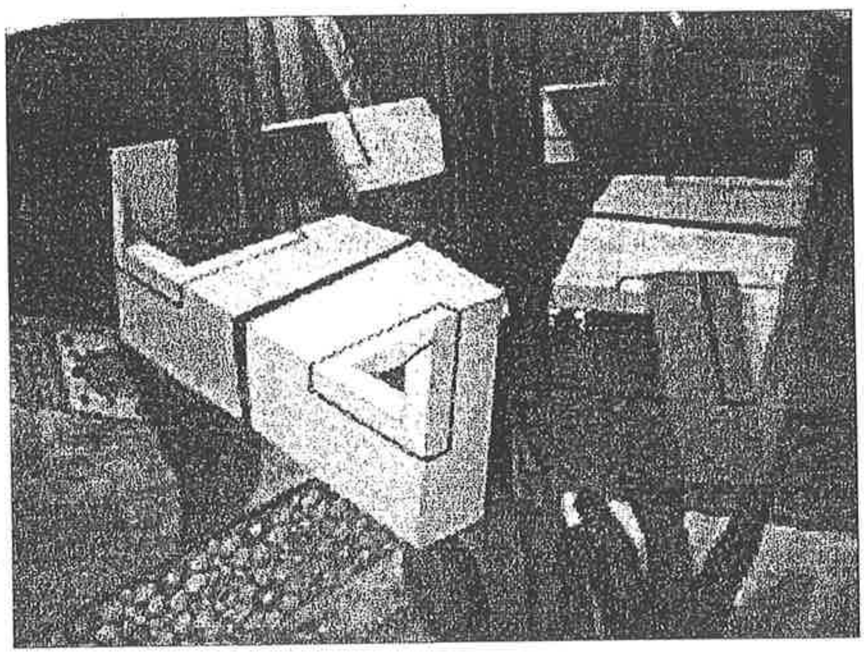
Think first about a case which I maintain is not paradoxical: perspective. Perspective is not a paradoxical percept. According to the percept, the parallel railway lines meet, and according to the camera too. What we know about the lines is that they never meet, of course. But that knowledge does not form part of the percept, it is not "projected onto" the percept so that there appears to be an impossible situation in front of our eyes. In passing, there are ways in which we represent parallelism other than having the lines meet at an horizon. We draw finite lines looking parallel, and then add arrows to indicate that the lines when extended infinitely never meet. This device is partly an exemplification (the finite lines do not meet as drawn) and partly conventional non-exemplification (neither the lines nor the arrow heads are infinitely extended).

Our subject matter here is different, however. Consider these images:



These *look* impossible. Unlike perspective, we have here paradoxical *percepts*. Qua percepts, moreover, the paradoxes are *not wholly linguistic*, in one important sense. The paradox cannot be presented entirely with words, you have to *look*. Try saying the words: "a three-sided object where each of the three sides recedes clockwise away from the viewer in a closed loop". Even if you could understand this description at all, it fails miserably to convey the oddity of the experience, which has to be seen to be believed.

To reinforce the point, consider the following photograph.



You can see from the two reflections that the object has an arm out the front, not joining up to form a closed loop. It only photographs as a closed loop from a single special direction, when the visual brain "joins up" the arm. When you know what it's a photo of, it can be *seen as* a possible object, with one arm sticking out the front. Even then, the paradoxical aspect doesn't go away. The image can still be seen as an impossible object, indeed that remains the natural way to see it. Elsewhere (Mortensen, 1997) I argued that the right way to think of the phenomenon is that the cognitive apparatus, specifically the visual module, projects onto the percept its expectation that points with a very small angular separation are spatially very close. Doubtless this is an evolutionarily useful thing to do. This results in the percept being describable by an *inconsistent theory*. Only an internalised inconsistency can account for our feeling that we see it but don't believe our eyes, I contend. Note too that the phenomenon here is one of perspective (one eye) rather than of stereoscopy (two eyes). The depth vision that comes from stereoscopy can serve to disambiguate images, but that is not available when we are looking at an image on a flat page.

In what sense do we have a paradox here? The impossibility in the image is unavoidably *partly verbal*, I think. If you try to describe what is impossible with it, you will come up with something like the following. "There are a series of points on the object, p1, p2, p3,. . .pn, such that p2 is further away that p1, p3 further away than p2, etc all the way around the figure so that p*n*−1 is further away than pn. The relation *is further away than* is transitive, so pn is further away than p1. But by inspection pn is closer than p1, and thus not further away than p1." This is an explicit contradiction, unquestionably. I cannot see otherwise how one would represent the oddity here. If you try to just gaze at the picture without having verbal thoughts *per impossible*, it just looks a bit odd. I think it takes the *words* to appreciate just how odd it is, its precise oddness. Of course, *the reality is not the words*. The words are *about* the geometrical reality. That is where representation with and without exemplification have their distinctive differences, both strengths and weaknesses. But what is odd about it needs negation to be expressed, and negation is a linguistic operator. Denial took language to be possible.

So we have here the necessary presence of the non-linguistic and the linguistic entwined together for this paradox to be grasped. It goes without saying that there are a very large number of other visual paradoxes that have been constructed, particularly in the twentieth century, starting with the great Oscar Reutersvaard in 1934. It is apparent that there are different classes of geometrical paradoxes to be described, with different kinds of contradictions. Instead of pursuing this line, however, I want to draw attention to non-verbal paradoxes in different modalities other than the geometrical.

5. Other non-verbal paradoxes

One familiar colour paradox is the colour sorites: the non-transitivity of the relation *looks the same colour as*. This is a paradox which is visual though not geometrical. We don't have much trouble with this familiar phenomenon. Perhaps we should. The air of paradoxicality comes from the use of the word "same", in "looks the same". What explains our inclination to use "same"? That's a transitive relation if ever there was one. Now the paradox tends to go away, if one uses instead the relation *cannot tell the difference between*. After all, there seems to be no a priori semantic reason why that should be transitive. However, that does not tell the whole story, since one can fail to tell the difference between things in more than one way. One can be totally blind, for example. Colour experience has a positive content, not mere negativity. Like the sorites elsewhere, this paradox amounts to the problem of matching a fuzzy concept of sameness onto the positive aspects of the world. The problem becomes sharper if we postulate necessarily conscious mental items. That is partly why traditional sense data theories lost favour in epistemology: if the reality exactly matched the conscious experience in its distinctions, then how could *looks the same colour as* fail to be transitive?

I think this phenomenon shows something interesting: that colour perception is partly a *default* mechanism. The colour brain adds a verbal judgement "same" to experiences that it cannot tell apart. But they must be different as percepts because they interact differently with a third party: the third colour sample. The non-verbal *relations between the percepts* are different, even though we cannot tell them apart by a *direct comparison*. This, I take it, is the verbal aspect of the paradox. As with the geometrical shapes, it looks to be unavoidable that words are essential in order to express the paradox. Thus we have here another example of the mixing of language and non-linguistic aspects to produce a paradox.

The final paradox is auditory: the ever-rising note. If you haven't ever heard it, I can assure you that it sounds just like that. How is it done? I don't know, but a reasonable conjecture is that it is a series of rising dominant tones which fade in, are at their loudest half way up, then fade out again. The perception of risingness is attentional perhaps: the attention fastens onto the rising tone and follows it up, shutting out other tones which are fading in and out. The oddness is quite primitive here, and arguably pre-verbal: there is a perception of constant change, a perception or memory of non-cyclicity, and at the same time a memory that later percepts are the same as earlier.

Even so, this paradox isn't wholly non-verbal, I would say. The oddity has to be said to be grasped: "a sound which sounds like it is always rising; while at the same time it isn't rising, because memory says that it's the same as earlier, and reason tells us that it can't be both changing (rising) and unchanging." On the other hand, as with all the cases we have been looking at here, it is hardly entirely verbal, if only because you'd be quite within your rights to doubt this verbal description if that's all you had. Only the experience can really justify it.

6. Conclusion

I am conscious of a profound feeling of dissatisfaction at how little one seems to be able to say in resolution of these puzzles. Still, there are a few general conclusions which follow from our discussion.

These examples all raise the question of how the verbal and the non-verbal are sewn together in our mental lives. This is a question which arises prior to paradoxes, when we reflect on the twin facts: (1) that we are thoroughly verbal creatures, words fill our mental social lives and define us as humans embedded in civilisation; and (2) that we are thoroughly bodily creatures, whose mental lives are saturated with the sensory. We thus want to know how it is that the verbal and the sensory interact. That they manage to interact is the power of words: they enable cross-modal communication and evaluation to take place. To make an unfashionable claim, words are what give us a CPU to integrate our actions, instead of our being creatures driven by a collection of modularised, non-integrated parallel sensory processes. It is plain that this could only be achieved by a mode of representation which was non-exemplifying. This problem becomes sharper, however, when paradoxes present themselves which trade on the interaction between words and the sensory.

Of course, a contradiction by itself isn't particularly interesting. For example, recall the old story of the Barber of Seville, who shaves all and only those in Seville who don't shave themselves. A simple argument takes one to the conclusion that the Barber shaves themself if and only if they don't. But this is no news: there simply is no such barber. They don't exist, it is a contradiction to say that they do. To be a serious paradox, as I said earlier, a *true* contradiction must be *threatened*: it must be at a minimum *difficult* to see what is wrong with the arguments for the opposing sides. So, for example, Russell's Paradox, which has an identical form to the Barber, is much more troublesome. Russell's Paradox describes the Russell Set, the set of all sets which are not members of themselves. In parallel with the Barber, the Russell Set is a member of itself iff it isn't. The reason why it is more troublesome, is because we have independent reasons to accept a natural principle which yields its existence: the principle that to every description (such as "red" or "non-self-membered") there is a corresponding set. This is why the presence of the non-verbal gives the perceptual paradoxes force: one side of the paradox is *manifest* in the percept. But how is that managed? Beyond the foregoing tentative remarks, I don't know.

Another conclusion is that these paradoxes escape Graham Priest's (1995) ingenious Inclosure Schema. As we noted at the beginning, Priest's account of paradox proposes the clash of two general principles, one which seeks to impose limits on thinkability or expressibility, and the other which seeks to go beyond, to burst outside the conceptual barrier. I find this a very plausible account of many of the thinkers described in Priest's book, though the strength of the paradox generated varies with the case: again we need to find the premises plausible before we have grounds to suspect a paradox. But the present non-verbal examples don't seem to fit at all. They are highly specific to sensory modalities, they do not seem to be about limits, and they are not at all about general conditions for conceivability and the like.

One suggestion put to me is that geometrical paradoxes all involve a clash between the local and the global. Locally we see that each point is further away than the next. In any local part, the diagram is consistent. But on the other hand, the diagram as a global *gestalt* is impossible. This seems right, but I fear it is more to describe the problem than to solve it.

References

Barwise, J., Allwein, G., 1996. Logical Reasoning with Diagrams. Oxford University Press, New York.

Dennett, D.C., 1969. Content and Consciousness. Routledge & Kegan Paul, London.

Goodman, N., 1981. Languages of Art (2nd ed). The Harvester Press, Brighton.

Kripke, S., 1975. Outline of a theory of truth. The Journal of Philosophy 72, 690–716.

Mortensen, C., 1989. Mental images: should cognitive science learn from neurophysiology? In: Slezak, P., Albury, W. (Eds.), Computers, Brains and Minds. Kluwer Academic Publishers, Dordrecht, pp. 123–136.

Mortensen, C., 1997. Peeking at the Impossible. Notre Dame Journal of Formal Logic 38 (4), 527–534.

Priest, G., 1995. Beyond the Limits of Thought. Cambridge University Press, Cambridge.

Quine, W.v.O., 1960. Word and Object. MIT Press, Cambridge, MA.

Smart, J., 1996. Not Quite Straight: A Memoir. Heinemann, Melbourne.

IT ISN'T SO, BUT COULD IT BE?

Chris Mortensen
Dept of Philosophy
The University of Adelaide
North Tce, SA 5005
Australia
Chris.Mortensen@adelaide.edu.au

Abstract: In his paper "Could Everything Be True?", Graham Priest argued against *trivialism*, the thesis that everything is true. Priest was aiming to show that it is not so easy to dismiss trivialism, but that in the end it fails. This paper focuses on a different but related thesis, namely that trivialism is *possible*. The possibility of trivialism is indicated by a more general thesis, namely that *anything is possible*, which is known as *possibilism*. Some of Priest's arguments indicate that he takes his arguments to refute the latter claim as well. This paper begins by surveying the advantages of possibilism. It then turns to argue that Priest's arguments fail against possibilism, and that trivialism, along with everything else, is possible. Finally, the explanatory advantages for modal semantics are briefly sketched.

Logique et Analyse, forthcoming (2006)

It Isn't So, But Could It Be?

1. Introduction

This paper contributes to a defence of the thesis of *possibilism*. Possibilism is the thesis that *anything is possible*. Possibilism is in opposition to the thesis of *necessitarianism*, namely that *there is at least one necessary truth*. Possibilism was named and defended first by Naess (1972), and later by Mortensen (1989). The present paper discusses the bearing on possibilism and necessitarianism of arguments due to Graham Priest (2000). Priest's arguments were aimed *prima facie* at a different thesis, the thesis that *everything is true*, which he called *trivialism*. However, as we will see, possibilism and trivialism are closely connected, especially given Priest's way of framing his attack on trivialism. Thus the present paper principally aims to identify Priest's arguments against trivialism, show how they represent a threat to possibilism, and demonstrate that they are unsuccessful. Before coming to that, however, it will be necessary to survey existing arguments in favour of possibilism, to establish its initial plausibility. In the final section of the paper, it will be seen that this perspective leads to a certain simplification of the semantics of non-normal modal logics.

2. Necessitarianism and Possibilism

The arguments for possibilism turn on a systematic attack on the opposition view, necessitarianism. The concept of "necessary" which applies here is somewhat loosely characterised as a group of notions around the ideas of logical necessity, metaphysical necessity, model-theoretic necessity, analyticity, and the like. It is not claimed here that these notions are all reducible to a single core. It is also not being claimed, as Quine would, that there are no coherent concepts at all in the vicinity. Rather it is proposed that various arguments against these notions have varying weights, depending on where they are directed. All the same, it is contended that they all have epistemological drawbacks.

The attack against necessitarianism comes from two broad directions: epistemological and ontological. Of these, the former carries the greater weight. We first survey the epistemological problem, then rehearse ontological arguments.

Arne Naess (1972) seems to have derived his view in turn from Popper's (1963) attack on what Popper called "conventionalism". According to Popper, conventionalism is the *practice* or *strategy* of defending a theory come what may against contrary empirical evidence or strong counter-arguments. Popper diagnosed the error as placing *conceptual restrictions* on theory-revision. Against this strategy, Popper and Naess argued that even concepts may need to be revised, and that a criticism which proceeds from a weaker conceptual base, one with fewer restrictions, is stronger since there are fewer ways to escape from it. Conversely, a criticism proceeding from unnecessarily strong conceptual or principles is actually weaker, in that it is easier to find places to reject. Now it is well known that when it came down to it, Popper was willing to exempt the principles of logic from his rejection of conventionalism. Similar to Quine, Popper ultimately found classical two-valued logic to be the correct logic. This introduces an *ad hoc* character into Popper's otherwise estimable methodological position. In contrast, Mortensen and Burgess (1989) argued that this was less than wholly general, and that a fallibilist like Popper ought to be saying that not even logic is exempt from revision; that is, that the set of specifically logical truths is null.

Here a cautionary note must be injected. Popper was a fallibilist, a distinguished tradition deriving from Peirce. Fallibilism is notoriously difficult to state, and this paper does not attempt to solve that particular puzzle; nonetheless it amounts to something like the claim that no theory is rationally unrevisable under the pressure of empirical science, that all theories can fail for good reasons. It would seem, then, that a fallibilist ought not to make an exception in the case of logic. However, Susan Haack (1979) argued persuasively that fallibilism must be regarded as compatible with necessitarianism. This is surely correct: after all, not even possibilism should be regarded as unrevisable. Thus, fallibilism should not claim to have the force of *apriori* disproof over necessitarianism.

But the epistemological argument against necessitarianism need not claim to have the force of *apriori* disproof, any more than any other scientific hypothesis. The argument is rather: Are there any reasons to believe the alternative? If not, possibilism has the virtue of the generality and economy of *epistemic monism.* This term, introduced in Mortensen (1989), refers to a wholly general method for establishing truths, namely the scientific method of empirical theory-choice using experiment, theory and observation. There is no need to cater for the knowledge-base of an entirely distinctive set of necessary truths. The problem here is not that necessary truths could not be shown to be true by ordinary scientific means, for they obviously can. The problem is how one would come to know *that they are necessary* (in any of the various senses of that term).

Perhaps this is to be done in the ordinary way of scientific theory-construction, by including a postulate of the form □A in a theory and appealing to its explanatory power? But it must be apparent at this point that, as Naess argues, *one adds nothing to the explanation of what hits our sensory surfaces, by putting a necessity box in front of any theoretical postulate*. "B because A" may be sensible, but "B because □A" gains nothing as an explanation, and even more obviously "A because □A" adds nothing also. We have no reason to use such statements in our theories. In sum, epistemic best practice indicates to us that we have no reason to believe in necessary truths, and the virtues of a uniform epistemic method are then overwhelmingly attractive in favour of possibilism and against necessitarianism.

So much for epistemology. But there are ontological currents as well. The main tide of ontological arguments against necessitarianism is the challenge to show how various accounts of the necessary could support a principled distinction between two fundamentally different kinds of truth, the necessary and the contingent; and to do it in such a way that the extension of each is non-null. In these waters, arguments tend to drift apart as different accounts of necessity are canvassed. So take for example the well-known empiricist reduction of necessity to *analyticity*, truth by meaning. This was a brilliant innovation in the theory of necessity, because it held out a plausible epistemology, namely knowledge of the meaning-conventions of words, which seems unproblematic or at least less problematic. But, as Hilary Putnam (1978) pointed out, there remains a gap: *how could it be that having a certain meaning would be enough to ensure truth, without the world playing a role?* That is not generally the way of it with a truth-making world. Indeed, if the world played no determinative role, what sense is there in describing it as *true?* As Putnam put it, you can't make something true by a convention unless it's already true. In passing, it should not be thought that these arguments depend on Quine's repudiation of the concept of analyticity: they are intended to apply to the *extension* of that concept without drawing its meaningfulness into question. In any case, as we have already noted, Quine retained for himself a core of logical truth, classical logic.

Or take a different account of necessity which has appealed to many, namely *model-theoretic necessitarianism.* Here the idea is that necessity is truth in all models (such as sets of consistent and complete worlds). Unfortunately, this does not survive long either, though for other reasons. It is too easy to construct models in which putative necessary truths fail. This must of course be accompanied by a survey of the numerous semantical studies which have produced counter-models, and this must be regarded as having considerable complexities when dealing with principles like the Law of Non-Contradiction ~(A&~A), let alone the Law of Propositional Identity A→A. These arguments are surveyed in Mortensen (1989). However, these arguments can be encapsulated briefly by noting that there is a general theorem covering all cases, due to Meyer-Routley (1977, 2004): *any sentential formula can be refuted in some two-valued model.* That is, model-theoretic necessitarianism must be accompanied by an argument to select out and privilege a distinguished subset of models, when it is conceded that the additional models exist. This is invariably not attempted. Again, the only fully general position is that which allows the widest class of models. But this yields the conclusion, not that there is no coherent concept of necessity here, but that its extension is zero.

These epistemological and ontological considerations are powerfully inclining, I suggest. But it must be conceded that the *intuitions* are ravaged by the denial of the necessity of such propositions as that *at least one thing is true,* or that *not everything is both true and false*, or simply that *not everything is true*. Something has to be done to pump up contrary intuitions, if anyone is to be persuaded. This brings me to the main topic of this paper.

3. Possibilism and Trivialism.

To recall, trivialism is the thesis that everything is true, so named by Graham Priest in "Could Everything Be True?" (2000). The name derives from the usual definition of a theory's being *trivial* if it contains every proposition, which is useful in disputes over the classical principle *Ex Contradictione Quodlibet* (from a contradiction everything can be deduced).

Priest characteristically sets himself to imagine the unimaginable, by taking trivialism seriously enough to need refuting. He aims to defend the thesis that not everything is true. Clearly, those of us who are not deranged agree that not everything is true. But it proves surprisingly difficult to justify that belief, as Priest ably demonstrates. Nevertheless, in the end the weight of argument is definitely favourable. We will review these arguments presently.

Given the main argument of Priest's paper, then, its *title* is misleading. For the title asks a different question: *is it possible that* everything is true, or perhaps is it *impossible?* Now of course philosophers sometimes ask whether something is so by asking whether it could be so: we're knee-jerk *apriorists* after all. And in this context few would be misled by Priest's title. Nonetheless, there are important issues under the surface here.

Possibilism has close connections with trivialism, in that if possibilism is true then it would seem that trivialism is possible (even if untrue). It might be thought that this is too much to conclude. After all, the truth of possibilism would seem to require only that there be, for each proposition, a world in which it is true. It is a further step to say that there is a single trivial world, one in which every proposition is true. The former might be called the "distributive" version of possibilism, and the latter the "collective" version. The distinction is conceivable enough, it relies on a traditional difference in two ways of taking the universal quantifier, which was applied in the analysis of the traditional fallacies of composition and division. Still, this objection can be sidestepped, I suggest. Trivialism would seem to be a meaningful position. That is one of the parameters of this discussion, as Priest would agree: it can be expressed in the logic of propositional quantifiers as $(\forall A)A$. Consequently, if *anything* (distributively) is possible, then trivialism is possible. That is, *everything* (collectively) is possible.

At any rate, whatever is right here, we can certainly say that if possibilism is true, then the answer to the *title* of Priest's paper is *yes*. If, on the other hand, Priest's arguments have the force of necessity, then one should conclude that what Priest is arguing against is not just false but impossible. If anything is impossible, then its negation is a necessary truth. That is, if possibilism is false then the answer to the *title* of Priest's paper is *no*.

Priest evidently takes at least some of his argumentation to have necessary force. For example, he writes:

"It is easy enough to show that trivialism is not true – indeed necessarily so. For it is either true or it is not. But if it is true, it follows that it is not true (everything follows). Hence, in either case, it is not true" (P190)

Priest comments that this would not show that there is something true which is rejected by the trivialist, because trivialism rejects nothing. Yet Priest himself is no trivialist, and he evidently regards this argument as successfully establishing the necessary falsehood, the impossibility, of trivialism. But does it?

Mortensen (1989) in defending possibilism maintained the possibility of trivialism. Any argument that trivialism is impossible will be either invalid or question-begging. Consider for example Priest's own argument. It is of the form "$A \vee \sim A$, $A \rightarrow \sim A$. Hence $\sim A$." Now suppose we try to strengthen the conclusion to: $\Box \sim A$. But that would not follow if the premisses as given are not necessary truths: the form "A∨B, A→B. Hence □B" is generally invalid if the premisses are contingently true. So one would at least need to strengthen one or both of the premisses to the stronger necessary form. But why should one accept that? It would be blatantly question-begging.

This observation does not establish by itself that trivialism is possible, nor that possibilism is true. The main aim at present is different, namely to consider Priest's arguments. He canvasses three arguments against trivialism. It is therefore worthwhile to consider how well they fare against possibilism.

Priest's *first* argument is as follows. Surely there are some propositions that we have to admit there are no good reasons to believe. Therefore, any trivialist would have to admit that there are parts of their position which there is no good reason to believe. Thus the belief in question, which the trivialist has because they believe everything, is irrational. But Priest allows that the trivialist can reply that there is some reason to believe any proposition, or at any rate many propositions. Consider any identity statement, such as that you are a scrambled egg. It is a familiar argument that by making small enough changes we do not change the character of a thing. So what began as you remains as you even when every molecule of your body is replaced by scrambled egg. By Leibniz Law, furthermore, it then follows that any thing has any property, since any thing is identical with something which incontestably has that property.

This is the familiar reasoning of the Sorites paradox. There must be something wrong with the Sorites, for it is contrary to observation. It is notoriously difficult to say just exactly what is wrong with the Sorites; but there must be something fallacious about it, or it would be that hairy is bald. Thus, any defence of trivialism which relies on the Sorites is unbelievable. But note that here Priest was trying out an attack on trivialism and finding it wanting. Thus, the conclusion of his argument was not intended as a refutation of trivialism. In point of fact, it is easy to agree with Priest's premiss that there are some propositions that we have no reason to believe. For example, there is the *phenomenological absence* of an observation or sensation. In the absence of a sensation, we have no reason to believe its *Protokolsatz*. But none of this threatens the thesis that triviality is *possible*, in any case. Even if the argument succeeded as an objection to triviality, the conclusion that our own world is not trivial does not begin to show what another might be.

Priest's *second* argument against trivialism is that it implies the meaninglessness of public language. Public meanings are learned, and learning implies contrast, some descriptions accepted and some rejected. But trivialism prevents rejection, since for the trivialist nothing is rejected.

It is clear that this argument does not succeed as an objection to possibilism. If trivialism is merely possible, then the meaninglessness of public language does not follow. If *our* world has contrasts, which it surely does, then it is our world in which the contrasts that fix meaning and learning in *this* language abound. If the failure of public communication is merely possible but not actual, then nothing follows about the inability to learn language in *our* world.

Priest's *third* argument is to the effect that it is phenomenologically impossible to believe that everything is true. This is because to live we have to make *choices*. Choices are goal-directed, they imply rejection of other alternatives. Since trivialism cannot accommodate rejection, there are no real trivialists.

Indeed so, but this does not even show that trivialism is untrue, as Priest acknowledges, let alone that possibilism is untrue:

"This does not show that trivialism is untrue. As far as the above considerations go, it is quite possible (sic.) that everything is the case; but not for me – or for any other person." (P194)

It might be that the fact that there are no trivialists counts against trivialism, but it surely does not count against possibilism. We make our choices, our actions and our rejections in our world, and this world is not trivial. That is quite compatible with another world being trivial.

The failure of these arguments as objections to possibilism illustrates a more general point. Defences of necessitarianism typically try to reduce possibilism to a contradiction. But all such arguments eventually fail, because *possibilism is a consistent position.* The simplest way to see this is to consider the matrices below, which extend classical logic with possibility and necessity operators.

&	T	F	~	◊	□
T	T	F	F	T	F
F	F	F	T	T	F

It is obvious that this is consistent if classical logic is. Hence, no argument that seeks to render possibilism to be a contradiction succeeds: the matrices tell us which premisses are false or question-begging. For example, Graham Nerlich argued in conversation that possibilism is committed to the possibility of necessitarianism. This is true, and indeed we have already registered the point in connection with Haack in Section 2. The matrices validate ◊□A for every proposition A. Necessitarianism is thus possible, it is a coherent position (or rather a group of positions for various accounts of the nature of necessity). But the matrices also show that possibilism is not thereby reduced to triviality or inconsistency.

There may be, of course, attempts to demonstrate that something is necessary other than by showing that possibilism implies a contradiction. I suggest, though, that they will all need a premiss of the form "Necessarily A" somewhere along the way, and then one would be inclined to wonder why this confers some explanatory advantage over A by itself. To illustrate this point, consider the objection raised in conversation by John Bigelow, namely that some account should be taken of our strong intuition that the truths of logic and mathematics are distinctively susceptible of *apriori* proof. I agree that this intuition as a mental state needs accounting for. But how would one progress the explanation beyond the usual causal explanation in terms of the occurrences of preceding mental states, by adding in a premiss that one of the causes of our mental state is necessary? How could necessary truth improve the explanation of any mental state, intuition or not?

4. Non-normal Worlds

The above matrices do not pop up out of nowhere. It is clear that they arise from the usual semantical assignment conditions for the modal connectives when applied to a model structure consisting of a single non-normal world. A single non-normal world may of course have all the so-called laws of logic holding true, such as $A \rightarrow A$ and so on, but no necessitated statement holds true. It should be noted in passing here that the issue of the reality of worlds is not at issue: it is not intended to take sides on modal realism versus various *ersatz* reductions of worlds.

This serves to deflect the objection that the matrices are cut loose from the meaning-constraints of alethic modality. It is plausible that when studying modal logic one identifies commonality in the concept under study (necessity, possibility, conjunction *etc.*) with commonality in the assignment conditions, so that variation in the worlds of the model structure represents varying accounts of the *same concept*. But that is exactly so here, the assignment conditions for possibility and necessity are the same as in the non-normal modal logics, only the case where there are normal worlds is unsatisfied and thus idle. For the same reason, it is pointless to object that the semantics of all normal modal logics validates the rule of necessitation: every theorem is necessary. Of course that is so, but the rule would need to be independently motivated. Needless to say, it is part and parcel of possibilism that the rule of necessitation fails.

There is one more point about the explanatory advantage for semantics in allowing a trivial world. The semantics of *non-normal* modal logics is anomalous in the way it treats non-normal worlds. Non-normal worlds are those (such as ours, if possibilism is right) at which all propositions of the form "Possibly A" hold. In standard modal semantics, this is regarded as *sui generis*, not arising from the accessibility relation in the way that other modal evaluations do. But that is *ad hoc*. Now there are two less *ad hoc* ways to produce or "explain" the above matrix. One could postulate a model structure in which our non-normal world had an infinite collection of accessible worlds, one for each proposition to hold in, so that $\Diamond A$ held on our world, for each A. This would correspond to the "distributive" sense of possibilism that we identified at the beginning of Section 3, each proposition would be possible but there would be no sense in which they were possible together. However, it is clear that a formally simpler way to improve things is to allow a single trivial world. Then any world from which the trivial world is accessible, is automatically a non-normal world. In addition to being technically much simpler, this would correspond to the stronger "collective" sense of possibilism which was adopted for preference in Section 3, and which was in accordance with Priest's own understanding of trivialism. Furthermore, the usual *ad hoc* assignment to non-normal worlds disappears in favour of the truth of all the ◊A being assured by the accessibility relation in the usual way.

5. *Conclusion*

We see, then, that possibilism resists Priest's arguments against trivialism, initially threatening though they might have seemed. We also see that possibilism has independent strengths. It is a consistent position, and there is no good reason to believe in its rival, necessitarianism. In positive terms, it is simple and plausibly motivated, being the only epistemically and ontologically general thesis in the field. Anything is possible, even triviality.

6.*Bibliography*

Haack, S, (1974), *Deviant Logic*, London, Cambridge University Press.

---, (1979), "Fallibilism and Necessity", *Synthese*, 41, 37-63

Meyer, R.K. and Routley, R (1977) "Extensional Reduction I", *Monist*, 60, 355-369.

---, (2004), "Extensional Reduction II", in R.Brady (ed) *Relevant Logics and Their Rivals Volume II,* Aldershot, Ashgate.

Mortensen, C, (1989) "Anything is Possible", *Erkenntnis*, 30, 319-337.

Mortensen, C, and Burgess, T, (1989) "On Logical Strength and Weakness", *History and Philosophy of Logic,* 10, (47-51)

Naess, A (1972) *The Pluralist and Possibilist Aspect of the Scientific Enterprise*, London, George Allen and Unwin.

Popper, K, (1963), *Conjectures and Refutations*, London, Routledge and Kegan Paul.

---, (1973), "A Realist View of Physics, Logic and History", in *Objective Knowledge*, Oxford, The Clarendon Press, 285-318.

Priest, G, (2000), "Could Everything Be True?", *Australasian Journal of Philosophy* 78 (189-195).

Putnam, H., (1978), "There Is At Least One A Priori Truth", *Erkenntnis* , 13, 153-170.

Acknowledgement.

Thanks are due for comments from a referee of this journal, as well as John Bigelow, Graham Nerlich, Graham Priest, Greg Restall and others when an earlier version of this paper was read at a meeting of the Adelaide-Melbourne Logic Seminar.

PART 5

Papers on The Philosophy of Mathematics and Physics

CHRIS MORTENSEN AND GRAHAM NERLICH

PHYSICAL TOPOLOGY[1]

1. INTRODUCTION

This paper is concerned with the ontological status of intervals or stretches of physical space.[2] Some writers[3] have adopted a view of intervals and points in space according to which the former are sets of the latter. We are particularly concerned, in this paper, to deny this claim with respect to continuous physical intervals in a continuous physical space.

There are at least two motives for treating the intervals of a continuous physical space as sets. One motive is topological. If physical space and spatial intervals are entities different from the bodies which occupy space, and if the topology of a physical space is a real property of it (in the sense that the space has exactly one topology), then there will need to be some way of stating the relations between physical intervals and, perhaps, physical points, which constitute the topology, or the ordering properties of those intervals. But the only existing machinery for constructing the topology of the real line is a set-theoretic one. If, then, spatial intervals are continuously ordered, we will need to call them and their orderings sets in order to have any hope of saying in what their continuity consists.

In this paper, this position will be disputed. It will be argued that the existence of certain set-theoretic ordering relations is insufficient by itself to distinguish between a continuous and a non-continuous physical interval. It will be argued that intervals in a continuous physical space are not sets of points. This raises the question: if intervals are not sets, in what does their continuity consist? An answer to this question will be offered.

As standardly treated, topology is a branch of set theory: topological spaces are structures of sets. Already, however, we have spoken both of the topology of the real line, and also the topology of physical spaces and intervals, which we claim not to be sets. In order to make this difference clear, we will (usually) speak of *set-theoretic* topology as opposed to *physical* topology.

Another motive for viewing intervals as point sets is metrical. In measure

Journal of Philosophical Logic 7 (1978) 209–223.

theory, measures are measures of subsets of a set, i.e., measures are measures of sets. In this connection, it is worth outlining how Grünbaum commits himself, for measure-theoretic reasons, to the thesis that intervals are sets.[4] Grünbaum argues that continuous space must always have a merely conventional metric, whereas discrete space need not. Very roughly, his argument is that the cardinal number of points in any continuous interval is the same as in any other. But the cardinal number of points of an interval of discrete space is not, in general, the same as in another, so each discrete interval can be metrically characterised intrinsically, by counting the number of grains which are members of it. Now, Grünbaum's conclusion seems open to a simple objection, urged by several writers.[5] The objection is that he assigns a size to the smallest granular parts of discrete space when he measures intervals by the number of grains they contain, but this assignment must be no less a convention than any assignment of measures directly to intervals of continuous space would be. Grünbaum dismisses this objection, claiming to be able to avoid it by treating intervals of discrete space as *sets* of their grains. This allows him to give a measure to the set (interval) by way of its cardinal number, without this entailing that the members (grains) of the set have a size in any sense. In response to Grünbaum, Nerlich[6] argues that in discrete space, the indivisible grains may only be treated as parts of intervals, and hence as intervals themselves, so the simple objection succeeds. However, Nerlich further contends that Grünbaum has improperly carried over to discrete space a set-theoretic notion (membership) which *correctly* holds between points and intervals of dense or continuous physical spaces, and it is this last claim we dispute. For these reasons we do not think that Grünbaum could accept our mereological account as the real metaphysical truth for which his set theoretic description is merely an (accurate) *façon de parler*. Grünbaum needs set theory for his defence of conventionalism against the simple objection just mentioned. But Nerlich could (and does) accept our arguments since he is no conventionalist. Later, we will offer an account of the measure-theoretic properties of physical intervals different from Grünbaum's.

2. ANY SET HAS MANY TOPOLOGICAL STRUCTURES

Our basic reason for claiming that the structure of continuous intervals of physical space cannot be accounted for by treating them as sets with certain set-theoretic ordering relations on them is this: too many orderings exist for any given set. The point can be expressed in the form of a theorem.

> Given any set of cardinality $2^{\aleph_0}$, there exists a binary relation on S which is a continuous simple ordering on S.

The proof of this theorem in standard set theory is not difficult. If S has cardinality $2^{\aleph_0}$, then, by definition, there exists a one–one correspondence f from S to $P(N)$, the power set of the natural numbers N. It is known that there exists at least one one–one correspondence, say g, from $P(N)$ to $(0, 1)$, the open interval of the real numbers R between 0 and 1. The natural ordering, $<$, on $(0, 1)$, which is a continuous simple ordering, then induces a continuous simple ordering O on S, by xOy iff $gf(x) < gf(y)$. The correspondence gf is clearly order-preserving between $(0, 1)$ under $<$, and S under O.

But, of course, S can be any set of the power of the continuum. In particular, S under its natural ordering can be disconnected, e.g., the union of two disjoint open intervals of real numbers. S can be discontinuously ordered, e.g., all the irrational real numbers between 0 and 1. Or S can be some arbitrary set of spatial points. It follows, then, that what makes the members of a set of spatial points make up a continuous spatial interval cannot be just the existence of a continuous simple ordering (set of pairs) on that set because, for any set of the requisite cardinality, there exists such an ordering.

A topological space in set theory is not simply a set, but rather a set together with a topology, which is a structure of subsets of that set. But this is to say that a given set has, in general, many topologies. Enough subsets, either of the set or its Cartesian product, exist to give the set a variety of topologies or orderings. Therefore, the ordering properties of intervals of physical space are not given simply by the existence of certain ordered sets, for this gives no explanation of why physical intervals should be ordered in one way, or have one structure, rather than another. But if the topological motive for identifying physical intervals with sets is to account for their physical structure in set-theoretic terms, the above argument seems to remove the temptation from topology to make that identification.

3. SOME ALTERNATIVE POSITIONS

There are several different positions which might be adopted in order to deal with the considerations of the previous section. In this section, we outline four. In the next section we will outline our own position.

First, we might deny, as Reichenbach[7] seems to do, that physical space has a unique physical topology. This move seems tantamount to denying that space is real (although it is not obvious that it follows from the denial of a unique topology), for the denial of a unique physical topology seems to remove all motivation for affirming the reality of space. As we see it, the principal motivation for claiming that space, in addition to bodies, is real, is to account for the physical topological properties which bodies display.[8] Thus, we are disinclined to adopt this alternative.

Second, one might say that only certain ordered sets of spatial points exist. The ordering properties of a physical interval are accounted for in terms of an ordering relation, which is a set, on that interval, which is also a set. What makes the interval ordered one way rather than another is that only certain ordered sets of its members exist. We might call such sets Aristotelian sets, an analogy with the difference between Platonic and Aristotelian universals. The objection to this approach is, of course, that the problems of constructing such a set theory are daunting, to say the least.

Third, one might say that an interval is an *ordered pair* of a set of spatial points and a suitable set-theoretic ordering. According to this view, the relation between points and intervals is not, as Grünbaum holds, that of being a member of, but that of *being a member of a member of*. Now this view does not suffer from the difficulty that any set of the requisite cardinality can be re-ordered so as to have a continuous simple ordering, because the orderings of intervals and non-intervals are built into them, as it were. But it suffers from a related difficulty, namely that if a certain pair $\langle S, O\rangle$ were, on this view, an interval, an infinite number of different pairs $\langle S, O'\rangle$ would also exist. Indeed, just two will do, and if O is a continuous simple ordering, the Axiom of Choice will guarantee us another, well-ordered. Then we should have the problem of saying why it is that we select just one of the pairs to be the interval containing the members of S. This point is, of course, a particular instance of the more general argument that a set has in general many topologies, so that set theory by itself doesn't tell us why we should select one topology rather than another as being descriptive of physical space.

Fourth, we might try to save the position just mentioned by introducing into the proposed set-theoretic description of the topology of space and intervals, a primitive function, which we might designate by '#', which selects from the set of all pairs $\{\langle S, O\rangle\colon O \subseteq S^2\}$ just one pair for each S, which gives the 'natural' ordering on S, and which might be written '$S^{\#}$'. This position was proposed in conversation by David Lewis, who argued that it would have the advantage of conceptual economy over the position we advocate, in that if we introduce the idea of non-set-theoretic relations (as we later do), we involve ourselves in messy metaphysics.

But this view, while in some sense descriptively simpler, leaves one with the same sort of problem as has been raised earlier. There are many such functions which select from any set $\{\langle S, O\rangle\colon O \subseteq S^2\}$ just one member. If U is the set of all physical points, then repeated applications of the Axiom of Choice to the set $\{\{\langle S, O\rangle\colon O \subseteq S^2\}\colon S \in P(U)\}$ shows that many functions like # exist. Why choose one as giving the natural ordering of points in physical space rather than another? If space has a unique physical topology, then the mere *existence* of the function # will not guarantee that space has one topology rather than another. Could it be, perhaps, in virtue of some property of the particular function #? But the attempt to say what that property is within set theory is clearly leading us into regress. The whole point is that set-theoretic apparatus *by itself* here cannot describe in what the physical topology of space consists. Of course, we do not deny that set theoretical tools may be of the greatest use to us in stating the essentially mereological facts and we make free use of them in what follows.

4. PHYSICAL TOPOLOGY IN TERMS OF PARTS AND WHOLES

If physical space is real with a unique physical topology, then this topology must be constituted by relations *which are not sets* between spatial entities (points or intervals). Having said what such relations are not, we wish to avoid saying what they are, for that would involve us in settling the general metaphysical problem of how predications are true. They are whatever makes relational sentences true, though it is our view that such entities have to be along the lines that Plato described. The point of the two approaches described in this section is that in dispensing with a set-theoretic description of the structure of intervals, we dispense with rhe need to speak of intervals as sets. This evidently has considerable intuitive advantages. This is not to say, however, that we dispense with sets, as will become clear. In the first approach, points are taken as basic. In the second, intervals are basic. The relation between points and intervals is that of part–whole. Points are unique parts, in that they have no proper parts. The two approaches do not differ substantially since, from an adequate definition of the continuity of intervals, atomic parts of these intervals (i.e., points) should emerge.

We assume a well-defined part–whole relation, for example along the lines of the Calculus of Individuals; but we do not attempt to construct analysis entirely within this framework, but use set theory freely. Success in the further venture of constructing analysis without set theory would show something stronger than we wish to claim, namely that set theory is unnecessary for physics.

(a) *Points*

We restrict ourselves in two ways. First, we aim to define only the continuity of intervals, and not further notions, e.g., openness, which are topological. Second, we restrict ourselves to one-dimensional space with the 'topology' of the real line. Multi-dimensional spaces constitute an important further problem. For example, evidently multi-dimensioned spaces cannot, consistent with our position, be treated as sets of n-tuples of points. Within these restrictions, perhaps the simplest way of saying what it is for space to be continuous is to say that spatial points are so ordered by a (not-set-theoretic) relation T, that there exists an order-preserving one–one correspondence between the set of points ordered under T, and the set of real numbers under its natural ordering, $<$. Alternatively, one could say that points are so related by *betweenness*, that there exists a relation-preserving one–one correspondence between the set of points related by betweenness, and the set of real numbers related by betweenness on that set. Clearly, it follows that there are $2^{\aleph_0}$ spatial points, and that T continuously simply orders them. These definitions have the advantage of underlining the fact that our approach does not deny sets, and does not prevent the set-theoretic study of the real number system from being useful to the study of the structure of physical space. The above one–one correspondences induce one–one correspondences between the set of sets of physical points and the set of sets of real numbers. We can then define a physical interval to be a whole whose point-parts are all the members of a set which is the image under the one–one correspondence of some real number interval. This ensures, for example, the property that any point between any two points of a physical interval is a point of the interval.

We now attempt to construct in greater detail such a structure of space, spatial points and intervals.

Betweenness defined for points in the one-dimensional universe suggests itself as a natural starting point with a solid basis in intuition. Betweenness satisfies various conditions, and we give five (where '$Bxyz$' stands for 'y is between x and z').

(1) $(x)\,(y)\,Bxxy \mathbin{\&} \sim Bxyx$

(2) $(x)\,(\exists y)\,(\exists z)\,Byxz$

(3) $(x)\,(y)\,(z)\,(Bxyz \supset Bzyx)$

(4) $(x)\,(y)\,(z)\,(x \neq y \neq z \supset (Bxyz \equiv (\sim Byxz \mathbin{\&} \sim Bxzy)))$

(5) $(x)\,(y)\,(z)\,(u)\,((Bxyz \mathbin{\&} Byzu) \supset (Bxyu \mathbin{\&} Bxzu))$

An *interval* is a sum of points, such that given any two points part of that interval, any point between those points is a part of the interval. The universe is an interval. The universe is the sum of all points, and the sum of all intervals. A *doubly bounded* interval is an interval such that there exists exactly two points, x, y, called the *end-points* of the interval, such that every point part of the interval not identical with x or y is between x and y. An *open* doubly bounded interval is a doubly bounded interval such that every point part of the interval is between its end-points. If x and y are the end-points of an open doubly bounded interval z, z is denoted by '$O(x, y)$'. A *closed* doubly bounded interval is a doubly bounded interval whose end-points are parts of it. If x and y are the end-points of a closed doubly bounded interval z, z is denoted by '$C(x, y)$'. A *half-open* doubly bounded interval is a doubly bounded interval just one of whose end-points is a part. If x and y are the end-points of a half-open doubly bounded interval z, and x is a part of z and y is not a part of z, then z is denoted by '$H(x, y)$'. $C(x, y) = O(x, y) + x + y = H(x, y) + y = H(y, x) + x$. $C(x, y)$ is called the *closure* of $O(x, y)$, $H(x, y)$ and $H(y, x)$.

Two points x and y are said to be *on the same side* of a point z, if z is not between x and y. A *singly bounded* interval is an interval x such that there is a point y, called the *end-point* of x, such that (i) every pair of points which are part of x and not identical with y are on the same side of y, and (ii), for any point z on the same side of y as some point u which is a part of x, z is a part of x. An *open* singly bounded interval is a singly bounded interval whose end-point is not a part of it. A *half-open* singly bounded interval is a singly bounded interval whose end-point is a part of it.

We will not venture into the territory of attempting to describe analogues of open sets, closed sets, neighbourhoods, etc., in physical space. Clearly, however, we cannot reproduce exactly the topology of the real numbers. There is no analogue of the null set to be open or closed. Furthermore, various of the conditions for closedness and openness break down in the case of the universe because of the non-existence of a 'complement' for the universe.

Density of an interval is straightforward: between any two points of the interval is a third. The universe is dense iff every interval is dense.

The test of this approach is continuity. The following seems to be adequate. A closed doubly bounded interval x is continuous iff three conditions are satisfied[9]

(1) x is dense.

(2) For any proper subinterval y of x which contains an end-point of x, either y is a half-open doubly bounded interval and x–y is a closed doubly bounded interval, or y is a closed doubly bounded interval and x–y is a half-open doubly bounded interval.

(3) There exists a countably infinite set S of points each of which is a part of x, such that between any two points of x is a member of S.

An open or half-open doubly bounded interval is continuous iff its closure is continuous. The universe is continuous iff every open doubly bounded interval is continuous.

It is not difficult to extend this definition of continuity to the singly bounded intervals. A *bounded* interval is a singly or doubly bounded interval. The universe is continuous iff every bounded interval is continuous. An interval is either a bounded interval or the universe. The universe is continuous iff every interval is continuous.

(b) *Intervals*

A second approach takes intervals as basic. A reason for doing this is what might be called the epistemological priority of intervals over points: separation and intervals between bodies are somehow visible in a way that points are not. We suppose that the set of all spatial wholes (constructed by closing the set of spatial intervals under the relations of part and whole) exists. Then we can say that space is continuous if there exists a one–one correspondence between the set of wholes and the set of subsets of the real numbers (excluding the null set) which maps physical intervals to real number intervals and vice versa; sums, overlaps and parts to unions, intersections and subsets and vice versa, and preserves betweennesses of wholes in betweennesses of sets of real numbers and vice versa. (These conditions are not intended to be independent). We require that any whole which has a proper part be mapped to a set with a non-empty proper subset. This gives us points: those spatial entities whose images are singleton sets have no proper parts, for, if they did, their images would have proper subsets in the mapping, which singleton sets do not (the null set is not the image of anything). There are $2^{\aleph_0}$ points. Singleton sets of real numbers have the same betweennesses as the real numbers, so, since the one–one correspondence preserves betweennesses, we have that points are continuously simply ordered. Physical intervals, then, are seen, e.g., to have the property that any point between any two points of the interval is a part of the interval.

We now attempt, in more detail, to construct continuity from the primitive ideas of intervals, part–whole containment and an ordering on intervals given by Charles Hamblin[10].

As a background logic, we take the usual first-order predicate calculus with identity, suitable axioms for the part-whole relation, and a single two-place predicate '<' to be read as 'precedes'. Quantifiers range over intervals. *Abutment* is defined as

Defn 1: $aAb =_{\mathrm{df}} a < b \,\&\, {\sim}(\exists c)(a < c \,\&\, c < b).$

We then have the following axioms

A1. ${\sim}(a < a)$

A2. $(a < c \,\&\, b < d) \supset (a < d \vee b < c)$

A3. $a < b \supset (aAb \vee (\exists c)(aAc \,\&\, cAb))$

A4. $(aAc \,\&\, aAd \,\&\, bAc) \supset bAd$

A5. $(aAb \mathbin{\&} bAd \mathbin{\&} aAc \mathbin{\&} cAd) \supset b = c$

A6. $(\exists b)$ (b is a part of $a \mathbin{\&} b \neq a$)

A7. $(\exists b)(a < b)$

A8. $(\exists b)(b < a)$

Hamblin[11] shows that these axioms are sufficient for dense, linear connected order if points are defined as ordered pairs of abutting intervals. We go on to construct continuity. We need the notion of *proper overlapping* holding between intervals. Following Hamblin, we define

Defn 2: $aPb =_{df} \sim(a < b \vee b < a)$.

This is sufficient to ensure that two intervals which proper overlap do so in an interval and not just a point.

y is a *proper side part* of x iff y is a proper part of x and $(\exists z)$ (yAz & $y + z = x$). Take some interval x and order the set of its proper side parts by the relation, being a part of. This gives us a simple ordering of the proper side parts. We call this ordered set x^*.

The density of the interval x may now be stated. The proper side parts of x are *densely nested* iff x has a proper part and, for any proper side parts, u, v of x, if u is a proper side part of v then there is a w such that u is a proper side part of w and w is a proper side part of v. x is *dense* iff the proper side parts of x are densely nested. This corresponds to the condition on dense intervals that between any two points there is, not merely a third point, but an interval which, of course, contains non-overlapping subintervals.

Consider just those proper side parts of x in the simple ordering x^*. The preceding condition ensures that there will be 1–1 functions $f^1, f^2 \ldots f^n \ldots$ each from the rational numbers in the open unit interval (0, 1) into the proper side parts in x^*, such that (i) for any two rationals $u, v, \in (0, 1)$ $u < v$ iff $f^i(u)$ is a proper side part of $f^i(v)$. (ii) for any proper side part a in x^*_i there are rational numbers $u, v, \in (0, 1)$ such that a is a proper side part of $f^i(u)$ and $f^i(v)$ is a proper side part of a. This last condition ensures that the images of the rationals 'spread end to end' in x. Functions such as these permit *cuts* among the proper side parts in their range. For any such function f^i, a *cut* among the side parts in its range is a pair of disjoint non-empty sets A and B such that each side part in the range of f^i is in A or in B

and each $a \in A$ is a proper side part of each $b \in B$. We denote an f^i cut by $[A, B]$.

We now state two conditions on an interval analogous to Dedekind's construction of the real numbers from cuts in the rational numbers. The conditions are necessary and sufficient for continuity. x is *continuous* iff (a) x is dense and for every f^i cut $[A, B]$ of x there is a proper side part z of x such that just one of the following holds:

(1) $z \in A$ & each a in $A(a \neq z)$ is a proper side part of z.

(2) $z \in B$ & each b in $B(b \neq z)$ contains z as a proper side part.

(3) $z \notin A, z \notin B$ & every $a \in A$ is a proper side part of z & z is a proper side part of every $b \in B$

and (b) there is a non-empty, countable subset S of x^* such that for any a and b in x^* such that a is a proper side part of b, there is a $C(C \in S)$ such that a is a proper side part of C and C is a proper side part of b. (Separability condition.)

This condition ensures that there are 1–1 functions $F^1, F^2 \ldots F^n$ from the set of real numbers in the open interval (0, 1) into the proper side parts of x which are in x^*, such that for any two real numbers $a, b \in (0, 1)\, a < b$ iff $F^i(a)$ is a proper side part of $F^i(b)$. We believe that our use of the *proper* side part relation prevents trivial cases where an open interval and its closure both figure as images of real numbers under the same F^i mapping. Our idea of proper side part is really the idea of a *preceding* side part. From all the devisions of the interval x into pairs of abutting subintervals we chose, in each case, the abutting (preceding) interval. But we can equally speak of the second of each of these pairs of subintervals as a succeeding proper side part. These proper side parts can also be simply ordered by the relation, being a part of, and we refer to this ordered set of succeeding proper side parts as x_2^*.

Now there are functions g^i, G^i from the rational and real numbers (respectively) of the open unit interval into the succeeding proper side parts of x which are in x_2^*. The g^i and G^i clearly resemble the f^i and F^i, but we choose, for example, the G^i so that for any real numbers $u, v, \in (0, 1)$, $u < v$ iff $G^i(u)$ contains $G^i(v)$ as its proper side part. Thus (0, 1) is isomorphic under G^i with x_2^* where the *inverse* of the relation being a proper side part orders x_2^* as '$<$' orders (0, 1).

Consider the inverses of all these functions from the domain of the reals into x^* and x_2^*. Call them $\bar{F}^i$ and $\bar{G}^j$. The products of the $\bar{F}^i$ with the $\bar{G}^j$ map from x^* into x_2^*. So there will be infinitely many 1–1 correspondences between the x^* proper side parts of x and the x_2^* proper side parts of it which are isomorphic under the part–whole relation and its inverse. Since x is a continuous interval, there is exactly one such correspondence C with the following properties. All the proper side parts of x^* and x_2^* paired under C (a) have a common part (b) do not proper overlap, i.e., have no interval as common part. Intuitively, for a closed interval x, C yields each pair of closed left hand and right hand proper side parts of x which overlap in the point which the two side parts contain in common. We postulate that the parts described in (a) and (b) above are points in the sense of earlier sections.

Clearly, these results can be generalised to spaces with different global topologies, for example, to the circle. Without beginning such a construction from scratch, we might proceed as follows:

A *cyclic interval* is a sum of open intervals each with the topology described but such that for any two points which are parts of the cyclic interval there are two distinct points such that they pairwise separate the first pair. Some axioms on pair separation for cyclic intervals follow, where we write $S(x, y/z, w)$ for 'x and y separate z and w'. We take 'separates' as a suitable primitive, analogous to 'between'.

(a) $S(x, y/z, w)$ if and only if $S(z, w/x, y)$.

(b) $S(x, y/z, w)$ if and only if $S(x, y/w, z)$.

(c) If $S(x, y/z, w)$, then it is not the case that $S(x, z/w, v)$.

(d) If $S(x, y/z, w)$ and $S(x, z/y, v)$, then $S(x, z/w, v)$.

(e) If x, y, z, and w are distinct points, then x is separated from one of the others by the remaining two.[12]

5. CONCLUDING REMARKS

Yet another alternative position to ours is as follows. Someone might wish to hold a 'mixed' theory, according to which intervals are sets of points, but what orders them in nature is a non-set-theoretic relation. Now unless this view denies the part–whole relation (a drastic step), it seems to us that it differs descriptively, but not ontologically, from ours. Both views hold that physical points are real, so both views will agree that any set constructible from those points exists. Both views agree that very many sets of ordered pairs of physical points exist, and both views hold that in addition, what is responsible for physical topology is the relatedness of points, not the existence of certain sets of ordered pairs. The difference between the two views is, basically, in the reference of the term 'interval'. We wish to say that things like *this* (the thing I am now passing my hand through) are the intervals and regions of physical space, and it does not seem that such things are sets. *This* is a whole, whose indivisible parts are points. The alternative view presumably wishes to describe *this* as a set, and on that point we wish to disagree, though perhaps this is not the place to argue it.

The type of argument given in section two has broader ramifications for the general theory of universals. There is considerable reason for supposing that properties and relations exist in addition to objects, in order to account for sameness, change, and why a universe of particulars should be one way rather than another. But an identification of properties and relations with sets, even sets of extensions in all possible worlds, would seem to be insufficient to determine just which universals are instantiated in a given world, for instance ours. Every set which can be made up from existing objects exists, and this seems to show that we need some further explanation for just why *these* objects are, for example, red, over and above the fact that they are members of a particular set in our world, which is in turn a member of a set of extensions in all possible worlds.

The account we have given of the structure of intervals gives an intuitively acceptable account of how matter can continuously occupy space. One can say that a continuous and connected one-dimensional rod is continuous and connected because (1) it is made up of points of matter, and (2) there is a continuous and connected physical interval such that every matter-point is at some spatial point of the interval, and every point of the interval has a point of matter at it. It is important to see, though, that the relation of being at, holding between matter-points and spatial points, is not a set-theoretic one–one correspondence. For, if it were, then, as in section two, any set of matter-points of cardinality $2^{\aleph_0}$ could be placed in one–one correspondence with the set of points of some suitable continuous and connected spatial interval. One might try to avoid this by constructing all over again the physical structure of rods, in parallel with the constructions of section four. But an easier way is to take *being at* as a primitive non-set-theoretic relation. Then we have the consequence that the topological properties of rods ride on the back of the topology of spatial intervals. This seems intuitively correct: the topological properties of rods are derivative from the topological properties of the space in which the rods are embedded (which is not to deny that the presence of material rods might not be causally relevant to topology or metric).

Finally, we specify an analogue of Lebesgue measure on one-dimensional intervals. We define a *co-ordinate system* to be an order-preserving one–one correspondence between the set of points and the set of real numbers. This correspondence induces a one–one correspondence between the set of subsets of points and the set of subsets of real numbers. To each set of points, except the null set, there corresponds a unique whole, so we have a one–one correspondence between the set of wholes of points and the set of subsets of real numbers excluding the null set, as before. Intervals are mapped to intervals, etc., and it follows that we can assign to the image of any set, the Lebesgue measure of that set. We have that the measures of points, and any spatial objects made up of denumerably many points, is zero (and this holds in any such co-ordinate system). The measures of the objects associated with the Borel sets (a, b), $(a, b]$, $[a, b)$ and $[a, b]$ are all $b-a$. The measures of the singly bounded physical intervals, and the whole physical line, are infinity.

Different measures assigning non-zero measure to a finite number of points can also obviously be constructed.

We note two points about the foregoing. First, a spatial interval can have an infinite number of unextended *parts* (unextended in the sense of having zero measure), and yet be extended. Second, differing co-ordinate systems confer different measures (in the above sense) on a space. So *to that extent* measure is conventional. But on the other hand, not any measure is possible. In particular, measures which assign non-zero size to points or denumerable spatial wholes prohibit co-ordinate systems and Lebesgue measures like the above.

We take this conclusion that any *point*, not just its singleton, has a measure *viz.* zero, to be directly contradictory to Grünbaum's thesis that measure cannot intelligibly be applied to members of sets of points in any space (*op. cit.* p. 575).

This is a point of some importance against Grünbaum. As noted in

Section 1, Grünbaum's defence against the objection that the metric of a discrete space must be as conventional as the metric of a continuous space rests on the claim that measure can only intelligibly be assigned to sets of points. But here we have specified something closely analogous to measure but which can be applied directly to the parts of intervals. It is up to Grünbaum to show that what has been done here cannot do the job of measuring the size of the parts of space. It goes without saying that we think that it can, though we leave the defence of that position for another occasion.

University of Adelaide

NOTES

[1] We are indebted for helpful suggestions to the following people: Robert Farrell, Henry Krips, David Lewis, Gordon Munro, Malcolm Rennie and J. J. C. Smart.

[2] Throughout this paper, 'space' is intended to be neutral between 'space' and 'spacetime'.

[3] Two are A. Grünbaum, especially in his 'Space, Time and Falsifiability', *Philosophy of Science* 39 (1970), 469–588, and Nerlich, G., *The Shape of Space* Cambridge University Press, 1976.

[4] The points below are argued in Nerlich, *op. cit.*, Ch. 8, Section 9–11. For Grünbaum's account, see especially Grünbaum *op. cit.*

[5] For example, in R. G. Swinburne 'Review of A. Grünbaum *Geometry and Chronometry in Physical Perspective*', *British Journal for the Philosophy of Science* 21 (1970), 308–311.

[6] Nerlich, *op. cit.* Ch. 8, Section 8 and Section 10.

[7] See, e.g., H. Reichenbach, *Philosophy of Space and Time*, Dover, 1958, Ch. 1, Section 12.

[8] See, e.g., Nerlich, *op. cit.* Ch. 7.

[9] For these conditions, see, e.g., R. L. Wilder *Introduction to the Foundations of Mathematics*, 2nd ed., Wiley, 1965, p. 150.

[10] See C. Hamblin, 'Starting and Stopping' *The Monist* Vol. 53 No. 3 (July 1969), 410–425; 'Instants and Intervals', *Studium Generale* 24 (1971), 127–134; 'The Logic of Starting and Stopping', in C. Pizzi (ed.) *La Logica del Tempo*, Torino, Moringhieri, 1974.

[11] See "Instants and Intervals" *op. cit.*

[12] See Bas C. van Fraassen *An Introduction to the Philosophy of Time and Space*, Rondon House (1970).

Ratio, XXV, 1, 1983

SPACETIME AND HANDEDNESS[1]

Chris Mortensen and Graham Nerlich

I: Preliminary

As it is well known, Kant argued that the difference between a left hand and a right hand forces us to conclude that space, considered as a thing distinct from hands and other bodies in it, exists. Evidently, this argument is relevant to the dispute between what is sometimes called 'absolutism', according to which space exists as an entity distinct from material bodies and the relations between them; and relationism, for which space consists in nothing but material bodies and relations between them. (We will presently qualify this description of the disagreement.) Kant's argument has given rise to a fairly extensive literature (see bibliography). A version of the argument was defended by Nerlich ((1973) and (1976) Ch. 2). One feature of the literature is that it has tended to concentrate on the implications of Kant's argument for the existence of space. In the post-relativistic era of spacetime however, this feature gives the discussion something of a dated air. In this paper, we propose to remedy the defect by enlarging the scope and realism of Kant's argument. We will argue that Kant's argument, set against a background where the primary ontological commitment should be to spacetime rather than to space, remains basically correct, at least when properly qualified.. Conversely, a proper development of Kant's argument in this setting is instructive for understanding the Special Theory of Relativity. Before we come to this, however, it seems to us important to sketch briefly and evaluate the broad outlines of at least some of the issues raised by Kant's argument, since our later argument essentially grows out of that discussion. The sketch will occupy the next section, and our main argument will occupy the subsequent section.

II: Moves in the Kantian Argument

We begin by entering two immediate warnings about the earlier characterisation of the dispute between 'absolutism' and relationism. First, as we described it, it is a dispute about the reality of space and its nature. But the phrase 'absolute space' was appropriated by Newton

[1] We wish to give particular thanks to Ian Hinckfuss, whose persistent criticisms have helped us to clarify our thoughts on the matters discussed in this paper. Thanks are also due to Jack Smart.

to stand for an entity which supplied the standard of nonaccelerated motion, and with it also an absolute distinction between motion and rest. It was a major dispute in the history of physics about whether there is absolute space in this sense. (See e.g. Rindler (1977) Ch. 1). We have no desire to remove the word 'absolute' from long-established usage, so we will describe the dispute that we are interested in, as that between 'realism' (about space, or spacetime) and relationism. It will emerge later that there are connections between the two disputes, however. Second, realism and relationism are both groups of positions. As is not uncommon in disputes about ontology, the lines can be drawn in different ways. For example, a relationist might be found who is prepared to assert the existence of space as an entity distinct from, i.e. not identical with, material bodies and their relations, provided that it is somehow a *construction* out of the latter, or is somehow not ontologically basic. Looked at like this, the issue is rather: if space is real, what is its nature? We will take it here that realism is at least committed to denying that space is a 'dependent' entity, in the sense that it is analysable in terms of, or reducible to, or constructible out of, resources acceptable to relationism. In particular, realism ought to hold that space is neither a set-theoretic nor a mereological construct out of material bodies and relations between them. This in turn raises the thorny issue of the nature of constructs and reductions, which, we will ignore on the principle that you can't talk about everything, and in the hope that the examples we select will not get us too caught up in the thorns.

One final preliminary point. We want to make a protest against physics worship. There is an attitude current, for which Quine has to take at least some of the blame, that the issue concerning the reality of space is a straightforward empirical one *to be solved by inspecting successful physical theory* to see what it quantifies over. Standard theoretical physics postulates a spatial or spatiotemporal manifold in which objects exist or events occur. The manifold looks for all the world like the realist's space or spacetime, and so we should conclude that space is real. We want to brand this 'physics worship'. What is wrong with it, is that it does not provide us with a fine enough seive for discriminating between kinds of entities which are, so to speak, *essentially* postulated by theory, and kinds of entities which are there *only at the convenience of the theorist*, or, perhaps, because theoretical practice in the area has not considered the possibility of their elimination by a reformulation of the theory. The best example we can think of is the use of set theory by excessively mathematically-minded theoretical physicists and philosophers. We have argued elsewhere (Mortensen and Nerlich (1978)) that set theory cannot by itself give a correct description of the structure of physical space, and that the use of set theory in this area gives a thoroughly misleading picture of what is really going on. Another example is the metric tensor. The use of the metric tensor apparently attributes a unique metrical structure to physical space, which uniqueness is incompatible with the metric conventionalism of Grünbaum and others. It would be a mistake, we think, to conclude from this alone that metric conventionalism must be false. Rather, metric conventionalism ought to be seen as offering a genuine research programme: to rewrite physics so as to do away with what must, if conventionalism is right, be a misleading theoretical formalism. We also have some suspicion of the use of phase spaces and Hilbert spaces in this regard. Our general point is that it is often *not* the concern of the theorist to look at possible reductive accounts of the kinds of entities he or she is interested in, so that certain kinds might get grabbed in the bag of values of bound variables when they need not have been. Developing a theory and taking a critical, reductive look at a successful developed theory are different games with different aims, played by different people. The latter game is one which might fruitfully be played by the philosopher of science. Thus, we want to urge that no matter how realist theoretical physics might *look*, relationism might come up with an alternative to realism, and it might even have theoretical advantages over realism. So, don't worship the physics that is : after all, it was made by mere physicists.

Kant's argument from hands to absolute space runs, very roughly, as follows: there is a difference between left and right which is neither intrinsic to hands, nor a relation between the one hand and the other, nor between a hand and any part of space it fills. The difference lies in a relation between the hand and space regarded as a whole. Here, properties which appear to belong to things are derived from space rather than vice-versa, as is usually supposed. This gives the argument its interest for absolutists or spatial realists. The argument can be extended and enriched by looking at what happens when handed objects are placed in various (non-Euclidean) spaces (as is done in Nerlich (1973) and (1976)). The phenomenon of handedness, incongruent counterparthood or enantiomorphism (to list some current expressions) is not confined to hands, of course. Objects with the sort of asymmetry needed are simply everywhere and this adds much to the force and interest of this style of argument.

Several relationist responses to these ideas are worth pursuing. We mention one and offer brief reflections on another. One approach is suggested by Lawrence Sklar, in his (1974) article. The relationist is envisaged as making use of possible as well as actual objects and motions. Sklar himself expresses reservations about this tactic. We think that such analyses are either inadequate or circular in ways spelt out in Nerlich (1976) Ch. 2 § 8.

In point of fact, the idea of congruous counterparthood is not really a temporal idea at all, and thus only in a Pickwickian sense one involving motions. Euclid's geometry was atemporal, but handedness and counterparthood are certainly possible therein. Handedness is possible in 4-dimensional spacetime, but in no sense of 'move' can spacetime worms be moved. The congruity of counterparts is a matter of a continuous function which 'projects' counterpart parts of space for every one of its arguments, and that is not a matter of motion, except perhaps at an intuitive level.[2] Rather, it is a matter of the counterparts being appropriately linked by paths, which are parts of space.

Another relationist strategy develops out of the observation that so far in the argument not a great deal has emerged about what kind of thing the space is which the realist has claimed. We have said that it is the sort of thing that has a topological structure, with continuous paths and dimensionality and also affine and metrical relations such as angle and distance between its parts. Now a relationist might be found to claim that a space of 'points', having all the above features, can be *constructed* out of items which are relationistically acceptable. We are guided here in part by Quine's discussion in *Word and Object*. (Quine (1960) Section 52). If viable, such a possibility at least shows that the Kantian argument does not establish the falsity of relationism, even if relationism so construed looks to have been somewhat extended. We therefore look a little more closely at the possibility.

Begin with the primitive idea of *spatiotemporal distance* (we concentrate on spacetime and not space because Quine does). Select five particle events, constrained only by the condition that there be no three-dimensional hypersurface containing all of them. (In passing, it is not *clear to us how* to make this constraint relationistically acceptable). The five events form a reference system, in the sense that were they to be embedded in a four dimensional manifold of points,

[2] Counterparts are incongruous if there is no continuous map from the closed real interval [0, 1] to the space containing the counterparts, whose values at the endpoints are the two counterparts respectively and whose value for any argument is a region of the space counterpart to the end-values.

any point would be uniquely specified by the quintuple of distances from each of the events. Reversing this idea, then, let a point be *identified with* every such quintuple of distances. With set theory, we can then set up the idea of a co-ordinate system on such a space; and also, since we are allowed distances, the metric tensor. Then we are away, with neighbourhoods, continuity, paths and so on in the usual fashion. The point of the story is that distances and set theory are supposed to be relationistically acceptable, so that even if handedness forces us to empty space with points, paths and topology, it need not force us to abandon relationism.

One point to make in connection with this strategy at the outset is that the mere fact of handedness does not establish that space is 'concrete' rather than 'abstract'. We leave these terms conveniently undefined, but an example will serve to explain what we mean. Enantiomorphism is possible in a 'space' whose points are n-tuples of real numbers. That much barely needs arguing. The set R^n of n-tuples of real numbers with Euclidean metric with some suitably shaped subset as the handed region serves as an instance. (In passing, this reinforces the point that handedness is not a matter of motion, for there is no sense to motion in such an atemporal abstract space). The Kantian argument thus has something to say about the *reality* of space, and about its *structure*, but it does not have anything to say about the *nature* of the parts of space. So the strategy we are considering does not count against the proposition: if there are handed objects, then there is space. What it does is to challenge the further proposition that if there is space, then realism is true and relationism is false. We are, however, inclined to belive this further proposition, and so we want to consider the scheme on its own terms.

We have quite a lot of worries about this kind of scheme, not the least of which is that it needs to be shown to be entirely adequate to the concepts of spatiotemporal geometry it is intended to capture. A second problem is that we are suspicious of the use of set theory in accounts of the structure of physical space, as we have argued elsewhere (Mortensen and Nerlich (1978)). One way to focus this suspicion is to try to spell out the ontology in more detail. What a spacetime point *is*, on this scheme, is a quintuple of numbers — a certain set. But can we literally take it that the space time point exists just if the quintuple does? Now, given standard set theories, the quintuple certainly does exist. In fact, all quintuples of numbers exist if any quintuple does. So is the scheme committed to a correspondingly rich array of spacetime points at all these distances?

That yields an infinite space with unnerving rapidity. We have no particular reason to believe that this conclusion is at all desirable, whether we regard it as a relationist or a realist result. Making spacetime points at distances exist for every quintuple of numbers enforces an embarrassment of riches on our ontology. However, there appears to be no principle which might "include out" some quintuples as spacetime points while "including in" others, and which is not circular in appealing to the structure of spacetime as independent of the quintuples.

It might be said that spacetime is some particular *set* of quintuples which is a subset (proper or not) of the set of all quintuples of numbers. Thus Minkowskian space time is a distinct set of quintuples from the set which makes up some non-flat Riemannian spacetime. But a similar problem crops up. Too many of these sets of quintipules exist. So that if a certain spacetime exists just if a certain set of quintuples exists then there strictly and literally are both flat and curved spacetimes, Newtonian spacetimes, and (for existing sets of $n + 1$ -tuples) spacetimes of n dimensions for arbitrary n.

This makes it look judicious to say that the sets are not (the very same things as) the spacetimes, but only represent them. We think it is judicious to say so. But saying so says nothing to the ontic purpose for it concedes and independent existence to spacetime.

We conclude this brief sketch of moves in the Kantian argument with the observation that we believe that the disease of circularity is a difficulty for any *kind* of relationist attempt to construct spatial hardware using modals, though obviously each attempt needs individual scrutiny. If e.g. possible relations between possible objects are claimed to be sufficient for the full paraphernalia of space, then the range of possibilities needs correct specification without appeal to the notions it is supposed to be reducing. Otherwise, we have no guarantee that the full structure of space has been produced without falsehood.

III: Spacetime and Enantiomorphism

As we said in Section I, the proper modern setting for Kant's argument are the notions of spacetime and relativity. To place it in that setting, we will pose a problem for it which grows out of relativistic considerations. The solution to the problem and its implications will occupy the rest of the paper.

We begin by observing, what many have observed, that following Special Relativity there are various senses in which there is no absolute space. (This is not to say that either Special or General

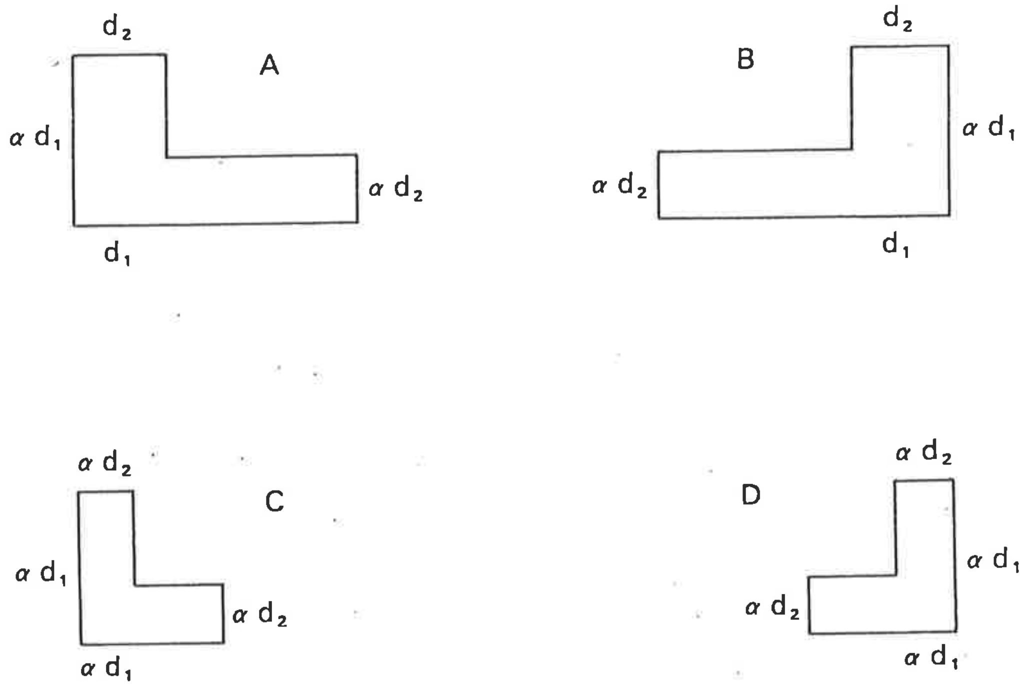
Relativity deny absolute spacetime). One sense, for example, in which absolute space is denied is that absolute spatial distances are denied. The end points of the same object will have different distances between them relative to different (Lorenz) reference frames. Another sense in which absolute space is absent in relativistic spacetime, is that is no standard of absolute position over time, as a hook for Newton to hang absolute acceleration on.

Now here is the problem. The Kantian argument appears to be intended to have *a priori* force. It is of the form: if there are handed things, then there is space. If correct, it would therefore have to hold no matter what the nature of the possible world being considered. Equally, then, the Kantian argument must have something wrong with it if it is possible for there to be a universe in which there are handed objects but no space. Now here you sit in your spaceship in our relativistic universe in which, as we have seen, there is no absolute space. Yet surely you have a left hand and a right hand, and they are incongruous counterparts of one another?

One way to misunderstand this problem is to take the premiss of the Kantian argument as being that there are handed or counterpart *spatiotemporal* regions. Relativity does not deny the existence of spacetime and so if there are incongruous counterparts in spacetime, spacetime must be real. Incongruous counterparts in spacetime would be suitably shaped spacetime worms, and if such worms exist, so must spacetime. Nevertheless, the problem we posed remains because left and right hands are incongruous *considered as spatial objects*. Hands in this context are not spacetime worms. Rather, they are timeslices of spacetime worms. What becomes of the space they are in then?

The beginnings of the answer lies in two things we have already seen. First, absolute space in the sense of absolute spatial distances is denied in relativity; that is, spatial distance is relative to choice of reference frame. Second, counterparthood is in part a matter of spatial distance, in that counterparthood implies sameness of distance between corresponding parts. It should come as no suprise then that congruous and incongrous counterparthood in a relativistic universe is frame-relative.

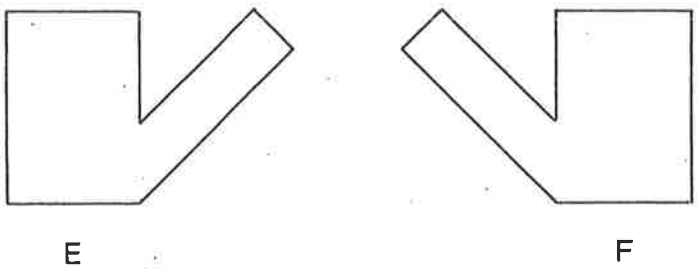
We illustrate the frame-relativity of congruous counterparthood by considering an example which seems to display the essential features of the matter. Consider a two-dimensional universe and the two-dimensional material bodies A, B, C, D below, with dimensions as indicated.



Bodies A and B are incongruous counterparts (in fact they look a little like left and right hands respectively). Thus A and B are handed, and opposite to one another. But C is not handed. A little reflection will convince the reader that any counterpart of C, such as D, is congruous with it. The number α is arbitrary, except that $0 < \alpha < 1$.

First consider a reference frame relative to which A and B are at rest. The dimensions of A and B are above, and relative to the frame A and B are incongruous counterparts and handed. Now choose a frame, moving away to the right from A and B. Relative to the second frame, A and B are moving left (along the negative x-axis). As is well known, in Special Relativity moving bodies undergo a length 'contraction' with respect to their rest length (*i.e.* their length relative to a frame in which they are at rest), and the length contraction is in the direction of motion. The length contraction is such that the greater the speed the more the length approaches zero. For any α, $0<\alpha<1$, we can choose some speed s such that relative to a frame moving at speed s any rest dimension d in the direction of motion is contracted to αd. Hence the rest dimension d_1, of A contracts so that relative to the new frame it is now αd_1. Similarly, d_2 contracts to αd_2. The dimensions at right angles to the motion, *i.e.* the dimensions up the page, remain unaffected. Hence the object A becomes, relative to the new frame, the object C. But whereas A is handed, C is not. Hence handedness, at least sometimes, is frame-relative (though we will presently see that this result is exceptional). By a similar argument, B becomes, relative to the second frame, D. A and B are incongruous counterparts, C and D are counterparts but congruous. Therefore, incongruous and congruous counterparthood are also frame-relative.

This has been an example where counterparthood is preserved relative to a suitable reference frame while congruity/incongruity varies. Commonly, counterparthood varies also. Real life hands, unless they are very unusual, are liable to be like this. For example, consider the incongruous counterpart hands E and F.



Choose a frame relative to which E and F are moving away together in the direction of E's (left) thumb. Then E's left thumb is shortened in length. (There will be other changes in E but we can ignore them). But F's thumb is inclined at right angles to the direction of motion, so it is not shortened. Hence E and F are not counterparts relative to this frame, and so counterparthood is frame-relative.

It seems to us straightforward to show that every pair of incongruous counterparts can have their counterparthood destroyed for suitable choice of frame. Being incongruous, they will not be reflections of one another for some choice of direction. Choose a frame moving in that direction and their dimensions will be modified differently. Interestingly, some pairs of figures have their congruous counterparthood invariant with respect to all choices of frame; circles in two dimensions, spheres and hyperspheres in higher dimensions. That does not mean that the idea of congruous counterparthood is sometimes not frame-relative. The basic relativistic notion is counterparthood-relative-to-a-frame. Only, some things are congruous relative to all frames.

However, while incongruous counterparthood is frame relative, handedness is, in general, not. Subject an actual hand shape to any Lorentz transformation and it remains left and handed, if it was left and handed before. A (three dimensional) left handed screw or spring remains handed relative to any (Lorentz) reference frame. More laboriously, the world-tube which is the 4-object associated with what is a left hand in some reference frame yields a left hand in the space of any other frame. So, in some sense, handedness *generally* is an invariant of Lorentz transformations though incongruous counterparthood is not. Thus handedness is not the same concept as incongruous counterparthood. One exception to the 'generally' is worth noting: the handed objects A and B above transform to the non-handed objects C and D respectively, so under some Lorentz transformations the handedness of some objects is not preserved. Specifically, those transformations which, while preserving the counterparthood of the particular shapes in question, vary their incongruence, also vary their handedness.

Incongruous counterparts focus usefully on the fact that enantiomorphic objects can be counterparts right up to reflection without ceasing to be oppositely shaped. Counterparthood as we define it, however, is a relational idea relating objects *to objects*: for any incongruous counterpart there is an object which is its incongruent counterpart. This makes sense of its being a metrical idea: there is always something to which it is incongruous, and counterparts are so related by sameness of internal distances. But handedness, though relative to a space (x is handed in space y) is not in our usage a relation of objects to objects. It is not, generally, a metrical notion. Every handed object is asymmetrical in shape, but how deep this asymmetry lies is to be gauged by the groups of transformations which preserve it. The transformations of the Lorentz group in Special Relativity are affine transformations from the point of view of shape transformations of spatial objects. Thus they destroy some asymmetries but not others. The asymmetry of a hand is not destroyed by affine transformations, so it is at least an affine asymmetry and we may speak of things that are affinely handed.

This brings us back, then, to the original problem: invariant spatial handedness is possible in relativistic spacetime; what becomes of space in relativistic spacetime if the Kantian argument for space is correct?

To understand what is going on, we need to understand what, in spacetime ontology, a space is, and what a frame is. Minkowski spacetime is a collection (we prefer 'whole', see (1978)) of spacetime points, with a certain metrical structure. A (Lorentz) frame is produced by cutting up that spacetime into 3 spatial dimensions plus a temporal dimension. For a time-constant, the frame gives us a three dimensional hypersurface, a spacelike slice of spacetime. In any given spacelike slice of a frame, the frame specifies spatial distances between points. Thus a frame specifies a collection of spacelike slices, one for each value of the time variable. Now there is a certain ambiguity in 'space', between space-at-a-time, and space-which-endures-through-time. The frame identifies the former as one of its spacelike slices. The latter is given in the frame by the fact that *same spatial position over time* can be defined as those points of the spacelike slices whose spatial co-ordinate relative to the frame is unchanging over time. Needless to say, such a space-which-endures-through-time is frame-relative. The class of all frames produced by the group of Lorentz transformations defines the class of permissible partitions of spacetime into a space (in either sense) and a time. Therefore, since spacetime is a collection of spacetime points, and since a space (in the sense of a spacelike slice or space-at-a-time) is a part of (or subset of) spacetime, space is a collection of spacetime points. Similarly, space-which-endures-through-time is a collection of spacelike slices relative to a single frame, thought of as having a relation of same-spatial-point-at-a-different-time defined on pairs of spacetime points in different slices. Thus relativity theory represents a significant ontological shift from the older Newtonian paradigm. Space though of as made up of points enduring through time is gone. So has sameness-of-spatial-point-over-time as a frame-invariant (real) concept. But space has not. Spaces as collections of spacetime points are still in the ontology. Not any collection of spacetime points is a space, of course. Only those associated with a permissable frame. But they are none the less real for that.

So the conclusion to draw is that the Kantian argument is not threatened by relativity, for space is still in the ontology on relativity. Spaces are made up of different entities (spacetime points) from what Newton thought. But recall what was said earlier, that the Kantian argument does not legislate on the *nature* of the points of the space in which handed objects are embedded. It is quite consistent with the Kantian argument that the points of space be fundamentally spatiotemporal.

As was pointed out before, a handed object is handed relative to an embedding space. In spacetime, that amounts to saying that a spacelike slice of spacetime can include a spatially handed slice of spacetime worm, and that the handedness is a matter of the character of the whole spacelike slice containing it. This means that the concept of handedness can be used to characterise not just space, but spacetime. The following can be said of Minkowski spacetime: the existence of deeply (*i.e.* affinely, projectively or topologically; in any case, invariantly) handed objects requires that spacetime has a definite character, *viz.* that every spacelike hypersurface is orientable. Thus, the space of every inertial frame is orientable too. Every deeply handed object has its character preserved in the space of every frame and that character is frame invariant.

Given that the geometry of spaces in SR is Euclidean, these definitions work and they are rather simple. Two spacetime objects are *absolutely enantiomorphic* if no continuous similarity mapping in any frame allows the space occupied by the one to be mapped onto the space occupied by the other (here, 'enantiomorphic' is a non-mertical version of 'incongruous counterpart'). An object is *absolutely handed* if for each frame there is a space into which it (or its space) can be mapped by reflective and similarity mappings but not by continuous similarity mappings.

These considerations can be applied also to the case of Newtonian spacetime without Newton's absolute space. This case enables a sharp distinction to be made between the space which is arrived at by Kant's argument, and Newtonian absolute space. This in turn justifies our use of the term 'realism' rather than 'absolutism', as discussed earlier. Newtonian spacetime without absolute space is, like Minkowski spacetime, a collection of spacetime points. It has however a different metrical structure, and the collection of spacelike slices associated with an internal frame differ from the SR case. In particular, incongruous counterparthood is frame-invariant as it is not in SR. Consequently, invariant incongruous counterparthood could not establish *absolute* space. It is thus a misunderstanding of the scope of Kant's argument to think that it could. But there is nothing about the case to suggest that the spaces of Galilean frames (Newton's relative spaces), in which incongruent counterparts are to be found, are not real. They are parts (or subsets) of spacetime; they therefore consist in spacetime points. The abolition of absolute space is not inevitably the abolition of space. Nor is it the triumph of relationism.

Whether or not appropriate objects are handed in spacelike hypersurfaces of the spacetimes of General Relativity can also be used to characterise these spacetimes. Sometimes, that style of approach can be just as illuminating as in previous examples. However, this seems often to result in laborious, inelegant descriptions. Riemannian spacetimes are so various, and the spacelike hypersurfaces, when they can be projected out, may vary so much even within one spacetime, that it seems not very illuminating to try to characterise them by means of their spaces and in terms of their orientability. But the style of argument we have been investigating here does not fail to yield its conclusion. Only its pragmatic or heuristic advantages are lost in such generalisation.

Department of Philosophy
University of Adelaide
G.P.O. Box 498
Adelaide
S. Australia 5001

BIBLIOGRAPHY

Hinckfuss (1975). I.C. Hinckfuss, *The Existence Of Space and Time,* Oxford, 1975.
Mortensen and Nerlich (1978) C. Mortensen and G. Nerlich, 'PhysicalTopology', *The Journal of Philosophical Logic* 1978, 209–223.
Nerlich (1973). G. Nerlich, 'Hands, Knees and Absolute Space', *The Journal of Philosophy*, 1973, 337–51.
Nerlich (1976). G. Nerlich, *The Shape of Space*, Cambridge, 1976.
Quine (1960). W. V. O. Quine, *Word and Object*, the M.I.T.Press, 1960.
Rindler (1977). W. Rindler, *Essential Relativity*, Second Edition, Springer-Verlag, 1977.
Sklar (1974). L. Sklar, 'Incongruous Counterparts, Intrinsic Counterparts and the Substantivality of Space', *The Journal of Philosophy*, 1974, 277–290.

Australasian Journal of Philosophy
Vol. 63, No. 1; March 1985

THE LIMITS OF CHANGE

Chris Mortensen

1. *Introduction.* What is an object's state of motion at the instant when it begins to move? If it is moving, when was it motionless? If it is motionless, how can it ever begin to move? If it is both moving and motionless, do we not have Hegelian contradictions in nature?

This problem has been discussed by Brian Medlin and by Charles Hamblin[1] I will have something to say about their solutions later. I will be concerned to show that the problem is an interesting one, one not to be set aside by a brief gesture in the direction of modern analysis and set theory, in fact one generated by their fruitfulness in physical theory. Nor is the problem one to be dissolved by analysis of our ordinary concept of motion, as we will see. One quick *caveat* however: I want to set aside quantum theory. The problem considered in this paper is about discontinuities, and quantum theory is all about discontinuities. Nevertheless, it is easier to make out the problem in a world not bedevilled by the complexities of quantum theory. Hence, the physics of this paper is classical.

2. *The Problem.* Both Medlin and Hamblin formulate the problem as one concerning the application of the putatively two-valued concepts of ordinary language, such as *motion* or *rest*, to a world of continua, especially continuous time and continuous space. Conceived in this way, it is not so surprising that Medlin wants to solve the problem by saying that at the instant of change the system is both in motion and at rest, and reject the apparent implication that such is a contradiction. Keith Campbell also thinks that if velocity could change discontinuously, then at the instant of change the system would be both in motion and at rest, though unlike Medlin he finds such a consequence uncongenial.[2] But it is easy to feel unmoved by this way of posing the problem. After all, if it *were* inconsistent to say that the system was in motion and at rest, we might feel that it was so much the worse for those concepts. What is needed to make the problem sharp is *a formulation in terms of the concepts of the best available physical theory*, which nevertheless still threatens a contradiction. Without such, it is difficult to feel that the attempt to save

[1] Brian Medlin, 'The Origin of Motion, *Mind* 72 (1963), pp. 155-175; Charles Hamblin, 'Starting and Stopping', *The Monist* 53 (1969), pp. 410-425; see also his 'Instants and Intervals', *Studium Generale* 24 (1971), pp. 127-134, and 'The Logic of Starting and Stopping', in C. Pizzi (ed.), *La Logica del Tempo*, Torino, Moringhieri, 1974.

[2] Keith Campbell, *Metaphysics: An Introduction*, Encino, California, Dickenson, 1976; p. 89.

our ordinary concepts is an interesting exercise.[3] This is why, even though Medlin is patently aware that the problem is quite a general one, it is curious that he persists with his solution. Even more curious is that he concedes in the end that his solution is less than general, and does not apply to parallel cases.

Parallel cases abound. Let q be any physical quantity which changes with time, and suppose that it is unchanging up to $t=0$ and then changes its value thereafter. What is the state of its rate of change, dq/dt, at $t=0$? To fix on a case directly related to motion and rest, let the velocity $v=dx/dt$ of a body moving along the x-axis be zero up to $t=0$, and be given by $v=t$ thereafter. Then velocity changes continuously, but acceleration dv/dt or d^2x/dt^2 is given by 0 up to $t=0$ and 1 after $t=0$. What is the value of dv/dt at $t=0$? From the point of view of the equations of physics, it does not seem to matter whether we say 0 or 1. But what is happening in nature at $t=0$? If it does not matter what we say, is nature arbitrary or indeterminate?

Nor is the problem confined to change in the time dimension, as Medlin notes. Consider the surface of a material object, and let the x-direction be normal to its surface, with $x=0$ at the surface. Then is there matter at the point $x=0$ and empty space for all $x>0$? or matter at all for $x\geqslant 0$? It might be arbitrary what we say, but what is there in nature?

3. *Hamblin and Intervals.* Hamblin is aware of the arbitrariness of saying that a discontinuously changing quantity has one value rather than the other at the instant of discontinuous change. His solution is to sidestep the problem by appeal to time structured only as a collection of temporal intervals rather than instants. As he notes, this move will of necessity involve appeal to a certain three valuedness: a quantity may have a value throughout an interval, or not have that value throughout the interval, or have it throughout some subintervals and not have it throughout others. But because instants are not allowed, the problem of discontinuous change at those instants does not arise.

The extent to which it does not arise, however, is the extent to which physics and common sense, which agree that change can take place *at* 12 noon, are rejected. A very simple case, on which physics and intuition agree I think, is a force function which varies linearly with time, given by $d^2x/dt^2=2t$. With suitable adjustment of constants, velocity $=t^2$, and so is *instantaneously* zero at $t=0$. Intuition seems to agree that instantaneous motionlessness is possible. (Consider braking to stop at the stop sign and instantaneously accelerating away, or an object thrown upward instantaneously pausing before falling downward.) Perhaps common sense is not worth saving, but physics is notably

[3] Thus Hamblin: 'Our problem is concerned with the *systematic* description of physical phenomena. It arises from the clash of the continuous-variable language of Mathematical Physics with the discrete two-valued language that we would like to make work as an alternative. *The mathematical description appropriate to an accelerating object does not raise problems for us*; but we would like in addition, to be able to apply the two-valued predicate 'in motion', or its negation, to anything at any time, and this is the project that runs into difficulties.' ('Starting and Stopping', *op. cit.*, p. 414; second emphasis mine.) To the contrary, I claim that it is the mathematical description where the real problems lie.

successful. So Hamblin would have to show, what he does not show, how enough of the theory of differential equations to describe such physical systems can be constructed within his assumptions. Without this, we should wonder how much physics has to be sacrificed in the interest of saving Hamblin's solution. I propose, then, to remain in the framework of time as a continuum of instants. At least, our discussion will show the options for solutions in that framework.

4. *Arbitrary Solutions.* I do not propose to rehearse various bad answers to the problem. Several of these are detailed by Medlin and by Hamblin. Instead, I want to focus on what I take to be the key issue, namely that of avoiding an *arbitrary* solution to the problem. By an arbitrary solution I mean one which assigns a particular value to some variable but can give no reason for preferring it to another, incompatible, assignment. Consider for example our earlier case of velocity $v=0$ up to $t=0$ and $v=t$ thereafter. To say that its acceleration is 1 for $t>0$ and 0 for $t\leq 0$ is arbitrary, if no reason can be given for preferring this to the description $d^2x/dt^2=0$ for $t<0$ and 1 for $t\geq 0$. (By a 'reason', here, I do not necessarily mean a 'physical' or 'physically significant' reason; a reason in the mathematics might suffice, or some pleasing metaphysical reason.) Now it seems to me that both of these solutions are arbitrary, though it may be that there is a pleasing metaphysical reason of which I am unaware for distinguishing between them. One reason for disliking arbitrary solutions, is the question of physical determinism. A variable's assuming at a time a given value rather than another is an event. So if every event is causally determined, it is causally determined that discontinuously changing variables have one value rather than another. But if it is arbitrary which value is taken, then either our best theories lack the capacity to describe the causal determinants of events, or such events escape the causal net of the universe.

There are however, at least four nonarbitrary solutions to our problem, all of which are either unnoticed or barely noticed by Medlin and Hamblin. I will discuss their strengths and weaknesses in the next few sections.

5. *The Paraconsistent Option.* Both Medlin and Hamblin flee from contradiction. They offer consistent solutions to what they think would otherwise be a 'paradox'. But Zeno, and following him Hegel, thought that motion was inconsistent (though they used this for different ends). Many other philosophers have been puzzled by change, motion and time. It is a *nonarbitrary* solution to say that when we have a discontinuous change in a physical quantity, the quantity takes both values (or perhaps all intervening values as well). It is nonarbitrary to say that a material surface is both occupied and not occupied by matter.

The last few years have seen philosophers beginning to take seriously the thought that the world might be inconsistent. If the world is inconsistent, then any logic (such as modern two-valued extensional logic) containing the law *Ex Falso Quodlibet* (From A and not-A to deduce every proposition) must be abandoned, since plainly not every proposition is true in our world.

But logic teachers have to labour hard to convince their first-year students of *Ex Falso Quodlibet*; so hard that it seems that sometimes this is the sole source of their own conviction. Alternative logics with natural motivations now exist.

Nonetheless, the strongest cases for inconsistency seem to exist in the logico-mathematical area. Various of the paradoxes, such as The Liar, or Russell's Paradox, call out for an inconsistent (or 'paraconsistent') solution, because of the unsatisfactoriness of alternatives.[4] In my opinion a compelling case for inconsistency in motion has yet to emerge, though it would be foolish to rule it out of court without argument. Nevertheless, nobody accepts inconsistencies easily. Perhaps this can be put in a slogan: *do not multiply contradictions beyond necessity*. This dictum implies that we should only go for the present solution if none of the others are satisfactory. I offer it to the unconverted in that spirit. Of course, if the other solutions are all unsatisfactory, and if we have an exhaustive list here, then there is inconsistency somewhere.

6. *Indeterminateness*. It is perhaps not entirely obvious that a nonarbitrary solution to the problem is to say that the discontinuously changing variable has no value at the instant of change. Philosophers have usually felt easier about accepting truth-value-less propositions than propositions which have both truth values, so to that extent the solution is to be preferred to the paraconsistent option.[5] Notice that indeterminateness, just because it is a nonarbitrary solution, does not seem to get into problems with physical determinism. If the variable assumes no value at t, then we do not have to worry about its being causally undetermined which value is taken. Moreover, some support might be obtained by considering the time derivative of any discontinuous quantity. The derivative of such a function is not defined at the point of discontinuity. It is, so to speak, indeterminate. So we seem to be stuck with indeterminacy somewhere.

A reply would be to claim that the derivative at a discontinuity is infinite. I do not wish to legislate that nature can have no infinite magnitudes. It does, however, seem to be a considerable cost to pay to avoid indeterminateness. If we do not want to pay the cost, then indeterminateness looks more reasonable.

On occasion when I have posed this problem to physicists, the reply has been 'it doesn't matter what you say'. This covers a multitude of sins. I take it that it often amounts to a claim of indeterminateness, since if nature is not indeterminate at the instant of change, then it surely *does* matter what you say. Of course, sometimes it means that it does not make any *experimental*

[4] The literature on paraconsistency is voluminous. Perhaps the best survey is in G. Priest and R. Routley (eds.) *Paraconsistent Logic* (forthcoming). See also R. Routley 'Ultralogic as Universal' in his *Exploring Meinong's Jungle*, Australian National University, 1980, pp. 892-959; or N. Rescher and R. Brandom, *The Logic of Inconsistency*, Oxford, Blackwell, 1979.

[5] To be sure, in the presence of the intuitively natural laws of De Morgan and Double Negation, the Law of Excluded Middle ~AvA is equivalent to the Law of Noncontradiction ~(A&~A); so perhaps this easy feeling is misplaced.

difference what you say. True, but unless one has an instrumentalist view of physical theory, that only the experimentally detectable is real or determinate, then it goes no way toward solving our problem. Perhaps it is a confession of lack of interest.

However, it is harder to choose the present option when the surfaces of bodies are considered. What sort of a surface is it that is neither occupied nor not occupied by matter? I do not know how to prove that this could not be so, but intuition seems to protest at it.

7. *All Change is Continuous.* The description of the problem requires that some variable, or its time derivative, or . . ., changes discontinuously with respect to time. But what reason have we to suppose that any such change occurs in nature? Perhaps all changes, their rates of change, and so on, are continuous. If this is so, then in our world the problem does not arise. Notice that this solution is nonarbitrary: velocities, accelerations, and their *nth* derivatives all have definite values at $t=0$.[6]

Hamblin gestures at this strategy. His argument against is not very convincing, however: 'If no derivative ever changes discontinuously, nothing ever changes'. ('Starting and Stopping', *op. cit.* p. 412). This seems to be quite straightforwardly false. There does not seem to be any reason why all physical quantities couldn't be described by C^∞ (infinitely differentiable) functions of time. For example, all physical quantities might be related to time by cosine, sine, or exponential functions. I do not mean that this is how the world is; only that such a world is a counterexample to Hamblin's quoted claim.

Intuition suggests that it is a contingent matter whether all change is continuous. So someone might feel disappointed that the present solution sidesteps the issue by failing to say what *would* happen if there *were* discontinuous changes. The point is, however, that the issue of whether a solution is arbitrary is bound up with how much we take the problem to arise in *this* world. We feel that there must be a right answer for our world, that one of the arbitrary alternatives is the correct one in nature; and we feel that it is unsatisfactory if we can give no reason for why we or nature should opt for one. In other worlds, however, we are free to say what we like, so to speak. In some of them there might be always a last moment, never a first, for discontinuous changes. In others, the reverse. What we say about other worlds varies with our descriptive whim. This point has to be handled with some care, since the assumption of classical physics at the beginning of the paper was an explicit assumption of a false theory. I take it, though, that the fact that classical physics was long believed to be true and is even now sufficiently close to the truth to be used as a good approximation of the behaviour of middle-sized objects, lends interest to what might otherwise be thought to be too speculative an exercise in metaphysics. In any case, quantum theory does seem to be involved in discontinuous changes *via* the 'collapse of the wave packet'; so conceivably we have a real problem about our own world. Of course, it might not be a contingent matter that all change

[6] Unquestionably, many physicists can be found to go for this option. Boscovich seemed to. (See Campbell *op. cit.* p. 89).

is continuous (then, too, quantum theory might break down). If all the other solutions discussed here have insuperable difficulties, then we have an argument for the conclusion that necessarily all change is continuous. That would be an interesting piece of metaphysics.

8. *Surfaces and Fractures in a Continuous World.* Nevertheless, there does seem to be a difficulty when we reflect that, as we saw before, any *general* solution must apply to all change with respect to any variable, not merely time. It is not easy to see how to apply it to the surfaces of material objects, for example. The reply that material objects, being made of atoms, have bumpy surfaces, only shifts the problem back to the surfaces of atoms. A special case of this difficulty is the 'fracture problem' (which I owe to Medlin in conversation): if any material object, such as an atom, is a plenum, then, in thought at least, it could be fractured. If space is continuous, one surface of the fracture must be topologically open, the other topologically closed. But no reason can be given for which should be closed and which open. Yet unless matter is composed of unextended points, at least some of it must be a plenum. In this section, we will see how to deal with these problems in a C^∞ world.

In such a world, there needs to be some C^∞ quantity responsible for the application of the concept *being occupied by matter.* Let us speculate how this could come about. Suppose that there is some C^∞ variable, call it *density d*, with the property that $d = 0$ in empty space, and inside a material surface density is always greater than zero but continuously approaches zero as the surface is approached from the interior. (We need not be concerned at the possibility of a reduction of density to mass/volume provided that these are also C^∞). Approaching the surface from the inside, matter becomes more and more attenuated as it were, like the edges of a mist. It follows that surfaces in this world would all be topologically open.

This still leaves the fracture problem. A discontinuous fracture at a point where $d > 0$ is a discontinuity logically incompatible with a C^∞ world; or to put it another way, discontinuous fractures cannot leave us with both new surfaces being topologically open. But the problem can be met provided that fractures are all continuous. And it is not difficult to describe density functions which are also functions of time, and which will produce breaks in surfaces in a finite time. So fractures in a C^∞ world which leave all surfaces topologically open are a possibility.

For the technically minded, let us digress briefly to demonstrate this. We restrict ourselves to two-dimensional objects, but the generalisation to higher dimensions is straightforward. Let density be given by the C^∞ function $d = \sin^2 x \sin^2 y(1 - \sin^2 x \sin^2 t)$ where x and y are spatial co-ordinates and t is time. We restrict ourselves to the range $0 \leq x, y, \leq \pi$ and $0 \leq t \leq \pi/2$. The function $\sin^2$ is chosen over sin because it does not take negative values. At $t = 0$, the body is a square whose surfaces are $x = 0$, $x = \pi$, $y = 0$, $y = \pi$. Density is zero at the edges and rises to a maximum of 1 in the centre $(x, y) = (\pi/2, \pi/2)$. For $0 < t < \pi/2$, the density gradually forms a 'hollow' along the line $x = \pi/2$. Parallel to the x-axis, in other words, density begins to look like the two humps on a camel. Imagine the body slowly being torn apart, the exterior surfaces being the same, but density is beginning to decrease around the fracture line $x=\pi/2$. Parallel to the y axis, density is still a single humped function, but its maximum height diminishes to a local minimum along the fracture plane. At $t=\pi/2$, the fracture is complete. The original body has separated into two with a hairline fracture along the line $x=\pi/2$. Each of the two new bodies is a rectangle with sides in the x-direction of length $\pi/2$ and y-directional sides of length π. Density for each of the bodies rises from zero at the edges to a maximum in their centres, just as at $t=0$. The crucial property of topological openness of surfaces is thus preserved.

I conclude, then, that surfaces and fractures can be dealt with in a natural way in a world in which all change is continuous. Let us move on to a fourth proposal to deal with the problem of the point of change. It has the merit of being more general than the one in this section.

9. *All Real Change is Continuous.* It might be held that functions other than C^∞ functions truly describe some changes, but that when differentiation proceeds far enough for a discontinuity to be reached, we have arrived at a quantity with no physical reality. Physically real changes, on the other hand, are always continuous.

Consider for example, a particle moving according to: for $t \leq 0$, $x=0$; and for $t>0$, $x=t^3/6$. Then dx/dt and d^2x/dt^2 are both continuous, but d^3x/dt^3 ($=0$ for $t \leq 0$ and 1 for $t>0$) is discontinuous. The proposal of this section would declare d^3x/dt^3 not to be a physically real quantity. We can expand this further along the lines of various modern theories of universals.[7] It is arguable that physical quantities are universals. But it should not be thought that the mere existence of a true linguistic description requires one to suppose that there is a universal underlying it, responsible for its truth. Indeed, an unlimited principle of abstraction of the sort: to every predicate there corresponds a universal possessed by exactly those things of which the description is true, cannot be consistently maintained. It leads to Russell's Paradox. Applying that idea here, we can say that even though certain linguistic descriptions, such as 'acceleration' and 'time', might have universals corresponding to them, it does not follow that any description constructible out of those descriptions, for instance 'd^3x/dt^3', also has a universal. It would solve our problems if such a quantity were not physically real: when d^3x/dt^3 changes, no change in any real quantity takes place in the world, except in those real quantities out of which d^3x/dt^3 is constructed. Hence, there is no need to bother about discontinuities in d^3x/dt^3. Nothing real changes, so it does not matter what you say. I do not mean to single out d^3x/dt^3 as a general example of a quantity which is not physically real, only for the purposes of discussing the above equation of motion. The point is that it is open to us to describe processes in nature by non-C^∞ functions, and declare discontinuous derivatives not to be physically real.

This proposal has some constraints on it. One is that we might believe

[7] For example, D. M. Armstrong's. See his *Universals and Scientific Realism*, Cambridge, Cambridge University Press, 1978. Only a small part of his theory is necessary here though.

that certain physical quantities such as *force* and *mass* are real, in virtue of the fundamental rôle they play in our explanations. What would we then say about a position function like: $t \leq 0$ iff $x = 0$ and $t > 0$ iff $x = t^2/2$, where d^2x/dt^2 was discontinuous? The answer, obviously, is that if all real change is continuous, and if force and mass are real quantities, then it follows that such position functions do not describe any actual processes in nature. So the present proposal does not say that continuity is sufficient for physical reality, only that discontinuity is sufficient for lack of physical reality.

This strategy has the advantage over that of the previous section that it does not need to commit itself to the strong thesis that only C^∞ functions are permissible descriptions of dynamic systems. This seems to me to be behind another response to the problem I have sometimes heard from physicists: 'Yes, but is the discontinuity physically significant?' I take this to mean that a discontinuity in a function involved in basic laws of nature would be physically significant and so generate a problem; but that discontinuities in *some* functions which might be produced from 'physically significant' functions (by, say, differentiation) need not be physically significant and so need not be problematic. They would be, so to speak, 'mathematical fictious'.

10. *Surfaces and Fractures as Discontinuities.* We saw a couple of sections back how to deal with surfaces and fractures in a C^∞ world. Needless to say, when we move to a more general case as in the last section, that solution is still available to us, since it postulates a C^∞ density function and any such function is allowed in the generalisation. In a way, though, surfaces and fractures are intuitively discontinuous structures.[8] It would be interesting, then, to see what can be said about them if we dispense with a density function in favour of discontinuous distributions of matter.

A way into this is *via* Boscovich. Boscovich thought that matter consisted of atomic material points, having position but no extension or magnitude, surrounded by force fields extending out to infinity. Such fields prevent any pair of material points from touching by becoming arbitrarily large sufficiently close to the unintended atoms. Keith Campbell points out that this solves the problem of what to do with surfaces: they are just concentrated but spatially discrete distributions of materially occupied points.[9] There is thus no issue about whether they are topologically open or closed. Again, fractures are no problem. A point cannot be fractured. Fractures are no more than an attenuation of the complex interlocking fields surrounding atoms (and as such are continuous).

The cost of this solution is that matter is ultimately unextended. That might be so, but it would be desirable if a nonarbitrary account of surfaces was available for an ontology of spatially extended atoms. Is there in this ontology some reason for preferring, say, closed surfaces to open? I think that there is, or at least there is if atoms are surrounded by force fields as in Boscovich's account.

[8] For instance, in explaining topology, transformations which are continuity-preserving (homeomorphisms) are often intuitively contrasted with tearing or fracturing.

[9] Campbell, *op. cit.*, p. 90.

Suppose that extended material atoms are surrounded by force fields with the following properties. They have no value at points occupied by matter, have a value at all points in empty space, and eventually become arbitrarily large at sufficiently small distances from the surface. It does not matter here whether these fields are repulsive (as in Boscovich) or attractive; the same argument works for both. Now if there were no matter *at* the surface because it was topologically open, then we would have the problem of saying what is the value of the field at the surface. Continuity suggests that the value of the field at the surface would be infinite. But now let us appeal to the principle that infinite values of real quantities in nature are to be avoided where possible.[10] The only way to do so, is for atoms to have topologically closed surfaces. Then the field simply approaches infinity as it approaches the surface, but never gets there. Thus, given such fields, the metaphysical principle that there are no infinite quantities issues in a nonarbitrary account of the topology of surfaces.

But then fractures are a problem again, for how can a plenum fracture into two closed surfaces unless matter appears from somewhere? The only solution, I suggest, is to say that atoms cannot be fractured. Recall how the fracture problem was first introduced: an extended object can be fractured in thought, so what nonarbitrary account of the topology of the fracture can be given? But of course it does not follow from the fact that we can *imagine* that an extended atom be fractured, that it can be fractured in reality. Indeed, to call it an 'atom' is already to imply that it is a simple thing. It is certainly consistent with an ontology of spatially extended matter that there be undecomposable atoms, furthermore. Provided, then, that an extended material thing cannot be fractured, nature need not be said to be arbitrary about which topology its parts are to have post-fracture.

Now while this gives a solution for atoms, it does not give a solution for extended nonatoms. By definition, if any continuous nonatoms exist, the fracture problem arises for them and is not be be solved in this way. But here I think we can get a clue by considering the time reversal of fracturing, namely joining. If surfaces of atoms are topologically closed, how can they join up to form a continuous whole? The answer would seem be be that, logically, they cannot. Their topology prevents it. Two closed surfaces cannot be made contiguous. Furthermore if atoms cannot be joined up, it seems reasonable to say that a thing cannot be made up of joined atoms. It follows that continuous extended non-atoms cannot be fractured either, since there are none. Moreover, this consideration provides reinforcement for the earlier suggestion that atoms cannot be fractured.[11] Were atoms to be fractured into objects with closed surfaces, then the time reversal of the process would be a joining of closed surfaces, and that it impossible. Thus, in a world where

[10] The principle has intuitive appeal, as noted earlier. In support, it might be argued that functional relationships between infinite quantities amount to an inability of the mathematical framework of the theory to model the causal relationships in nature, and thus theoretical inadequacy.

[11] Indeed, this follows logically for any universe in which the laws are time reversible: a process is possible iff its time reversal is possible.

extended things have closed surfaces, it is impossible that there be fractures, joinings, and continuous extended nonatomic objects. Needless to say, this does not mean that familiar middle-sized objects could not be broken. Their breakings would simply be through the empty space that surrounds atoms. They would be pullings-apart of agglomerations of atoms, but not real fracturings of continua. Of course this is how we break middle-sized things in our world. You don't have to split atoms to chop wood.

11. *Conclusion.* It is clear that the attempt to find nonarbitrary answers to variants of our general problem has not issued in a blanket solution. Different things are to be said when different assumptions are made. In a continuous world changes in motion do not seem to be problematic. With a continuous density function as well, surfaces are topologically open and fractures and joinings are continuous. The same can be said for a world where all real quantities are continuous. When we dispense with continuous distributions of matter in favour of discontinuous matter/nonmatter surfaces, things become more complex. But under the assumption of Boscovichian fields surrounding matter there is a 'pleasing metaphysical reason' to conclude that surfaces are topologically closed. With or without that reason, if all surfaces are topologically closed, then we seem to be forced to say that matter consists of unfracturable and unjoinable atoms.

It is harder to see what to say when these assumptions are relaxed, for instance in a world where real discontinuities abound. What if the world is such that anything can be fractured given enough energy? Perhaps that would be an indeterministic world whose indeterminateness was not detectable by experiment. But then it might also be possible to describe a physics which causally differentiated between two kinds of surfaces, closed and open, thus avoiding arbitrariness. I do not know how to do this, but I am unable to prove that it cannot be done. Perhaps that is the best place to stop talking.[12]

University of Adelaide

Received December 1983

[12] I am indebted to comments by Philip Cam, Keith Campbell, Alan Lee, Graham Nerlich, Jack Smart, Joseph Wayne Smith, and a referee of this *Journal*.

THE CANADIAN JOURNAL OF PHILOSOPHY
Volume 16, Number 4, December 1986, pp. 713-722

Explaining Existence

CHRIS MORTENSEN
The University of Adelaide
Box 498, G.P.O.
Adelaide 5001
South Australia

I

The problem of why something exists rather than nothing is doubtless as old as human philosophising. Of comparable antiquity is the observation that one cannot hope to explain why something exists rather than nothing by appealing to the existence of something else, on pain of vicious circularity.

In this paper, I distinguish between the question of why anything exists, and the question of why particulars exist. These two questions are equivalent only if the only things that exist are particulars. Certainly many have held that universals as well as particulars exist.[1] I take it here that there is a *prima facie* distinction between universals and particulars. It follows that the former question is *prima facie* more

1 See e.g. D.M. Armstrong, *Universals and Scientific Realism* (2 vols.) (Cambridge: Cambridge University Press 1978).

general than the latter. I will initially concentrate on the latter, taking a hint from some recent theorising about the physics of the Big Bang. I will argue that, properly understood, there is a sense in which the existence of particulars might be explicable. That, it seems to me, represents some progress on the problem. For instance, it is arguable that when people ask why anything exists they have in mind the question of why particular things, or one big particular thing such as the spatiotemporal universe, exist. Insofar as that is the problem, I suggest that we can make inroads into it. I then go on to ask how these considerations might be applied to the more general question of why something exists rather than nothing. I will suggest that there are several ways the world might be, in which even this question might have an answer of sorts.

II

Current intense levels of theorising about the Big Bang continue to push explanation closer to $t=0$. Recently the physicist Edward Tryon has proposed a theory of the Big Bang according to which it begins as a 'quantum fluctuation' out of nothing. Tryon has described his theory as a theory of creation *'ex nihilo.'*[2] Now *'ex nihilo'* is a loaded phrase for philosophers, conjuring up debates about whether something could come out of nothing unless God created it. I do not think that it is necessary to confuse philosophical readers with the technical details of Tryon's proposal. It should be said, though, that inspection of those details reveals that the initial quantum fluctuation takes place in otherwise empty space and time.[3] Now it has been argued that empty space and time, or spacetime, are particular existing things.[4] If that is true,

2 Edward Tryon, 'Is the Universe a Vacuum Fluctuation?' (hereafter UVF) *Nature* 246 (1973) 396-7; also 'What Made the World?' (hereafter WMW) *New Scientist* 1400 (1984) 14-16.

3 Tryon: '... some pre-existing true vacuum,' WMW, 15; or '... the vacuum of some larger space in which ours is embedded,' UVF, 397. It is fair to say that much of the physicist's interest in such a theory is in the accounts of how a big universe could come out of a little quantum fluctuation and of how conservation principles can be held true, which do not concern us here.

4 See e.g. Graham Nerlich, *The Shape of Space* (Cambridge: Cambridge University Press 1975).

then Tryon's theory is not a theory of particular things beginning out of literally *nothing*, as the phrase *'ex nihilo'* suggests.

Mind you, the situation is complicated by what I take to be an implied relationism about space and time in Tryon.[5] Briefly, relationism is the doctrine that space and time are mere constructs out of spatiotemporal relations between particular material bodies and events. One consequence of this doctrine is that unless some of the latter exist, space and time cannot. Thus, if relationism were true, Tryon's theory would be of a beginnning literally *ex nihilo*. But I do not believe that relationism is true, as has been argued elsewhere.[6]

III

Even so, Tryon's theory provides us with the opportunity for speculation. So let us ask what kind of theory there could be which gave an account of how particular things and events exist or occur, in terms other than by postulating the existence of other particular things or events. To avoid the complicating factor of relationism, let us speculate about what a probabilistic theory of Tryon's kind, but which lacks commitment to pre-existing space and time, could do in explanation of particularity. So let us try to postulate a theory wherein all particulars begin to exist a finite time ago, and wherein there is some initial state which has some likelihood in virtue of some probabilistic laws such as those of the quantum theory. It goes without saying that the present exercise is speculation; I am not suggesting that it is true.

I do not know how to describe this possibility in the kind of detail physicists go in for. But obviously it would be desirable, if possible, to supply more detail about the kinds of laws which would give a 'physics of nonexistence.' We will proceed in two stages. First, we will con-

5 In addition to the use of *'ex nihilo,'* we have, for example, '... some pre-existing true vacuum, *or state of nothingness*,' WMW 15, emphasis mine.

6 See Nerlich, Ch. 2; also his 'Hands, Knees and Absolute Space,' *The Journal of Philosophy* 70 (1973), 337-51; also Chris Mortensen and Graham Nerlich, 'Spacetime and Handedness,' *Ratio* 25 (1983) 1-13; and 'Physical Topology,' *The Journal of Philosophical Logic* 5 (1978) 209-23. Note, too, that it is not apparent how to make Tryon's own words consistent here: how in a state of genuine nothingness could anything pre-exist?

sider the possibility of eliminating a pre-existing space but not time, so that the resulting theory might be regarded as explaining the existence of both space and also events in it. Then we will go on to look at the problem of extending the account to eliminate pre-existing time as well.

In General Relativity, there are what is known as the 'vacuum field' solutions to the field equations. Informally, these say that in a universe with no matter and energy, spacetime still has a definite metrical structure, in some cases that of Euclidean flatness. As has been pointed out by Grünbaum and others,[7] the existence of the vacuum field solutions suggests that relationism is false, in that in the absence of matter, spacetime would continue to be an existing thing. We should avoid any theory like that here, because of the previous complication that space, or time, or spacetime, are arguably particulars. But avoiding it does not look to be in principle impossible. We can suppose that there is some quantity which is a function of time and which measures the distribution of energy or matter (call it mass $M=M[t]$), and if M takes the value zero the theory says that the metrical structure of space (but not time) is *undefined*. This seems a reasonable way of saying that under the condition $M=0$, space would not exist.

So, let us imagine that our laws include the consequence that if $M=0$, then events $E_1, E_2, \ldots$ have probabilities $p_1, p_2, \ldots$ respectively. My claim is that if the universe 'begins' with event E_1, then this law will explain the occurrence of E_1 as well as anything is explained in the quantum theory or in Tryon's theory. Before getting to that, however, there are a number of complications to explore. One is this. Arguably, the events $E_1, E_2, \ldots$ would be spatial events, in the sense that if any of them occurs at a time then space exists at that time. So we might imagine that the condition $M=0$ obtains for an interval of time and then a 'quantum fluctuation' occurs, and space and matter are born. That seems to me to be an intelligible possibility. But the form of the above law does not require that the possible events E_1, E_2 have to occur *after* the condition $M=0$. So is it a possibility that we have a law '$M=0 \rightarrow E_1, E_2, \ldots$ have probabilities $p_1, p_2, \ldots$' where the E_1, E_2 could *simultaneously* with $M=0$? It is, if the 'events' $E_1, E_2, \ldots$ did not

7 A. Grünbaum, 'The Philosophical Retention of Absolute Space in Einstein's General Theory of Relativity,' in J.J.C. Smart, ed., *Problems of Space and Time* (New York: Macmillan 1964) 313-17.

need the existence of space; or equivalently, did not need $M \neq 0$. That would be possible, for example, if then E_i were conditions on the *derivative* of M, say '$M=0 \rightarrow dM/dt=x_1,x_2,\ldots$ with probabilities $p_1, p_2,\ldots$'.[8] Now this form of law allows for various possibilities. If at some time *t*, we have $M(t) = 0$ while the quantum fluctuation $dM/dt=x, \neq 0$ occurs, then for an interval of time after *t*, $M \neq 0$. Thus, the history of the universe for times when space and matter were present, would be the set of times $\{t' : t < t' \leq now\}$, which has no first instant. This is of course, topologically possible with a continuous *t* variable. Again, one of the finite possibilities when $M=0$ at some time *t* might be $dM/dt=0$ also. (Perhaps even this *has* to be one of the probabilities.) Then, if $dM/dt=0$ comes off, we would have the situation described earlier, of $M=0$ for a stretch of time after *t* until one of the other possibilities $dM/dt \neq 0$ occurs and space begins.[9]

Since there does not seem to be any contradiction in the supposition that laws might be as above, I conclude that at least a pre-existing space is dispensible from an account somewhat like Tryon's. It seems to me that a theory like the present one would give as good an explanation of the existence of space and particulars (other than times) as any in probabilistic physics. The radioactive decay of a single atom is not explained in current theories via prior sufficient causal determination. But it is explained nonetheless, to the extent that we demonstrate how it is governed by laws showing that events of that kind are to be expected, with a precise degree of expectation. A somewhat random universe need not be a chaotic one. If our universe is such that this is the best kind of explanation we can ultimately hope for, then the origin of space and matter need not be worse off in this respect than anything else. In Tryon's words, 'our universe is simply one of those things which happen from time to time.'[10]

8 A mathematically more sophisticated theory would deal with the events E_i and their probabilities using integrals over finite intervals of time, and would also need to give conditions on higher derivatives of M, which would in turn be a tensor quantity; but we will not bother about these complications here.

9 Independently, we can consider whether the whole of time stretches infinitely, or only finitely, into the past. One way, but not the only way, in which the latter could happen, is if $M=0$ at a first instant. Time would then be structured isomorphically with a finite closed interval of the real numbers, $0 \leq t \leq$ now (ignoring future times).

10 Tryon, UVF, 397

Can we get rid of pre-existing time as well? I think that we can. First, let us strengthen the condition $M=0$ to mean that in the absence of mass, neither space *nor* time exists. This does not, of course, amount to relationism, no matter how a relationist's heart might be gladdened if such were the case: the constant conjunction of space, time and matter does not entail that they are identical. Now it seems to me that it could still be a law that $M=0 \rightarrow dM/dt=x_1, x_2, \ldots$ etc. Here, though, we might have to understand the 'obtaining' of the 'initial' condition $M=0$ & $dM/dt=x_1$ (say) somewhat differently, on the grounds there would be no instant 'at' which $M=0$. Imagine that time is finite into the past but lacks a first instant. This would be so if the set of instants corresponds to some finite half-open interval of the real numbers, $0<t\leq$ now. Then we can understand the proposition '$M=0$ & $dM/dt=x_1$' as meaning that the *limit* of M as we go backwards in time (toward $t=0$) is zero, and the limit of dM/dt is x_1; or to put it differently, as *t approaches* zero, M approaches zero and dM/dt approaches x_1. Things would behave in the early part of the universe *as if* dM/dt really were x_1 at an earlier time.

We might in addition want to regard the condition $M=0$ & $dM/dt=x_1$ as a 'mathematical fiction,' in the sense that $M=0$ is not an event which occurs at a time. I am not persuaded that we must do this, however. An argument that we must, would appeal to the necessity of the principle that *whatever obtains, obtains at a time;* and it is not clear how one would argue for its necessity (its mere truth being insufficient to prevent speculation). Furthermore, against such an argument we might invoke a counter-principle which has been widely held in the history of philosophy, that no particulars exist necessarily. For then, since the previous argument would establish that temporal instants exist necessarily if any proposition is necessarily true, then temporal instants fail to be particulars.

In any case, we seem to have been able to dispense with pre-existing temporal particulars. So I suggest that a theory something like Tryon's is conceivable, in which the existence of all particulars is on equal footing in respect of explanation, and in which the probabilistic explanations are of the sort ordinarily available in probabilistic physics. Furthermore, conceivably this is just the right way to deal with the Big Bang. What bothers theorists about the actual instant $t=0$ is, I suggest, that current theories predict a spatial (perhaps even spatiotemporal) singularity there. Tryon trades this in for pre-existing space and time, and matter/energy singularity. The present suggestion does seem to allow for at least a spatial singularity, perhaps even a spatiotemporal singularity; but only in the sense that the usual laws of physics hold on a finite half-open time with no first member. Perhaps then there is less reason to find singularity at the origin of things perplexing.

Robert Nozick[11] also considers the possibility that the existence of something rather than nothing be explained by some kind of probabilistic partitioning of possible states (one state being that nothing exists and so being satisfyingly egalitarian in his sense). He is concerned that any a priori partitioning of possible states for this purpose would be arbitrary and so need explanation, i.e. be inegalitarian. I think that Nozick is not always sufficiently careful about the difference between explaining why something exists, and explaining why a proposition, such as a universally quantified law, is true (though he does address himself to the question of truthmakers for laws). On our present model, our laws would be responsible for the particular probabilistic partitioning that there is. This seems to be standard scientific practice in more limited domains. So, too, on the present model it is the truth of laws which would explain existence, or at least the existence of particulars. Another of Nozick's ideas, that there might be certain 'natural' states, including its being a natural state that nothing exists, can be given a law-based probabilistic gloss: natural = high probability. The present account also avoids a point of Michael Burke's.[12] Burke argues that were time finite into the past with no first element, one should not conclude merely from the fact that every event had an explanation in terms of prior events, that it had been adequately explained why there is something rather than nothing. Whether this is true or not might be disputed; and if it is not, then the present model explains existence in a stronger sense than I have claimed. I am inclined to agree with Burke; but even if he is right, it is being claimed here that the extra explanation is provided by (probabilistic) laws.

11 Robert, Nozick, *Philosophical Explanations* (Oxford: Oxford University Press 1981)

12 Michael Burke, 'Hume and Edwards on Why Is There Something Rather Than Nothing,' *Australasian Journal of Philosophy* 62 (1984) 355-62

IV

We have been considering the possibility of a lawlike explanation of the existence of particular things and events. It will doubtless have occurred to the reader that, whatever the ontic status of a pre-existing space and time in Tryon's account, he is still left with the truth of the laws of probability physics as unexplained. Now someone might confusedly think that hence such laws would 'exist,' so that the existence of some things would remain unexplained. But on the face of it, at least, laws are not the right kind of thing to exist. They are, rather, the kind of thing which is true or false. The latter does not rule out the former, though the claim that laws exist would need an argument. But even if laws do exist, it might be argued that they would not be particulars; so that particularity, at least, remains explained.

A more promising line of argument is this. It might be asked how a law could be true if nothing exists to 'ground' it. We might invoke a slogan: *no difference without a difference in what exists.* If L_1 and L_2 are different sets of laws, and L_1's being true is a different state of affairs from L_2's being true, then some things must exist and have a certain nature in order to constitute the difference.

Here we see the reason for the earlier distinction between explaining particularity and explaining existence in general. For there is a current theory, due to Armstrong, Tooley and Dretske,[13] according to which laws are true in virtue of relations between underlying existing universals. I do not propose to discuss the details of the theory. The difference between Armstrong and Tooley is interesting for our purposes, though. Armstrong's universals are Aristotelian, Tooley's are Platonic. For Tooley, the reason why a law or counterfactual can be true even when nothing exists instantiating the terms of the law, is that the truthmakers for the law are relations between Platonic universals, the mark of which is that they continue to exist uninstantiated. An Aristotelian like Armstrong holds that universals only exist in their instances, and do not exist uninstantiated. My preferences in this matter lie with Tooley, but here I only want to contrast the way the two

---

13 Armstrong; see also his *What Is a Law of Nature?* (Cambridge: Cambridge University Press 1984); Michael Tooley, 'The Nature of Laws,' *Canadian Journal of Philosophy* 7 (1977) 667-98; Fred Dretske, 'Laws of Nature,' *Philosophy of Science* 44 (1977) 248-68.

views apply to our present discussion. If Tooley is right, then the existence of particulars might well be explicable along the lines of this paper; though the existence of something rather than nothing is not, since for the explaining laws to be true, universals must exist. If Armstrong is right, the matter is less clear. It is arguable that Armstrong's theory cannot allow that there be laws which hold when no particulars exist, in which case it does not look like the kind of explanation of particularity we have been considering would be available. But perhaps laws can be true while no particulars or universals exist. Then we would have the stronger result that the existence of anything at all would be explicable in such a universe. Of course I am not saying this is how things are, only how they might be.

So there is a difference between asking why particulars exist and asking why anything at all exists. The former might be answerable along the lines discussed; but even an answer to the latter is not wholly unthinkable if laws could be true consistent with nothing existing. But now we can observe that presumably the explaining laws would be contingent. For both Tooley and Armstrong, for example, the truthmakers for laws are *contingent* relations between universals. So something remains unexplained: why contingent laws are thus and not so. Conceivably, of course, it is incorrect to think that the laws of nature are contingent. The kind of reasoning which led to the Theory of Relativity can be made to look surprisingly a prioristic. If entailment is necessary for explanation, then since necessity distributes over entailment, this course abolishes contingency altogether.

Perhaps it is not essential to be so heroic in the quest for Total Explanation. Nozick considers extensively the hypothesis that ultimate contingencies might be self-subsuming and so in a sense explain themselves. His conclusion seems to be that inegalitarianism cannot be avoided even here. One contingency-retaining possibility not considered by Nozick is as follows. Suppose that the laws of nature are necessary but probabilistic, with a finite probability going to the condition that nothing exists, $M=0$. Then, I suggest, if anything exists it would exist contingently. But on the other hand existence would be explained as well as any probabilistic explanation explains, and by necessary laws. The idea that a probabilistic theory such as the Quantum Theory might be necessary is a kind of dual to the above suggestion that the Theory of Relativity might be necessary. Since presumably necessities would not need explanation, the probabilistic idea has the advantage that it allows both for contingency and also for the explanation of every fact.

I do not regard the necessity of either of these theories a particularly tempting option, it must be confessed. But even here we should not be too hasty in our rejection. If there is any lesson in this paper, it is that explanations might be pushed further back than we hitherto thought.[14]

Received September, 1984

14 Thanks to Michael Bradley and Graham Nerlich for helpful comments.

"Arguing for Universals"

***Revue Internationale de Philosophie*, issue on *Realism in Australia*, 160 (1987), 97-111.**

ARGUING FOR UNIVERSALS

Chris MORTENSEN (*)

1. INTRODUCTION

The problem of universals comes in various forms, often with versions in both a formal mode (problems about predication) and a material mode (problems about being). Thus corresponding to the formal problem of how a predicate can be true of more than one thing, there is the material problem (sometimes called One over Many) of how different things can be somehow the same or similar. Arguments for universals along these lines can be given a scientific realist (or causal realist) twist, as has been pursued by several authors ([1]) : in order to make intelligible how the equations of scientific/causal theory apply to types of situations, indeed uninstantiated types of situations, we need to suppose that there are certain real samenesses or universals behind those types. Again, corresponding to the formal problem of how more than one predicate can be true of one thing, there is the material problem of how a thing can have more than one aspect. Needless to say, while there are correspondences of sorts between material and formal modes, we should beware of concluding too readily with Carnap that formal and material versions of a problem are equivalent.

Arguments for the existence of universals are inevitably tied to (formal or material) variants of the One over Many problem, because universals,

(*) Thanks to David Armstrong, Michael Bradley, Peter Forrest and Graham Nerlich for many stimulating ideas about universals.

(1) See David ARMSTRONG, *Universals and Scientific Realism*, Cambridge, Cambridge University Press, 1978 ; and *What Is a Law of Nature* ? Cambridge University Press, 1983 ; or Michael TOOLEY 'The Nature of Laws', *Canadian Journal of Philosophy*, 1977, 667-698.

whatever their nature, are definitionally the kind of entity which is the same in different things. Undoubtedly the universe contains things similar to one another (predicates have multiple extensions), and however one argues for universals, one must be arguing that some similarities are produced by samenesses. A strong thesis would be : all similarities are sustained by or produced by the existence of samenesses or universals. This does not have to mean that whenever things are similar in a respect, there exists a universal instantiated in all and only those things similar in that respect ; we will presently see the troubles about that. But it should mean at least that underlying samenesses are *responsible* for all similarities ; though the responsibility might be so to speak indirect, for instance *via* the logical operations (2). Less specifically, the relationship between natural language predication and the laws and predicates of physical theory designed to account for all basic causal interaction is doubtless quite complex, but it is a not uncommon view that in the end such basic interactions are responsible for all change or differences, and all stability or samenesses, in the universe. It will be convenient here to avoid questions about this complex relationship, by restricting consideration as much as possible henceforth to basic causal theory itself. We can sum up the above strong thesis in a slogan : no similarity without sameness in what exists, to which it is convenient to give a name, the Similarity Principle. I take it that the principle includes a cluster of more precise versions, both material and formal, which I will not bother too hard to distinguish. Notice, too that the principle is typically intended to apply as much to polyadic predicates as to monadic predicates : what is similar in the situations *a* is *between b and c* and *b is between c and f*, is accounted for in terms of a (ternary) universal *betweenness*, which *a* has a special role in connection with, as does *d*.

One independent argument for universals is a characteristically causal realist one. Successful physical theory *all* quantifies over universals. It is worth having a perspective on just how ubiquitous the practice is. An equation like $F = Gm_1m_2/r^2$ asserts an identity, the expressions on either side being singular terms. As if that were not bad enough, the expression on the right hand is a functional expression, and there is no known theory

(2) Compare Armstrong *Universals and Scientific Realism*, especially vol. 2. Compare also with the thesis that atomic facts are responsible for all truths via the operation of the logical connectives.

of multiplication and division which does not treat them as *operators*; and operators, as is well known, require quantification over that which is operated on. Moreover, the operators here work on dimensioned quantities (5 *grams*) which can be possessed by more than one particular, which is the kind of thing a universal is. Now of course, there might be some wholesale reconstrual of basic causal theory which avoids the use of equations and operators, and I will in Sections Five and Six respectively consider the possible use of set theory and of number theory to do this. For the moment I only want to stress the *naturalness* of the theory of universals. It constantly puzzles me when the objector complains about the *oddness* of entities like universals. On the contrary, universals and their behaviour are *very familiar* to us from scientific theory. In this paper, I will be running two separate lines in favour of universals: questions about the Similarity Principle and related principles, and the just-mentioned obviousness and familiarity of universals from physical theory. I consider the former group of problems first, and return to the latter in Section Six.

2. A Problem

There is a threat to the Similarity Principle and thus any argument for the existence of universals, which standard arguments against universals frequently seek to exploit. This is, that if any similarity obtains between things with demonstrably no universal to be responsible for its obtaining, then it is reasonable to ask why we should ever believe that any similarity is produced by underlying samenesses. This threat can be made more concrete in several ways; we mention two here. First, Russell's Paradox for universals. Suppose that to any similarity there is a universal instantiated in exactly those things which are similar in that respect. The first order expression of this is the schema (∃u) (x) (x instantiates u ↔ Fx). But then, substituting '~ x instantiates x' for the schematic 'Fx', it is a brief and well-known argument to the conclusion that *u* instantiates *u* & ~ *u* instantiates *u*. Second, common regresses employed against universals often focus on the role played by instantiation itself: if to every similarity there is a common universal instantiated in just the right things, then how precisely does instantiation glue together particulars and their universals? If instantiating *u* is a similarity in things, then it needs a further universal to be responsible for that similarity, and so on.

Setting aside currently fashionable paraconsistent solutions to the Russell contradiction, or cumbersome irreflexive hierarchies ([3]) of universals inspired by Russell's Theory of Types and subject to all the objections thereto, we can say that these objections alike emphasise the problematic role which instantiation has to play in any theory of universals. That problematic role is precisely why we cannot sit content with the simple observation that *having a corresponding universal* is not preserved with respect to compounding by the logical operations such as *not* and *or* ([4]). To expand this point, someone might argue that while 'Fx' might have a corresponding universal, it does not in general follow that '∼Fx' does ([5]). Thus while 'x instantiates x' *might* have a corresponding universal, we are not entitled to conclude that '∼ x instantiates x' does ; and it is substitution for the latter in the Russell schema which gets the contradiction going. Against this, there are two points. First, it is not just a matter of negative universals causing the trouble, since Russell's Paradox can be replaced by what might be called Curry's Paradox for universals ([6]). Substitute for 'Fx' in the schema not '∼ x instantiates x', but rather 'x instantiates x → snow is black'. It is then an easy argument ([7]) to prove the falsehood that snow is black. Second, it is true that one need not maintain the *general* schema (∃u) (x) (x instantiates u ↔ Fx) to get the contradiction ; that it is enough to have (∃u) (x) (x instantiates u ↔ ∼ x instantiates x) or its cousins (e.g. Curry). But then, regress/hierarchy troubles to not turn on having a universal corresponding to the reflexive case 'x instantiates x', so much as applying some version of the Similarity Principle to the general situation in which *x* instantiates *u*.

(3) For a hierarchy, see Graham NERLICH 'Universals : Escaping Armstrong's Regresses' *Australasian Journal of Philosophy* 54 (1976), 58-64. On irreflexivism, see Armstrong, *Universals and Scientific Realism op. cit.*

(4) Aside from Armstrong's arguments for this observation, an advantage of adopting the principle that universalhood is not preserved with respect to truth functional compounding is that Popper's qualitative verisimilitude ordering is apparently resurrectable in a natural way by counting universals. A decent theory os verisimilitude has seemed essential to some realists (Smart, Putnam, Popper), and recent theories fall well short of this. See my 'Relevance and Verisimilitude' *Synthese* 55 (1983) 353-64.

(5) Armstrong *op. cit.* Notice that this is consistent with the thesis that universals are responsible for all similarities or predications via the operation of linguistic mechanisms.

(6) On Curry's Paradox for sets, see *e.g.* Robert Meyer, Richard Routley and Michael Dunn, 'Curry's Partadox', *Analysis* 40 (1980), 124-8.

(7) And one, moreover, which poses problems for paraconsistentist solutions to the problem ; see MEYER, Routley and Dunn, *op. cit.*

So the problem is that it cannot be that there is any sound argument for universals sufficient to show that similarities always have to be explained in terms of samenesses, since instantiation appears to be a counterexample. And yet one wants to save as much as possible the intuition that such explanations are *sometimes* appropriate, while stopping short of endorsing anything so strong that it produces a contradiction. While the Similarity Principle has to be abandoned, there seems to be *something* right about it, namely that samenesses sometimes *explain* similarities. But how to save the latter while drawing a not-too-*ad-hoc* line at the former? Various people have complained about Armstrong's *ad-hocery* concerning instantiating (8). But is the cost of avoiding it an abandonment of universals altogether?

3. Difference

A standard argument against resemblance nominalism is that our best theories about the universe assert *more than one similarity* between things, that is *different* similarities; and it is difficult to construe this assertion without quantification over similarities. That is to say, difference plays a central role in arguments for universals. Causal realism needs causal *variation*, or causal explanation of differences and potential differences between situations, no less than it needs causal samenesses. And so we might have a companion to the Similarity Principle which we can call the

(8) M. Devitt 'Ostrich Nominalism or Mirage Realism?' *Pacific Philosophical Quarterly* 61 (1980), 433-39; W. V. O. Quine, 'Soft Impeachment Disowned', *Pacific Philosophical Quarterly* 61 (1980), 448-9; Gail Fine, 'Armstrong on Relational and Nonrelational Realism', *Pacific Philosophical Quarterly*, 62 (1981), 262-71; D. Lewis, 'New Work for a Theory of Universals', *Australasian Journal of Philosophy* 61 (1983), 343-377. David Armstrong suggests in correspondence, appealing to Lewis' paper, that 'what is a difficulty for all positions is not a difficulty for any'; that *any* answer to the problem of universals requires *some* fundamental connective to be exempt from abstraction. It is true that unlimited abstraction has to be abandoned, but that is the beginning of the story. The point is that the Similarity Principle threatens us with unlimited abstraction, and yet it is hard to see how to avoid using it, or to place limits on its use. My suggestion is that it *is* possible to avoid placing too much weight on it. Again, as Armstrong notes, if instantiation is exempt from general abstraction in order to avoid regresses. Armstrong is hardly in a position to wield his own regress arguments against his opponents. So, then, one would seem to need different arguments for Armstrong's position. Perhaps it is not possible to get to Aristotelianism, but I do claim that one can get to universals without relying on Armstrong's regresses.

Dissimilarity Principle: no dissimilarity without a difference in what exists.

I suggest that the Dissimilarity Principle can get us where we want to go using the Similarity Principle, but at less cost. A common argument for universals, also a related argument against resemblance nominalism, is essentially an application of it. We can imagine a single particular to have more than one causally efficaceous property, e.g. charge and mass. A test particle with both charge and mass might be involved in a gravitational interaction in which only the latter is causally operative. But unless some things exist which are different to constitute the difference in these properties, this is unintelligible. Differences in behaviour and potential behaviour are attributed to independent variations in applicable laws. If this did not issue from differences in what exists, then nothing would be explained: the nonexistent cannot make for a difference in the existent. All dissimilarity is nonidentity.

Showing that more than one predicate can be true of a thing because the thing has different existing aspects, does not show what is the nature of aspects. So far as the argument has gone, it can consistently be said that everything has, for example, *parts*, and that some of the parts of a thing are its aspects ([9]); or perhaps that aspects are *abstract particulars*. Here is where we can put some weight on the Similarity Principle, though hopefully not too much. Because evidently those existing aspects of a thing which differ from one another, are precisely those things which are the same in different things: the mass 5 grams is different from the mass 6 grams, but more than one thing can have each. Thus, the Dissimilarity Principle delivers the existence of *universals* which differ from one another, the differences serving to constitute the lawlike differences in the world. But since they are universals, we have saved, at least to some degree, the intuition that there exist genuine samenesses between different things in the universe. But also, we have had to rely neither on any principle to the effect that all similarities ultimately derive from samenesses, nor any demonstrably unsound argument for it.

4. Instantiating

To reinforce this last point, we can ask whether there is some difference between things marked by the relational concept *instantiating*; and which

(9) A position which, for instance, a bundle theorist of universals might accept.

thus, on the Difference Principle, needs accounting for by an existing universal. It is hard to imagine that there could be. Certainly there is a difference between *a's instantiating u* and *a's* instantiating *v*, when *u* and *v* are different universals responsible for differential causal behaviour. But that difference is carried by $u \neq v$. The question is, rather, whether there is something common to *all* instantiations in virtue of which *a's* instantiating *u* differs from its not being the case that *a* instantiates *u*. But what we are trying to distinguish here is instantiating *something*, from not instantiating anything at all; since as we have just seen distinguishing instantiating *u* from instantiating *v* can be carried by $u \neq v$. The answer has to be that there is no extra difference between things to be accounted for here. Being a *mere* instance does not confer differential causal potentialities; what does that are the different universals instantiated. In short, it seems reasonable from the Dissimilarity Principle to draw the causal realist conclusion that those and only those universals exist which suffice to explain the differences in lawlike behaviour and causal potentialities in things, but that *mere instantiation* is not one such. The advantages of this conclusion for the problems of the previous section are evident.

A quick disclaimer must be made here. It is not being contended that causal realism rules out higher order universals in Armstrong's sense. At least those higher order universals will exist which are necessary to constitute the natures of universals involved in causal theory, and furthermore it is not unreasonable to say that these contribute indirectly to causal processes. We already admit different roles in causal processes, for instance particulars versus universals. That higher order universals might not appear in physical theory should not concern us overmuch if we think that metaphysics is also a legitimate source of truths.

It is by no means plain sailing from here, however. If you lean heavily on the Dissimilarity Principle, you need to be ever vigilant that some difference might not creep up for which you cannot account by differences in what exists. For example, what of the different predicates 'x exists' and '($\exists$u) (x instantiates u)'. Is there a difference here that will cause trouble ? The obvious and gratifying manœuvre here is simply to say that there is no real difference between existence and being an instance of a universal. This is a conclusion with which Armstrong and Kant would variously sympathise.

5. Against Sets

I want to go on to consider a complication. But before doing so, it is necessary to say some things about the use of set-theoretic constructions in philosophy, so that we will not be inclined to take the wrong route around the complication.

It is well known that the theory of universals has many parallels with the theory of sets, for instance Russell's Paradox above. When problems about universals are cast in the formal mode as problems about predication, it has often seemed tempting to reach for the Completeness Theorem for first-order classical logic as offering a set-theoretic account of predication, and thus an extensionalist reduction of universals to sets in line with the desire to account for similarities by samenesses. That set theory is an extensional theory of universals is both its strength and its weakness. It is a strength because the identity conditions for sets give no more problems than those of their members, and a weakness because those same identity conditions would then identify universals which are intuitively *and causally* distinct, so long as they were co-extensional. It is also common to object that the theory of universals-as-sets gets the order of explanation wrong. Far from its being the case that things are red because they are members of the set of red things, rather it is that S rather than S^1 is the set of red things *because* all and only S's members are red. Since a set is extensional, mere brackets so to speak, its identity is derivative from the identity of its members ; so that it takes on whatever character it has in virtue of the character of its members. Thus, underlying sets there must be universals to explain predication, samenesses and differences.

This point is worth amplifying. It has been argued ([10]) that the standard mathematical description of the topology of space, as a collection of sets of spacetime points, cannot be right because set theory gives us many more actual alternative collections of sets of points as well. Set theory *by itself* has no way of choosing which of the existing alternatives is the correct description of reality. Mathematical topology is an extensional attempt to describe the structure of space or spacetime entirely in terms of points or sets of points. And as *a description of the extensional features of that structure* no-one could quarrel with it. But, too many other set theoretic collections which equally well count as a topology also exist (since sets

(10) Chris Mortensen and Graham Nerlich, 'Physical Topology', *Journal of Philosophical Logic*, 1978, 209-223 ; also 'Spacetime and Handedness', *Ratio* (1983) 1-14.

do ; remember that we are evaluating universals as sets here). The trouble is that these other collections give an incorrect description of reality. So whatever constitutes the structure of space must be something which *underlies* sets, and which makes it possible to select one set theoretic story among others as the correct extensional story. In the spirit of the present paper, that something must be relations between points, polyadic universals. Thus, universals are necessary to explain why certain sets and not others do the job they are expected to do, because there are too many other collections of sets around which could do the job but happen not to. Universals explain why only certain collections of sets are the right ones to pick, and *that explanation needs to be made.*

This argument sits comfortably with causal realism. It is certainly a mistake to accept incautiously every entity (particularly mathematical entities) postulated by physical theory. It is not typically the concern of physicists to fine-tune their mathematical apparatus to suit the ontic scruples of philosophers. Causal realism is not physics worship ([11]). Indeed, it has always been irksome that the strictly mathematical items necessary for physical theory do not play a causal role in those theories ; epistemology has always been a stumbling block for the philosophy of mathematics. That is no news either, but it is worth seeing in the causal realist framework. It is hardly satisfying to argue, as many have, that the epistemology of mathematical items is no more problematic than successful postulation, wherein you get numbers and sets for free along with the electrons. The causal realist wants to postulate only those items which make a *causal difference.* This is where sets and universals differ, to the advantage of the latter. The possession of a universal can be causally efficaceous in a way that being a member of a set is not (or at best only derivatively). Instantiating a universal alters the causal efficacy of a thing ; that is precisely the role given to universals in physical theory. Needless to say, this rejection of items which make no causal difference puts causal realism at an opposite pole from the extreme of Quine's Pythagorean universe. As an aside, it would seem to be that causal realism views the geometrical structure of spacetime, since it makes a dynamic difference, as no less a causally relevant feature of the universe than any other feature of it, contrary to a thesis of Nerlich's ([12]).

(11) On physics worship, see 'Spacetime and Handedness', *op. cit.*
(12) Graham Nerlich, 'What Can Geometry Explain ?', *British Journal for the Philosophy of Science*, 30 (1979), 60-74.

It is not difficult to see that these arguments apply equally to the set-theoretic constructions on possible worlds which have been proposed to solve several philosophical problems. I do not mean to argue against modal realism here, so much as to reject the attempt to inject an element of intensionality into various universals, by identifying them with sets of extensions in possible worlds. If a world is a collection of particulars, distinguished by their properties and relations, and if properties and relations are nothing but sets whose membership varies from world to world, then there will ultimately be nothing which distinguishes a set in a world as the set of red things in that world rather than the set of green things. All the sets you'd ever want would exist, but only some would be correctly identified as sets of red things, and set theory itself doesn't tell how to do the identification.

Twentieth century mathematics could some justice be called the Age of Set Theory, to the extent that set theory has become its *lingus franca.* But set theory might also be viewed as a disease of modern mathematics, which has produced a secondary infection in philosophers, the etiology of which is mathematical logic and its Completeness Theorem. There are signs that mathematics is curing itself of the tendency to reduce *structure* to *containment*, with the development of Category Theory ([13]). It is to be hoped that philosophy will be more inclined to use set theory as a mere tool without being mesmerised by its ontic commitments.

6. Numbers and Quantities

We return to a matter raised at the end of Section One, namely the fact that our best theories, indeed all our theories, quantify over universals ; and that furthermore the prospects for nominalist reconstrual of this look very dim, since our equations exploit mathematical operations which make sense only on terms.

Now it might be argued, as Quine does ([14]), that predications at the basis of the equations of physical science, for example 'x's mass = 5 grams', can be reconstrued not so much to avoid commitment to universals, but to commit one only to certain kinds of universals, namely numbers. As follows : 'x's mass-in-grams = 5'. A further move might be

(13) See Robert Goldblatt, *Topoi,* North Holland, 1979.
(14) *E.g.* Quine, *Word and Object,* M.I.T. Press, 1960.

made to reduce the right hand side to sets, but let us not concern ourselves with that blind alley.

This position has its attractions, among which is that mathematical operations on numbers can be drawn on naturally to explicate their use in the equations of physical theory. Indeed, how else to do it ? What sense are we to make of writing 'Force = (5 grams × 6 grams) + (8 cm)^2'. Wouldn't it be more intelligible to write 'Force-in-appropriate-units = (5 × 6) + 8^2'.

Now it seems to me nevertheless that the Quinean view is not the *natural* view of the matter. The natural view is that x's mass is a universal in x, and the same universal might be in y ; that is, that the construal most natural in accordance with physical theory is 'x's mass = 5 grams = y's mass'. Worse is the fact that the alternative is at odds with causal realism, as we emphasised in the previous section, in that it requires the existence of numbers which make no causal contribution to the universe. One might begin to make sense of why x accelerates the way it does if told that x's mass = 5 grams, because the mass 5 grams is an entity the instantiation of which confers differential causal activity on x in accordance with physical law. But what contribution could the *number* 5 make to x's behaviour, different from the contribution the number 6 makes ? We seem to be in the same kind of problem as we saw for sets : both 5 and 6 would exist, so how could some *function*, mass-in-grams, serve to relate x differently to 5 from x's relation to 6, and in such a way as to confer differential causal capacity on x ?

So admitting *quantities* into one's ontology makes more causal sense than admitting *dimensionless numbers.* And to this conclusion one can add the weight of two arguments against the dimensionless numbers story. The first is the simple fact that it falls foul of the substitutivity of identity. For suppose that x has a mass of 14 grams and a charge of 14 volts. Then, on the theory, we are to construe this as x's mass-in-grams = 14 and x's mass-in-grams = 14 and x's charge-in-volts = 14. From which it follows that x's mass-in-grams is identical with x's charge-in-volts. But now suppose that x is involved in a gravitational interaction with an uncharged particle. Surely physical theory says that x's charge was causally *irrelevant* to x's subsequent behaviour, while certainly x's mass was not. How could that be if x's mass-in-grams was *identical with* x's charge-in-volts ? On the other hand, if x's mass = 14 grams and x's charge = 14 volts, we have no such problems. Of course, what we can conclude, if x's mass = 14 grams and y's mass = 14 grams, is that x's mass = y's mass, which, far from being problematic, is precisely in accordance with the view of them as universals.

A second argument against dimensionless numbers draws attention to the semantic structure of the expression 'x's mass-in-grams'. Having this expression denote a number, conceals the fact that it has a structure of implicational relations which would be difficult to explain. If x's mass-in-grams = 2000 then x's mass-in-kilograms = 2, but evidently mass-in -grams is a different function from mass-in-kilograms. The relationship between them obviously has something to do with x's having the *same mass,* that is that mass is a semantic component of mass-in-grams, but this is hard to make sense of on the present construal.

But do we have a problem here ? After all, how does one account for 'x's mass = 2000 grams ↔ x's mass = 2 kg' ? Evidently by saying that the universal 2000 grams is identical with the universal 2 kilograms. How this comes about, I suggest, is that change in a system of units for the one kind of property (grams to kilograms) amount to a systematic change in *names for the same properties.* '2000 grams' is part of a systematic set of names for properties, '2 kilograms' is part of a systematic set of a different set, and names the same property. This fits nicely with the intuition that *change of units is mere terminological change.*

A number of questions remain in connection with the present account, which I will mention but not pursue in this paper. One is to account for the fact that even in the preferred identity 'x's mass = 2000 grams', the number 2000 is a semantic component. So a job remains to be done of saying what contribution it makes to the theoretical function of '2000 grams'. In particular, this semantic contribution is bound up with the ability to perform arithmetical operations on dimensionless numbers which *thereby* has consequences for the values of dimensional quantities of which they are components. Clearly the answer will have something to do with how physical theory makes it felicitous to choose continuous ranges of properties falling under a property-kind. Beyond remarking that on the face of it physical theory does just exactly that, I will not pursue the matter. A related puzzle is why physical theory permits the multiplication and division of properties from different property kinds, but not their addition and subtraction: 6 grams × 6 seconds = 36 gm sec., 6 grams ÷ 6 seconds = 1 gram per second, but 6 grams ± 6 seconds doesn't make sense. Clearly whatever the answer is, control over arithmetical operations here rests with the dimension rather than the arithmetical operations on the numbers, a fact which I suspect the Quinean will have trouble dealing with. So causal realism suggests some thesis to the effect of 'abandon dimensionless mathematical entities in favour of quantities', since it is the latter which have causal bearing on the universe, according to causal theory. But perhaps this goes too far, since physical theory might provide for quantification over dimensionless numbers. The kind of case which strikes me might be troublesome is 'The number of electrons in the universe is finite'. Beyond noting the possibility that numbers might need to be admitted *as well as* quantities I will not pursue this. But I trust that this section has done enough damage to a certain metaphysical manœuvre which would otherwise cast doubt on my claim that we need to see basic causal theory in perspective, as a theory which deals with universals as *unavoidable and unproblematic.*

7. The Adicity of Instantiating

Now we consider a complication, which comes from the familiar problem of the irreducible adicity of predication. It was argued that one can account for differences in a thing by differences in entities instantiated in the thing. Let us now ask which adicity instantiation has. I spoke casually as if instantiation was a binary relation, but that will do only on the too-simple assumption that the aspects of a thing are the unary properties of a thing. More specifically, the difference between *a's* being heavier than *b* and *a's* being larger than *b* comes down to the instantiation of universals such as *heavier* and *larger*, which do not seem to be analysable into unary properties. If not, then we have to say that *heavier* is instantiated in *a* and *b*, where the order of *a* and *b* matters ; so that it seems pretty inescapable that instantiation is functioning here as a ternary relating. Conceivably it has something to do with *heaver* being instantiated in the *ordered* pair <a,b>, but that has the disadvantage of introducing set theory into a metaphysic so far free of it. What then shall we say of the earlier binary predicate 'x instantiates *u*' ? Is it really ternary after all ? That seems otiose, especially when as should be apparent there is no limit to the edicities we must allow. If it is not really ternary but binary, then we have two instantiation predicates of different adicity, *e.g.* 'x instantiates$_1$ *u*' and 'x and y instantiate$_2$ *v*. A similar argument evidently applies for any adicity for which there is a universal not analysable into combinations of lower adicity. Hence instantiating$_1$, ..., instantiating$_n$, ..., perhaps even up into the transfinite if there are universals of infinite adicity.

Now that is not so bad : it is a hierarchy and not a regress. Does the Dissimilarity Principle do any damage here ? I am inclined to think not, though it does seem arguable, since there would be an acknowledged dissimilarity between instantiating$_1$ and instantiating$_2$ marked by the indices. Nevertheless, even if there is no problem here, there is a certain lack of economy in the infinite hierarchy of instantiation predicates. I suggest that there is a natural way to avoid the hierarchy.

That can be done only if we have a single instantiation predicate, and this can be done only if either we can assign a fixed adicity to instantiation, or instantiation need not have a fixed adicity. Prospects for the former of these two possibilities are bleak, I should think. It means finding an upper limit to the adicity of instantiations and then analysing instantiations of higher and lower adicity in terms of it, all of which looks *a priori* unlikely. However, prospects for the latter are not at all bleak.

The theory of anadic predicates, predicates of no fixed adicity has been extensively studied, *e.g.* by Richard Grandy and by Barry Taylor ([15]). One example is 'x, ... x_n are surrounded by y_1 ... y_m'. There is no fixed number of places, either before the verb or after it. Yet the predicate does seem to constitute a single semantic unit, so that breaking it down into the infinite number of predicates 'x_1, ..., $_n$ are surrounded $_{n,m}$ by y_1, ...y_m' would be a distortion.

Utilising the insight of the theory of anadic predicates, we can say that the instantiation predicate is anadic. Indeed, if anadic predicates are possible at all, and if any such is analysed into the instantiation of a corresponding universal, then we will be forced to say that instantiating is anadic. If we do say this, then the infinite hierarchy above collapses immediately to the single case 'x_1, ..., x_n, ... instantiate *u*'. This view has the further advantage that it is consistent with the attractive thesis that there need only be *one* basic kind of multiple attribution of relational predicates which is not ultimately explicable in terms of the existence of universals, namely the assertion of instantiation.

(15) Richard GRANDY, *Advanced Logic for Applications*, Synthese Library, Reidel 1977. Barry Taylor, 'Articulated Predication and Truth Theory', in B. VERMAZEN and M. HINTIKKA (eds.), *Profiles of Philosophers : Donald Davidson*, ... Note though that the methodology of these is thoroughly set-theoretic.

8. Conclusion

One further matter relates to the dispute between Aristotelian and Platonic theories of universals. One difference is that the account of laws with uninstantiated subject terms looks to be considerably simpler in the latter than in the former ([16]). On the other hand, Aristotelianism does seem to have a better account of being a particular. The Dissimularity Principle perhaps creates trouble with the difference between '*x* is a particular' and '(∃u) (x instantiates *u*')', but an Aristotelian can, instead of appealing to a brute difference between particularity and universality, presumably analyse the former as '~(∃y) (y instantiates x)' i.e. particulars are things which have no instances. That is not available to anyone who thinks that there exist uninstantiated universals.

Department of Philosophy.
The University of Adelaide.

(16) Tooley, *op. cit.*, or Armstrong, *What Is a Law of Nature ? op. cit.*

Semiotics and the foundations of mathematics

CHRIS MORTENSEN and LESLEY ROBERTS

Introduction

Semiotics has largely been neglected as a vehicle for foundational studies in mathematics. A notable exception is Brian Rotman, who has, in the name of semiotics, offered accounts of mathematics in general and number in particular. Rotman nominates his basic semiotic vocabulary as being principally Saussure's, though informed by Peirce and later developments such as Eco: sign, signifier, signified, language or code, discourse, metasign, and subject (Rotman 1993: 31). We therefore begin this article with a brief overview, particularly of Saussure's version of semiotics, taking the opportunity to do some critical reconstruction along the way. We then briefly survey the problem of the nature of mathematics as conceived by analytical philosophy, drawing on Hilary Putnam's work. These preliminaries set the scene with the theoretical concepts we use. Our main interest in this article is in numbers, including especially the number zero. We proceed to describe and criticize Rotman's theories on these matters, and then to offer an alternative understanding that draws on some of his insights. We then consider the problem of infinity, arguing that Rotman's position on this issue also has its drawbacks. We conclude by broadening the focus to the nature of mathematics in general, discussing the work of Edwin Coleman and René Thom in the context of the prospects for a semiotically-informed philosophy of mathematics.

Preliminaries: Saussure's semiotics

Saussure's (1916) linguistic signs were made up of a signifier and a signified. These do not occur separately: Saussure used the analogy of a piece of paper — the sign is like the sheet of paper with the signifier as one side and the signified as the other. Saussure said that the signifier is 'the form which expresses the word' and the signified is the meaning.

Semiotica 115–1/2 (1997), 1–25

0037–1998/97/0115–0001

Now Saussure's account is not primarily about external, physical items. Rather, it is a mentalist account; while Saussure allowed a material substrate, both the signifier and the signified are located in the mind. The mentalist account of the signifier (as opposed to the signified) is one which we think it would be better to avoid, if only because it privatizes a centrally important linguistic concept. We follow Wittgenstein in holding that language is importantly public, and that only if some central linguistic building-blocks are public can one explain the singular usefulness of language in communication with complete strangers. Nonetheless, this issue would seem to be largely one of nomenclature, we suggest, if it is conceded that signifiers have a physical base.

Saussure pointed out that the signifier and signified are arbitrarily linked, or as one might say, the link between syntax and meaning is conventional. What is less obvious is Saussure's further claim, that both the signifiers and the signifieds are arbitrary. He maintained that there is no intrinsic property which determines a particular signifier or a particular signified, but rather a signifier is defined by its relations to all other signifiers, and a signified by its relations to all other signifieds. Language on this view turns out to be a kind of relational algebra; what is important are structural relations.

It is the system of relations that establishes identity conditions for signifiers and signifieds. Saussure explained this by using a chess analogy. If you are playing a game of chess and you misplace a knight in the middle of the game, then you can replace it with something which bears no physical resemblance to the original piece, such as a button, provided that the same relations hold between it and the other pieces as held between the knight and those pieces. If the structure does not change, if two linguistic objects stand in the same relations, then they count as being the same. Thus the account is thoroughly holistic.

Saussure used the term 'difference' to talk about relations between different items. We believe that this has led to some confusion in later interpreters. One of his famous quotes is 'in language there are only differences' (1974 [1916]: 120). We claim that this has to be understood in the way we have indicated, as the claim that the identity conditions of linguistic units are relational and holistic. If Saussure meant by 'difference' literally simple disidentity, then he would be open to a conclusive technical objection. This is, that there is a simple proof, using quantifier logic plus identity as a model language, which shows that no collection of nonidentity statements is sufficient to imply any identity statement which is not already part of one's theory.[1] This shows that in order to fix the identities of any collection of items, more is needed than their differences construed as disidentities.

Following Rotman, we will say that a code is a language or sublanguage exhibiting this typical internally differentiating pattern of differences. What is it, then, that structures a code, if it is not simple disidentity? Saussure wanted to define two kinds of relations called syntagmatic and associative. He gave some clues as to what these relations are, but in neither case were they clearly defined. Syntagmatic relations were based on the 'linear nature of language'. He said that 'a term acquires its value ... because it stands in opposition to everything that precedes or follows it, or both' (1974 [1916]: 123). Presumably such relations would give us the grammar of the language, i.e., which signs can be combined with which. The definition of an associative relation is not much clearer. A word, he said, will 'call to mind' a number of other words which are 'related in some way' (1974 [1916]: 123). The different kinds of relationships are formed by different associative relations. But just what are these different kinds of relationships? Saussure was not clear. He said that associative relations are 'based on the comparing of terms which have something in common' (1974 [1916]: 125), and he included antonyms, synonyms, rhymes, and having the same prefix, suffix, or root. Thus, for example, 'black' might be associated with 'white' and also with 'tack'. Now while Saussure may not have been very clear about association, we can certainly assimilate any other work on association, natural or conventional, from Pavlov onward. In more modern terms it is the thesis that meaning is a pattern of distances in cognitive space.

Saussure's concept of the sign is different from what we ordinarily mean by 'sign'. Something is a sign if it stands in for something else: smoke is a sign of fire, a footprint in the sand is a sign of human presence. General semiotic theories which adopt this more general conception of a sign (such as C. S. Peirce's, for example) are general theories of representation; and their associated semiotic accounts of language typically come with a theory of reference. But Saussure does not explain this stand-for relation. The relation between signifier and signified is not a stand-for relation; one does not represent the other.

It is almost a cliche to point out that Saussurean semiotics suffers from Saussure's omission of a theory of reference. This strikes us as somewhat unfair to semiotics. It is certainly true that without a theory of reference one runs the risk of thinking that there is nothing outside the text. Contrary to the apparent view of some post-structuralist thinkers that there is no extra-linguistic world, it is unavoidable that a general theory of language must take into consideration reference, that is, the naming relation between signs and the physical and social world, on pain of denying us linguistic and epistemic access to a mind-independent reality. But this is hardly a major problem: for example, Peirce's version of semiotics admits it. Nevertheless, there can surely be no objection to pulling apart the sign in order to study the signifier-signified relation, which we gloss here as the relation between syntax and the varieties of meaning. Indeed, there is likely to be a special imperative to do so for the case of mathematical codes: as we see, it is their referentiality which is open to question here, so that we are interested in the possible contributors to signification in the absence of referentiality. One such contribution is likely to be from surrounding discourses, or what Wittgenstein called language-games; though we should always be aware that language-games or discourses typically have a socio-political-historical aspect, as well as a structural-relational aspect, in their contribution to meaning. Further, while there are strong prima-facie objections to Saussure's mentalist account of the signified (meanings), nonetheless we concur with Dummett (e.g., 1977) that the emphasis on the concrete epistemic phenomenon of *understanding* locates meaning correctly with respect to its cognitive foundations: any theory has to be compatible with its own epistemology.

Further preliminaries: Analytical philosophy of mathematics

A central problem for any general account of the nature of mathematics is 'what is mathematics about?' One answer is given by Platonism. On a Platonist account, mathematics is about real, existing abstract objects such as numbers and sets, which are referred to by mathematical words or signifiers such as the count-nouns. In a well-known overview, 'Philosophy of mathematics: A report' (1979), Hilary Putnam discusses major accounts of mathematics including Platonism, conventionalism, formalism, and intuitionism. Putnam notes two definitive issues for any position. One is the truthmakers of mathematical propositions: given that we think that mathematics is straightforwardly true or false, and mathematicians give every sign in their practice that they think just this, what makes the true propositions true? Here Platonism has a ready answer, namely, that the truthmakers are the abstract objects and their properties and relations to other such abstract objects, to which mathematical nouns refer. The ability of Platonism to give a ready account of the truthmakers of mathematics is unquestionably its major strength.

The second central issue for any philosophy of mathematics is the epistemology of mathematics: how we can know any mathematical truths? On this point Putnam is, contrary to some well-known theorists including Godel, uncompromisingly naturalistic: our knowing mechanisms are finite nerve nets, so that Platonism in particular has the serious difficulty of how finite nerve nets can have any connection with abstract objects sufficient to know them (Putnam 1979: 389). There are two related epistemological problems here. First, as we have just seen, any account must explain or explain away how we come to know the 'objects' of mathematics and truths about them. Second, it must explain how we as finite beings come to understand infinite mathematical structures. Even the set of natural numbers {0,1,2,3 ... } seems too large for mere mortals to comprehend.

We set aside the special problem of infinity until a later section, although we notice that debates about the nature of infinite processes in mathematics go right back to the ancient Greeks. Two opposing positions have been (1) the claim that the appeal to infinite processes and series is only ever an appeal to the *potential infinite*, that all an infinite process ever amounts to is an instruction to continue; as opposed to (2) the claim that *actually infinite* mathematical items such as infinite series and sets exist. The debate became especially acute from the seventeenth century when calculus was discovered, and the fate of infinite items became linked to the fate of their reciprocals, the infinitesimal numbers. The debate was widely held to be settled in favor of the potential infinite by the nineteenth-century Cauchy-Weierstrass definition of limits and the derivative; though later that century Cantor's paradise of infinite sets, and particularly Robinson's important discovery in the 1960s of nonstandard analysis, sharply revived the question.

On the question of general epistemology, Putnam mentions the approach of Quine, that numbers and sets including infinite sets are postulated in the very same way that electrons are, and for the very same reasons, namely, as unavoidable parts of the explanation of the success of the observations of physicists (Putnam 1979: 390). It is worth making a terminological distinction at this point between Platonism and realism: mathematical realism is the general position which asserts the existence of mathematical entities; while Platonism is the extra claim that mathematical entities are abstract, by which we will mean that abstract entities are not in spacetime nor are they causally relevant to events occurring therein. The Quinean argument is certainly realist and referentialist about mathematics, though not obviously Platonist (Putnam calls it holist; see 1979: 390). At first glance, moreover, it proposes a better epistemology for any realist about mathematical entities, namely the familiar hypothetico-deductive method of natural science. However, it also raises the important distinction between pure mathematics and applied mathematics.

Realism can appeal to supporting arguments wherever it can find them. In Quine's case we have *the argument from applied mathematics and* *physics*. While we do not agree with Quine's argument, we will support a different argument from natural science for a realist conclusion. Whether it is a Platonist conclusion is a matter which we discuss. But notice that it is left open that if physics were different, then perhaps the mathematics necessary to describe it would not need to assert the existence of numbers and sets. We would then have no Quinean reason to think that they existed; we might as well invoke Ockham's Razor and deny them. Why one initially resists this conclusion is because it leaves untouched the whole question of pure mathematics. After all, to the extent that entities are needed in physics to explain the causal order, to that extent they are causally implicated and thus not abstract as we are using the term. Such realism is not at all inevitably Platonism. On the other hand, it seems likely that there are areas of pure mathematics which will, as a matter of the construction of the universe, never be necessary for physics (consider the example 'There exists at least one nondenumerable inaccessible number'). At any rate, much pure mathematics is done at some remove from practical applications; even if, as with differential geometry, it is sometimes later discovered to have a use in physics. Pure mathematics has thus an aspect of necessary truth which the contingent truths of physics do not explain, just as arguments have an aspect of absolute logical validity, indeed at its most noticeable in pure mathematics, which logic studies. This is *the argument from pure mathematics and logic* for full-scale Platonism, as supplying the truthmakers for the true propositions and valid arguments of pure mathematics. Putnam is aware of this point, and tries to broaden his data to be explained to include 'combinatorial facts' (1979: 390). But he concedes that such Platonism, if it asserts the existence of acausal abstract objects, is back with the epistemological problem, since the hypothetico-deductive method does not seem to engage with entities whose presence would make no difference to the contingent causal order.

We do not take this up here, since our target is Rotman; but we wish to register disagreement with one more aspect of Putnam's discussion. One major rival to Platonism has been Brouwer's intuitionism (see, e.g., Dummett 1977), according to which mathematics is a mental construction which does not exist until constructed and isn't to be regarded as true until constructed: mathematics is made, not discovered as the realists would have it. Brouwer is to be saluted for placing the epistemology of mathematics at the center of a philosophical understanding of mathematics. But Putnam points out that the mentalism is at odds with the a priori character of pure mathematics and logic (1979: 394): how could a mind (finite nerve net), whose operations are causally determined by the chemistry of the brain, create the necessary constraints on mathematical constructions? For example, it is well known that there is nothing to stop a brain or mind from being inconsistent; whereas, according to Putnam, we regard inconsistency in a mathematical construction as a sure sign of error.

While we agree that pure mathematics and logic have an aspect of necessity which needs to be explained or explained away, we think that Putnam has gone too far here. As Imre Lakatos (1976) and others have made abundantly clear, the history of mathematics, far from a paradigm of formal clarity and rigor, is the story of a boat afloat on a sea of anomalies. In a particular case, the discovery of the paradoxes of set theory and semantics in contexts where deduction seemed to be at its most innocent and pellucid, has been a driving force in mathematical logic this century. However, its noticeably ad hoc outcomes have recently led some theorists (e.g., da Costa, Priest, Routley) to react against the constraint of consistency. This has led ultimately to the discovery of inconsistent mathematics, on which see, e.g., Mortensen (1994). That is to say, Putnam is wrong in the detail of this criticism of intuitionism.

Another less-than-satisfactory aspect of intuitionism is its revisionism. It is notable that, in consequence of his analysis of what a mental construction could be, Brouwer concluded that many existing constructions in mathematics were illegitimate, and that certain apparently reasonable logical principles such as the Law of Excluded Middle (either A is true or not-A is true) were unsound. Now Quine's realism bites, because physics plainly needs *some* mathematics; hence Brouwer had to show that he can get enough in his truncated mathematics to deliver the successful predictions of physics. Whether intuitionism can do this satisfactorily is still open; though it is clear that Brouwer's attempt, brilliantly creative though it was, was also remarkably ugly and restricting. But there is a deeper objection here to any revisionist philosophy of mathematics. Indeed, it applies also to any philosophy of mathematics, such as Hilbert's formalism (see, e.g., Putnam 1979: 388), which accords a notion of 'canonical mathematics': that some mathematical formalisms are more canonical as mathematics than others. A revisionist is in the business of claiming that some accepted mathematics is correct, while some other accepted mathematics is false or misleading or less canonical, and thus should be discarded or taken less seriously or revised. However, such a position can only be part of the story. At most, what are addressed are the questions of what are the truth-makers and the epistemology for mathematics. These are, of course, perfectly legitimate questions which have been addressed in the foundations of mathematics this century, but left unaddressed is the question of *what makes the supposedly false or uncanonical parts of mathematics still mathematics*? Surely, until we can answer this question we will not have understood something at the core of mathematics. And this is why semiotics has an appealing prospect here, since it offers the possibility of understanding that true and false mathematics are both mathematics in that they are both a *distinctive* kind of *text*. We return to this point in the final section, particularly in the discussion of Edwin Coleman's work.

Putnam concludes that so far 'nothing has worked', that every existing account suffers from some fatal flaw. He finishes with what is from our point of view a significant observation: he urges philosophers of mathematics to cease doing formal mathematical logic, and to start investigating the history of mathematics, plausible reasoning, and the philosophy of language, in 'discussion of the deep metaphysical issue of realism as a theory of truth and reference' (1979: 395).

Rotman's account of mathematics and number

In his paper 'Towards a semiotics of mathematics' (1988), Brian Rotman begins by looking at what makes up a mathematical text, and notes that in such a text one does not just find mathematical notation, but also natural language. Rotman claims that the signifiers and signifieds of mathematical discourses are the 'scribbles' and 'thoughts' of mathematicians. That is, the signifiers are what the mathematician writes down, and the signifieds are whatever is going on in the mathematician's head when she is doing mathematics.

Rotman posits three 'semiotic subjects'. The three semiotic subjects are the Mathematician, the Agent, and the Person, who work together when mathematics is practiced. The Mathematician is the subject who does the scribbling, and who follows the inclusive imperatives of mathematical discourse, such as 'consider', 'define', and 'prove'. The Mathematician imagines the domains within which mathematical actions take place and she also imagines the second semiotic subject, the Agent. The Agent is a kind of skeleton diagram of the Mathematician, who carries out mathematical instructions within the imagined domains. The Agent, according to Rotman, is 'free from the constraints of finitude and logical feasibility' (1988: 13). Thus, when proving the statement 'for all numbers x and y, $x+y=y+x$', while the Mathematician cannot manipulate all possible instances of x and y to determine that $x+y=y+x$ in all cases, the Agent has no trouble with this infinite task. Rotman's Person is the subject who answers to the pronoun 'I', and who can articulate the connection between the Mathematician and the Agent. It is the Person who, in virtue of this capacity, can be persuaded by a proof. In each line of a proof, Rotman's Mathematician observes the imagined Agent performing some task. The Mathematician 'becomes convinced — persuaded somehow by the thought experiment — that were he to perform these actions the result would be as predicted' (1988: 14). Even for finite equations such as $2+3=3+2$, Rotman says that the Mathematician must be convinced that the equation holds for all tokens of 2 and 3. Once convinced, the Mathematician 'scribbles' a new line of the proof. The Person, who grasps the underlying narrative, the connection between the Mathematician and the Agent, is persuaded by the proof.

Rotman gives an account of whole numbers with which he wishes to show the falsity of Platonism, and claims that his account shows that mathematical discourse creates its own objects. He argues that numbers appear when there is a subject who counts: counting is a recursive process, and an analysis of number should begin with a close look at I, II, III, IIII, etc. Counting begins when 'I' is taken as a signifier and 'etc.' as an instruction to copy the last signifier and add another 'I'. On Rotman's account, 'etc.' is an instruction to the Mathematician who will imagine her Agent performing the algorithm. Numbers, he says, are 'things in potentia' (1988: 32). They are all the possible signs which can be produced by the Mathematician and her Agent.

There are several problems with Rotman's account. To begin with, it is not at all obvious how the falsity of Platonism or even realism follows from this account. While the signifier might be inseparable from the signified, it does not follow that there is no object to which number signs refer, any more than it would follow in the general case of arbitrary signs where reference evidently takes place.

In his book, *Signifying Nothing: The Semiotics of Zero* (1987), Rotman offers a 'deconstruction' of number discourses in an attempt to show that numbers do not exist prior to the numerals or count-nouns, but are in fact produced by them. Rotman's deconstruction relies on the introduction of the (number) variable. On a Platonist account of mathematics, a variable can be replaced by any numeral which names a pre-existing number. Rotman argues that variables can be replaced by number signs and these must be produced by the process of counting, hence variables must be explained in terms of a counting subject. A variable thus ranges over all the possible signs produced by the 'one who counts' (1987: 3).

This account is reminiscent of those anti-Platonist accounts of number which turn on what is known as the 'substitutional' interpretation of variables, as substitutable by count-nouns which are not referential (see later discussion, especially Note 2). We do not propose to consider this line at this point: we are more interested here in Rotman's argument for his position. Now a Platonist could readily admit that variables can be replaced by number signs which are produced by counting, but still insist that those signs are referential. Rotman's deconstruction seems to rest on a prior claim, that numbers are produced by numerals. His argument appears to be like this: (1) number signs are produced by counting, (2) number signs do not refer to but are constitutive of numbers, and (3) variables can be substituted for by number signs. Hence, (4) numbers are produced by numerals. The trouble is that when (4) is unpacked, it is not clearly different from (2); that is, Rotman has assumed what he is trying to show.

The most serious problem faced by Rotman's general account is the epistemological problem. We saw that the Platonists have the problem of explaining our knowledge of the abstract entities they claim do exist. We argue that Rotman's account suffers from a similar difficulty.

At first glance, Rotman's semiotic subjects might seem to provide solutions to both epistemological problems. Mathematical objects turn out to be the signs which we produce ourselves, and hence there is no problem in knowing them. In a proof, the Mathematician observes the Agent performing an infinite number of tasks and becomes convinced that were she to perform them, her answer would be the same. The Mathematician/Agent creates the signs of mathematics, and the Person believes in truths about them. The Person 'knows' the relationship which holds between the Mathematician and the Agent, and hence 'knows' the mathematical objects which are produced by the Mathematician through her thought experiments and her scribbles. So the objects of mathematics are knowable because they are the creations of mathematicians. The problem of infinite structures is similarly solved. Though the Mathematician is finite, she has an infinite Agent who does the mathematical leg-work for her. Since the Person has an understanding of the relationship between the Mathematician and the Agent, she can understand infinite structures.

Still, this account has two obvious and related problems. The first is how a functional part of a finite Person, the Agent, could manage to perform an infinite task. The second is how another functional part, the Mathematician, who *ex hypothesi* cannot perform an infinite task, could be rationally persuaded by the operations of the Agent. At any point in time, when the Mathematician observes the Agent, only a finite number of these tasks will have been performed. As Rotman acknowledges, the Mathematician is a finite being, and finite beings simply do not have the time to imagine infinite processes. In other words, Rotman seeks to explain our understanding of infinite structures by positing an Agent who can do the work for us. But he does not provide a convincing account of how we are to understand the infinite processes of the Agent.

The special case of zero

The sign for zero plays, on the face of it, a special role. As a number, zero is stranger than other numbers, if only because zero things of a given kind are not any number of things of that kind at all. The same is true of other 'null signs' in mathematics, such as the null set, initial objects in category theory, or such things as null sequences or null lists in computer science.

Rotman concurs in giving zero a special place; but we will argue that the details of his position are unsatisfactory. Rotman's most consistent account would seem to be that zero is a meta-sign which signifies the absence of other signs (see, e.g., 1987: 3). By a meta-sign, Rotman means a sign which appears like other signs but which has a role as signifying aspects of those other signs, particularly aspects which involve a subject. As a count-noun, zero signifies the 'origin' of counting (1987: 13). (In his later book *Ad Infinitum* [1993], Rotman appears to have a somewhat different nomenclature: there the mathematical metacode is identified as informal mathematics; and is distinguished from code, which is mathematics considered strictly formally, though it might be argued that metacode in this sense continues to fill the earlier role of meta-signs. We consider the later book more closely in a later section.)

Now to say that zero plays an excluding role with respect to other signs is, from a Saussurean point of view, no news: we saw that for Saussure all signs are constituted by their difference from others within the code, but this precisely does not distinguish zero from other signs. Furthermore, it is initially strange to speak of zero as the origin of counting: counting does not start at zero, but at one. What Rotman seems to mean here by the origin, is the originator, perforce a subjectivity and thus a subject. Zero signifies the trace of the originator of counting in a way that one does not, presumably precisely because no one starts at zero when they count and because zero plainly is not referential. Still, if the *subjective act* of counting is the origin of numbers, then it is not obvious why zero should play a greater role in signifying that, than one or any other number should. All results of counting ought to signify the counting.

We are not confident we have understood Rotman on what semiotic closure or completion amounts to. It seems to be a kind of limit, and if that is so then perhaps it is simply another way of expressing the idea of zero as the origin of counting. Rotman uses the analogy with algebraic variables and equations. These too are meta-signs in that they are to be understood as signifying a subjectivity at work, the algebraic subject, who is performing a kind of virtual counting, counting at one remove as it were. But equations containing variables are also the semiotic completion of counting in that they signify the results of all possible acts of computation through counting. Note again the substitutional conception of variables as arising from more primitive semiotic acts of the same linguistic type, as opposed to the Platonist conception of variables as a kind of variable name for abstract mathematical objects.

Rotman finds a role for zero in other places in mathematical discourses. He discusses the vanishing point in perspective in paintings, or line at infinity in projective geometry (1987: 17,39). The line at infinity in projective geometry is the place where the difference between parallel lines disappears or ceases to exist. Rotman argues that the device of perspective in paintings is a meta-sign in that it signifies a subjective point of view initially different from the viewer's, namely the artist's. However, we submit that there is nothing particularly subjective or meta-linguistic in this conception. After all, we are all familiar with the public visual aspect of the horizon, which is furthermore describable by a simple and uniform collection of mathematical transformations. Indeed, there are more 'perspectives' than there are subjectivities. An analogy is in the Special Theory of Relativity, where there are more frames than observers: frames are perfectly objective aspects of reality in Special Relativity, and talk of observers was an accretion of outdated Positivism. That is not to deny that zero and infinity, as reciprocals of one another, are intimately related. On the other hand, there is an aspect of the denial of existence of a difference at the horizon which we will see is on all fours with what we say about the semiotic role of zero in the next section.

Another place to find zero at work, according to Rotman, is in exchange, money, and credit (e.g., 1987: 5,46). Zero arises once bookkeeping makes possible balanced books. This strikes us as quite right, and indeed a good place to look for one of the important semiotic functions for zero. We will take it up in more detail in the next section. Having said this, we wish to turn to give a positive account of numbers in general, which will locate the semiotic role of zero within it. We will see, however, that we will eventually have to tolerate a measure of number realism of a sort. But we will also see that zero has a special place, and that realism about zero and other null signs is to be avoided at all costs.

Numbers: The true story can now be told

In considering the semiotic nature of mathematical signs, it must be stressed that prior considerations about what the nonhuman world must be like in order to support mathematical practices are relevant. It will be seen that the origins of realism are to be found in a special semiotic context, namely, that of natural science, which in turn offers reasonable hope for epistemology.

So we begin with the count-nouns 'one', 'two', ... , and ask for their role in counting. The salient point which springs to mind is this: that a simple activity like counting, teachable to any preschool child, surely has nothing whatever to do with arcane abstract objects. This point alone is telling against a Platonist-referentialist view of the count nouns. Platonic numbers simply do not arise from counting.

This intuitively reasonable argument is reinforced by a technical consideration from logic. When counting the number of marbles in a tin, we get an answer like 'There are two marbles in the tin'. It is well known that this can be rendered in a standard and systematic way in the usual apparatus of quantificational logic, without mentioning entities such as numbers, like this: 'There is an x and there is a y such that x and y are marbles in the tin, x is not identical with y; and for any z which is a marble in the tin, either z is x or z is y'. The last clause, beginning 'and for any z ... ', obviously says that there are no more than two marbles in the tin. If the result of counting was that there are three marbles, one could render this in a systematic way as 'There is an x, a y and a w such that x,y, and w are marbles ... etc.'. Notice that there are no abstract numbers spoken of here, only marbles and the tin.[2]

A second salient point about counting concerns the unique role of zero in this process. To count the marbles in the tin, start with some marble and say 'one', move to another marble and say 'two', and continue appropriately. Here is how NOT to count: point to an empty space and say 'zero', then continue as before. That is, zero does not arise from counting in the way that the (other) count-nouns do: in an important sense, zero is not a count-noun since it is not the outcome of a counting. Could zero be dispensed with altogether in counting then? Now zero finds a use in a counting-related activity, as one possible answer to the question 'How many marbles in the tin?' However, 'There are zero marbles in the tin' does not refer to an abstract entity, zero; it is not even the same *kind* of answer as 'There is one marble', 'There are two marbles' ..., etc. The latter are all constructed from the basic beginning 'There is one marble' by adding additional existence claims, whereas the answer zero is a denial of existence. It is the answer one gives when all the existential claims are false. But neither is it any kind of *meta*-assertion. It is an answer which one sometimes gives because the series 'There is one marble', 'There are two marbles' ..., does not logically exhaust the possibilities by itself; while adding the possibility 'There does not exist a marble' does provide an exhaustive set. This in turn raises the general issue of the role in mathematics of null entities such as the null set. We will be suggesting that they are nominalized ways of denying existence, and thus no matter what Platonist conclusions one comes to about other entities in mathematics, zero ought not to be construed as naming any sort of existing thing, abstract or concrete.

It can be fiendishly difficult to rid oneself of the tendency to think that nonexistence in general is a kind of thing, a null thing. Natural languages typically contain a nominalization of nonexistence: in English one can speak of the absence of nonexistence of a thing, and Sartre wrote about *le néant*, nothingness. But we submit that this tendency ought to be resisted, in the name of common sense over false profundity. It is well known that a nominalized place does not guarantee referential status. In defense of our position, we offer a slogan: *nothingness does not exist.*

Thus the natural number nouns 0,1,2 ... arise in counting, but their use there is not referential; and in any case zero is a special case, though not a meta-case. But this now leads us to ask from whence come the negative numbers. When the negative numbers are taken into account, there arises the additive group of integers ... $-2,-1,0,1,2,$ ... . For our purposes, the important part of describing this collection as an additive group is that every number now has its negative counterpart with the property that adding the two together produces the zero. Negatives and zero, then, go together. But negatives do not arise from the semiotic activity of counting, or at least not without a *directionality* to the counting. With directional counting one then has the possibility of a result zero to the counting: count in one direction, then count the same amount in the other direction. Direction suggests geometry, but it is too soon for geometry. The obvious model is exchange, barter, the market, giving and receiving as group inverses of one another. Nor are there any realist numbers, negative or positive, needed for exchange (at least not if the exchange involves merely counting units of things as opposed to measuring continuous quantities of things, see later this section). We now see a further role for zero emerging, that of an absence of change, an inability to make a difference, the end result of two operations which 'cancel each other out'. Note again the distinctive pattern of zero being involved as a denial of existence. Not only does nothingness not exist, but *there are no null entities.*

With exchange comes money. Rotman notes that money comes in two stages, gold and post-gold (1987: 46). The former had intrinsic value. The latter comes about because the intrinsic value of the former can be debased below its face value. Post-gold money counts as a meta-sign for gold. Money plays an important intermediary role in facilitating exchange. But notice a further fact so obvious as to need explanation, that no one ever made a genuine coinage with a face value of zero dollars (pounds, francs, lira ...). (We have seen play money with such a denomination, but it is no sort of legal tender.) The role of zero in denying existence provides the explanation, for if a coin were worth zero dollars then its purchasing power would be nonexistent, and it would thus be wasting money to manufacture it. (There could still be existing things having no monetary value, things which it did not cost to manufacture or obtain, but they would not be coinage.)

Post-gold coinage leads inevitably to a different code, credit. According to Rotman, the role of a meta-sign such as post-gold money is to signal and facilitate a change of codes. Signs indeed have the power to change codes, which makes them in a sense prior to things and 'creative' of them (Rotman 1987: 49). We think that it is worth emphasizing that the Hegelian/Marxian tradition had the useful concept of a dialectical process, signifying a mutual interaction, sometimes with a struggle. Too much emphasis on one aspect of the contradiction can lead to according it a false priority, and even to thinking of the other aspect as a mere epiphenomenon.

With credit and the keeping of financial books, negative numbers and zero gain a further role as the inverse of assets: debts. Then the 'fiction', net worth, can be defined as assets minus debts. Now we have a puzzle; for while having zero debts means that there are no debts, and having zero assets means that there are no assets, having zero net worth does not mean that there is no worth. Zero net worth is a positive state, better than negative net worth, for example. It need not involve the nonexistence of purchasing power; and will not, while assets continue to exist. Zero net worth just means that amount of assets = amount of debts. There is, however, an implied *conditional* denial of existence: that if assets were used to discharge debts, then neither assets nor debts would exist at the end of the process.

So far we have seen that the correct way to understand the use of count-nouns does not require realism about numbers, that negative count-nouns arise in conditions such as exchange where directionality of counting has a place, that the count-noun zero has a special role associated with the denial of existence, and that these facts can be discerned from the formal properties governing the mathematical structures in question. But now we have to consider the role of continuously distributed numbers such as $\sqrt{2}$ and pi, that is, the real numbers. The position we develop is essentially Newton's, though it derives ultimately from Plato. We also draw on recent developments and improvements on the status of quantities and numbers done by, e.g., Forrest and Armstrong (1987), Bigelow (1988), Bigelow and Pargetter (1990), and Mortensen (1987, 1989, 1994).

We will see that the anti-realism we have been developing about numbers must be abandoned, and that the correct understanding of real number signifiers is referential.

With the rise of the theoretical sciences, especially physics, physical chemistry, and chemistry, quantities arise. Quantities come with a number and a quantity-kind: 3 sec, 5.2 gm, $\sqrt{2}$ cm. But the primary reality is the dimensioned entity, the quantity; while the dimensionless entity, the pure number, is derivative. Should one treat quantity-signifiers as referential? The natural answer is yes, in view of the role they play in natural laws. For example, consider Newton's law of gravitation: any pair of bodies with any masses m_1, m_2, separated by a distance r, attract each other with a force F given by $F=Gm_1m_2/r^2$, where G is the gravitational constant. The quantities F, G, m_1, m_2, and r all come in dimensions or quantity-kinds. The laws of nature relate quantities together, and in virtue of that explain the observed behavior of bodies. So lawlike relationships between quantities constitute fundamental aspects of the universe. This is the best of reasons for taking fundamental laws literally and hence referentially and realistically.[3]

What sort of entities, then, are quantities? Evidently, they are quite like the universals of Plato and Aristotle.[4] One and the same quantity, such as 5 gm, can be possessed by a multiplicity of things, in different locations. This was the principal mark of a universal. Because universals are unlike spatially located bodies, they have seemed controversial entities to many through the ages, but this suspicion seems misplaced to us. Universals are no more mysterious than the quantities familiar to us from elementary physics, and shouldn't be feared.

Our mathematical interest, however, is in the numbers which come as a part of quantities, the '5' in '5 gm'. Here we draw on Newton (1728, cited in Bigelow and Pargetter 1990: 60). Universals themselves can have properties and relations. Pure, dimensionless numbers arise as ratios between quantities: $5.1=(10.2\text{ gm}/2\text{ gm})=(204\text{ cm}/40\text{ cm})$, for instance. Ratios are thus *relations of comparison* between quantity-universals. Being relations, they are also universals, since one of the marks of relations is that they too can relate differing collections of universals: the comparison between 10.2 gm and 2 gm is the same as that between 204 cm and 40 cm.

Is this realist position on numbers Platonism? In the sense that it is a variant of Plato's view on universals, the answer is yes; but in the sense of 'abstract' we defined earlier, numbers are not abstract, since they emerge from the quantities whose possession is necessary for the behavior of bodies as described by the laws of nature. In that sense, the answer is no.

What are we to make of zero as a comparison between quantities? Here one can once again draw attention to the unique role for the zero signifier, namely, that unlike other real number signifiers it is *not* referential. Zero would be an existing universal like the other real numbers provided that it arose as a comparison-ratio between quantities. But what is one to make of zero quantities such as zero grams? We suggest that there is no such quantity-universal as zero grams. To describe something as having mass zero grams is to assert nonexistence again, the absence of any power to influence the behavior of bodies involved in interactions in which mass is causally relevant.

But here we have an interesting point: that the story about zero is not quite the same in every causal law. It is like the situation when directional counting was considered. Some but not all quantity-kinds (kinds of universals), such as mass, length, temperature, and duration, come with an absolute zero. There are no negative masses, lengths, temperatures, or durations. So having zero mass, length, temperature, or duration is the absence of mass, length, temperature, or duration. But there are positive and negative charges. So zero charge means something else, something more like zero worth. Here we can distinguish charge from 'net charge'. To describe a body as having zero net charge is not necessarily to describe it as having zero charges, it is not a simple denial of existence. It is to say that such positive and negative charges as it might have 'cancel out', in the sense that it behaves exactly as a body lacking all charges behaves. Notice a difference from the concept of zero net worth described earlier: having zero net worth does not mean that one behaves the same as if one had no assets and no debts. But notice also the core feature of a denial of existence, albeit a conditional denial: *if* the charges are allowed to neutralize one another, *then* the body lacks a charge of any kind.

There is considerably more that could be said about numbers and universals, but enough has been said to make our point. There does exist a fairly straightforward account of how number signs function in the world, and it is a referential account. But its referentiality derives from the scientific discovery of continuously distributed quantities and the physical laws in which they appear; and especially its referentiality does not arise from counting or exchange of discretely-occurring commodities. Furthermore, the zero sign plays a distinctive role in denying existence, sometimes conditionally, and thus is not referential. Finally, these semiotic facts can largely be inferred from the nature of the mathematics describing the properties of the numbers themselves. From the semiotical point of view, this is hardly surprising since mathematics is seen as a textual activity arising from human practice in the world. But notice also the particularly *simple* codes in play here: number-nouns, addition, subtraction, multiplication, division, and equality. This argues for a *simple* account, arising from *universal* human practices such as counting, exchange, and natural science. The nature of the end product ought to contain clues to its origin, as in any respectable scientific theory. That is only to say that the situation is a dialectical interplay of theory and practice.

Infinity

In *Ad Infinitum* (1993), Rotman offers a somewhat revised account of the semiotic subjects involved in mathematical activity. Instead of Person/Mathematician/Agent we now have Person/Subject/Agent (see, e.g., 1993: 8). The Person is the entity which refers to itself with the indexical expression 'I', and which understands mathematical metacode, which is to say informal mathematics. The Subject is that entity or subfunction which understands formal mathematical propositions or code, while the Agent retains its role as a manipulator of signifiers at the subcode level. One significant difference in the role of the Agent, however, is that it is no longer required to perform infinite tasks or simulations. Rotman is unequivocal in appealing to the Nietzschean Philosophy of the Body (or, in analytical terms, naturalism and materialism) to reject the actual infinite. He argues that the Agent would have to be imagined as embodied or corporeal, which rules out infinite tasks (see, e.g., 1993: 10,16). But as John Bell (1995) argues, this is to place a severe and immediate block on what an Agent could do in the way of imaginative manipulation. But it can be argued from a weaker premise that no natural function of a finite material person, such as an Agent would have to be, could be expected to make infinite constructions.

Rotman links belief in actual infinities with Platonism, and sees his task to reject both (see, e.g., 1993: 10,158). But it is not at all clear that belief in the actual infinite is incompatible with a careful non-Platonist realism. For example, our most successful theories about the mathematical properties of the manifold of physical spacetime assert that every spacetime interval contains a nondenumerable infinity of points, while quantum field theory has long endeavoured to manipulate infinite quantities appearing in its equations in an operational fashion, using the device of renormalization. Note that this is a species of the realist hypothetico-deductive argument from applied mathematics.

Even so, one might still feel that infinities remain an epistemological problem for the naturalist for a deeper reason. After all, how could a finite creature even *understand* the infinite if it has only a finite number of bytes to encode the concept? Now this can look insurmountable if one conceives the task of understanding infinity in mathematics as the task of encoding an infinitely large construction. In the sense of separately representing and encoding an infinite number of distinct atomic parts, this would seem to be beyond finite creatures, or at any rate creatures with a finite number of functional parts, such as discrete automata. Nonetheless, we should reflect on the similarly between infinity and zero. Like zero, infinity signifies a negative existential: there *does not exist* a counting of the whole collection (while unlike zero there do exist countings of subcollections). So, given that one has a concept of negation, the problem of understanding the infinite in general terms is reducible to the problem of understanding finite counting in general terms, which is something we seem to have a good purchase on. The matter does not rest there, however, because it can obviously be argued in reply that we cannot fully grasp finitude until we grasp all the infinite distinct instances of finite countings. We do not propose to pursue this any further, only to caution that the epistemology of infinity might not be so impossible to deal with.

There is, however, another point against Rotman. This is that, contrary to the prospect which semiotics might be thought to hold out for understanding what distinguishes mathematics from other textual activities, Rotman appears to succumb to revisionism. Despairing of the infinity of the infinite series of natural number signifiers 0,1,2, ... , he proposes to replace it by a finite series of signifiers 0,1,2, ... $, where $ represents an unspecified upper limit of actual finite countings. This interesting approach constitutes what he calls 'non-Euclidean arithmetic' (e.g., 1993: 115). While Rotman is unclear about what specifically Euclidean is being rejected, the approach falls within the established and respectable problematic of mathematical finitism. Nonetheless, it seems to amount to a proposal to declare some existing parts of mathematics false, the parts which would refer to actual infinities. As such, it is open to the objection urged earlier against revisionisms, namely that it fails to tell us what is mathematical about the rejected parts, false though they may be.

We conclude, then, that in rejecting infinities Rotman ultimately falls back into the analytical problematic of truthmakers he was seeking to transcend. This brings us back full circle to the task of bringing semiotics to bear on understanding mathematics in general, and in the final section we turn to that.

General mathematics

Here we want to say that we view the situation with mathematics in somewhat the way John Passmore viewed aesthetics in his well-known paper 'The dreariness of aesthetics' (1951). Passmore rejected the aestheticians' preoccupation with high-level theorizing about art on the grounds that its very generality was self-defeating in that it tended to lose what is interesting about art, that is, its richness and complexity. Passmore did not claim that no very general answers might fortuitously be found, only that such discoveries were apt to be platitudinous and not really why we were in the game anyway: what aesthetics is and ought to be about, is strongly continuous with art theory.

To apply this to mathematics needs an application of the idea of *difference*, namely a sense of the richness and diversity of mathematical texts. Edwin Coleman, in his Ph.D. thesis, *The Role of Notation in Mathematics* (1988), pointed out that even a brief perusal of the differences between, say, a page from Euclid, a page from Whitehead and Russell's *Principia Mathematica* (1910), a page from a text in business mathematics, a page from a standard calculus text, and a page from a mechanical engineering text, will convince one that these differences are richly textual and at the same time the very stuff of mathematics (see Coleman 1990: 131–136). Consider, for example, the varying role of diagrams, and of natural language text therein (*Principia Mathematica* used precious little of these). Now on the one hand, it is an essentialist mistake, identified by Wittgenstein among others, to think that there *must* be something general and yet interesting in common between all of these. But on the other hand, a central question about mathematics is how the varieties and possibilities within text contribute to mathematics, indeed, how textual understanding *constitutes* mathematical understanding; or to put it conversely, how the mathematical understanding is distinctively textual and symbolic. This is very much in the spirit of Rotman's *Signifying Nothing*, but even *Ad Infinitum* supports the point (1993: 33):

> ... no account of mathematical practice that ignores the signifier-driven aspects of that activity can be acceptable. It is simply not plausible — either historically or conceptually — to ignore the way notational systems, structures and assignments of names, syntactical rules, diagrams, and modes or representation are constitutive of the very 'prior' signifieds they are supposedly describing.

Coleman, who must be credited with grasping the bearing of semiotics on the foundations of mathematics independently of Rotman, argues at length that mathematics is very significantly a textual phenomenon. The complex varieties of mathematical texts and codes, with their various uses and language-games, argue for an explanation which avoids a simplistic referentiality. While of course there are many connections between mathematical codes and the extra-linguistic world, it is precisely the varieties of these connections with the varieties of useful distinctively mathematical styles of texts, which need understanding. Reference must thus play at best a very secondary role in that understanding.

These observations point to an expanded conception of the role that the general theory of signs, symbolism, and notation ought to play in understanding the nature of mathematics. But they are also applied against a more traditional view of the general nature of mathematics: formalism. Hilbert took the view that in order to display the correct formal relationships which justify mathematical propositions, mathematics should be reconstructed as a formal uninterpreted calculus employing axioms and rules for mathematical theories. In a similar fashion, though with somewhat different philosophical aims, Whitehead and Russell took it upon themselves to provide formally correct definitions of the basic mathematical concepts so that mathematical propositions could be seen as deriving from these by means of logically valid arguments, the program of logicism. But we commit an error if we think that such formalizations reveal the real mathematics. Hilbert and Whitehead and Russell, being twentieth-century thinkers, are comparative latecomers on the mathematical scene. *Principia Mathematica* looks quite *unlike* almost all of mathematics written beforehand, or currently for that matter. Even the idea of logical consequence or following from, on which *Principia Mathematica* was based, must undergo a radical shift in light of the function of diagrams in mathematics, which Barwise (1994) and others have pointed out recently. Understanding mathematics in general requires understanding mathematical difference as much as formal sameness. What needs to be understood are the possibilities for difference while remaining within recognizably mathematical codes, and these differences are notably textual. But also, importantly, there follows a certain anti-foundationalism about mathematics. Elucidation of mathematical difference is anti-reductionist in spirit, unless the varieties of mathematical text yield to a *simple* underlying explanation, which looks to be more unwarranted essentialism. Coleman has gone on to develop these themes in a number of studies (1990, 1992, 1995).

We note two other thinkers who have addressed textual issues in a way which we think has been relevant to understanding the complexities of mathematics. One is Nelson Goodman, in his justly admired *Languages of Art* (1981). Goodman does not discuss semiotics explicitly; but he refers approvingly to Peirce and Morris, and his themes are certainly semiotical in orientation, while having a distinctively analytical viewpoint. The other is René Thom,[5] well known for his contributions to catastrophe theory. Thom's aims are somewhat different from ours. One might characterize the difference by saying that we apply semiotics to understanding mathematics, especially foundational themes; while Thom applies mathematics to understanding semiotics, especially Peircean concepts. Having said that, we must acknowledge that Thom was aware of the limitations of the Hilbert Program in catering for geometric objects, linking it with the tendency to see only a chain of signification which we have also criticized:

> ... une théorie très en vogue sur la place de Paris prétend: il n'y a pas de signifié, il n'y a que signifiant; chaque signe réfère à d'autres signes en une régression sans fin ... Les mathématiciens, dans leur souci d'éliminer tout appel à l'intuition géométrique, ont connu, avec le programme de Hilbert, la même tentation. (Thom 1980: 197)

Supporting this point Thom offers, *inter alia*, formal accounts of the Peircean concepts of icon, index, and symbol, with the aim of showing that spatiality is amenable of semiotic analysis and application. In a later work (1983, especially ch. 14, 'Semiotics'), Thom develops this direction, arguing that in iconicity there is a particularly evident mutual generation between signifier and signified.

However, enough has been said by now, we believe, to make our point that the way is clear for a textual and notational conception of mathematics. The view of mathematics as symbol, notation, and text points away from narrow formalism and logicism, as we have seen. It also points away from a Platonism of abstract objects. Texts are produced by humans interacting with the world and one another, while abstract objects are unworldly. It would be perverse to turn one's back on the primary phenomenon to be explained, in seeking an explanation of it.

We also conclude that while there are difficulties in the details of Rotman's views, we must acknowledge the interesting direction of his contribution.

With some apprehension about avoiding the Scylla and Charybdis of essentialism and platitude, we offer some concluding observations on the central role of symbolization in mathematics. Certainly mathematics has a power to exclude the extraneous in order to achieve generality and purity of focus. But philosophy is also general; and philosophy and mathematics alike aim at certitude, clarity, and rigor. What distinguishes mathematics from philosophy are the symbolic means by which this is achieved. By this, for want of a better not-too-theoretical term and *pace* Peirce, we mean the ordinary sense of symbolic which contrasts with everyday language. This is where spatial representation must be taken into account, for it is very difficult to give a mathematics lecture which is purely spoken, while very easy to give a mathematics lecture which is purely written. Note that there remain large distinctions to be elucidated between the varieties of the geometric and the otherwise merely symbolic in mathematics, as well as the large roles played by the computational and the deductive. Note also the difference between mathematics and ordinary language or discourse: while Coleman is undoubtedly correct that mathematics derives from ordinary discourse and makes heavy use of it, it is nonetheless generally not so difficult to give a lecture on nonmathematical themes which is purely spoken. Of course the nonspoken use of iconography might be otherwise relevant: our mental lives are thoroughly permeated with the nonverbal, just as they are with the verbal.[6] It is, on the other hand, generally much easier to give a wholly spoken philosophy lecture than a wholly spoken mathematics lecture, as Socrates amply demonstrated. Another dimension of the difference between spatial representation and auditory representation is manifested in the relation between mathematics and music: music, one might say, is auditory geometry. This would in passing account for why many mathematicians have felt an affinity with music, and why mathematics has an aspect of the beautiful.

Notes

1. Theorem: No theory of first order logic plus identity, whose only nonlogical axioms are negations of identity statements (i.e., $-(a=b)$, $-(b=c)$, etc.) will contain any consequences which are identity statements (i.e., of the form $(e=f)$), unless these are already theorems of first order logic plus identity (such as $(a=a)$, $(b=b)$, etc.).

   Proof: There will always be a countermodel to $e=f$. Let the domain be all names, as in standard completeness theorems for first order logic. Let the interpretation of an identity statement be true just in case the names flanking the identity sign are the same symbol, otherwise false. Then the interpretations of $-(a=b)$, etc., are all true, but $(e=f)$ is interpreted as false, unless its interpretation was already fixed by first order logic. QED.
2. But if the primary phenomenon in counting is assertions of the form 'There are n Fs in W', then why mislead the youth by having *nouns*, which look like they function referentially? The presence of the noun 'two' looks like it fills the place after the equals sign in 'The number of Fs in W= … '. In standard logical notation, such a construction permits one to deduce 'There exists an x such that x is identical with two', and it is near enough to Platonism to be told that the number two exists. But the explanation lies in the human tendency to nominalize, to turn such place-holding devices into a series of nouns. After all, what varies in the collection of statements of the form 'There are two marbles in … ', 'There are three marbles in … ' is the place occupied by 'two', 'three', … . That is the part we are interested in, in answer to the question 'How many Fs in W?'; the rest is constant in all the answers. It simply saves time, then, to answer with one of the collection of words which fit the place. The logical device which records our ability to speak in an apparently referential way yet avoid existential seriousness is substitutional quantification. 'There is an x which is G' (such as 'There is a number which is the number of marbles in the tin') is construed as meaning 'Some noun can

be substituted for x in "x is G" in order to make a true statement'. There remain existential questions about what are the truthmakers for statements of the form '(noun) is G', but at least we are freed from thinking that the human use of nouns is *unavoidably* referential. There also remains the accompanying epistemological issue of how the atomic sentences of the form '(noun) is G' are to be known. The literature on substitutional quantification is extensive; on its use in avoiding numbers see, e.g., Priest (1983), who offers a conventionalist epistemology of the atomic sentences of pure mathematics.

3. There are further technical reasons for holding that quantities are prior to numbers; see, e.g., Mortensen (1987).
4. There was a notable difference between Plato and Aristotle on the status of universals which are not instantiated, such as a mass so heavy that nothing happens to possess it (cf. also recent questions in physics about the 'missing mass' of the universe). Plato believed in uninstantiated universals, Aristotle denied it. This has led to some interesting recent debates concerning the status of laws governing the possible behavior of bodies possessing uninstantiated properties, which, however, goes beyond our present semiotical concerns. See, e.g., Armstrong (1978).
5. We are indebted in this paragraph to comments by Thomas A. Sebeok.
6. See, e.g., Mortensen (1989).

References

Armstrong, David (1978). *Universals and Scientific Realism*. Cambridge: Cambridge University Press.

Barwise, Jon and Etchymendy, John (1994). *Hyperproof*. Cambridge: Cambridge University Press.

Bell, John L. (1995). Review of Rotman's Ad Infinitum. *Philosophica Mathematica* 3, 218–221.

Bigelow, John (1988). *The Reality of Numbers*. Oxford: Clarendon Press.

—and Pargetter, Robert (1990). *Science and Necessity*. Cambridge: Cambridge University Press.

Coleman, Edwin (1988). The role of notation in mathematics. Unpublished Ph.D. dissertation, University of Adelaide.

—(1990). Paragraphy. *Information Design Journal* 6 (2), 131–146.

—(1992). Presenting mathematical information. In *Designing Information for People*, R. Penman and D. Sless (eds). Canberra: ANU Press.

—(1995). *Signs and Powers: The Semiotics of Mathematical Practice*. (Forthcoming.)

Dummett, Michael (1977). *Mathematical Intuitionism*. Cambridge: Cambridge University Press.

Forrest, Peter and Armstrong, David (1987). The nature of number. *Philosophical Papers* 16, 165–186.

Goodman, Nelson (1981). *Languages of Art*. Brighton: Harvester.

Lakatos, Imre (1976). *Proofs and Refutations*. Cambridge: Cambridge University Press.

Mortensen, Chris (1987). Arguing for universals. *Revue Internationale de Philosophy* 160, 97–111.

—(1989). Mental images: Should cognitive science learn from neurophysiology? In *Computers, Brains and Minds: Essays in Cognitive Science*, P. Slezack and W. Albury (eds.). Dordrecht: Kluwer.

—(1994). *Inconsistent Mathematics.* Dordrecht: Kluwer.
Newton, Isaac (1728). *Universal Arithmetick*, second edition. London: Longman.
Passmore, John (1951). The dreariness of aesthetics. *Mind* 60, 318–325.
Priest, Graham (1983). An anti-realist account of mathematical truth. *Synthese* 57, 49–65.
Putnam, Hilary (1979). Philosophy of mathematics, a report. In *Current Research in Philosophy of Science*, 386–398.
Rotman, Brian (1987). *Signifying Nothing: The Semiotics of Zero*. London: Macmillan Press.
—(1988). Toward a semiotics of mathematics. *Semiotica* 72 (1/2), 1–35.
—(1993). *Ad Infinitum*. Stanford, CA: Stanford University Press.
Saussure, Ferdinand de (1974 [1916]). *Course in General Linguistics*, ed. by C. Bally and A. Sechehaye, trans. by W. Baskin. London: Peter Owen.
Thom, René (1980). L'espace et les Signes. *Semiotica* 29 (3/4), 193–208.
—(1983). *Mathematical Models of Morphogenesis*, trans. by W. M. Brookes and D. Rand. Chichester: Ellis Horwood. (See especially ch. 14, 'Semiotics'.)
Whitehead, Alfred North and Russell, Bertrand (1910). *Principia Mathematica*. Cambridge: Cambridge University Press.

Chris Mortensen (b. 1945) is a Reader in Philosophy at the University of Adelaide, Australia ⟨cmortens@arts.adelaide.edu.au⟩. He is interested in logic, metaphysics, and the philosophy of science. His publications include *Inconsistent Mathematics* (1994) in addition to numerous articles and reviews.

Lesley Roberts (b. 1963) is a graduate student in the Department of Philosophy at the University of Queensland ⟨lroberts@lingua.cltr.uq.oz.au⟩. She is principally interested in logic and the philosophy of language.

Australasian Journal of Philosophy
Vol. 76, No 2, pp. 182–197; June 1998

ON THE POSSIBILITY OF SCIENCE WITHOUT NUMBERS

Chris Mortensen

Part 1: Field on Numbers

I. Introduction

Hartry Field's well known book *Science Without Numbers* (1980) was an important contribution to the debate on realism and platonism about numbers. Field's exploitation of the notion of conservativeness was a particularly significant innovation. However, there is a difference between realism and platonism; and one aim of this paper is to support the former while disputing the latter. I will explore Field's nominalist strategy, and argue that it is both unnecessary and unlikely to bring about the results he desires. After Field's position is disputed, we see that non-platonist numbers play a distinctive role in securing metrical realism or anticonventionalism in basic physical theories.

By realism about numbers is meant the claim that there are numbers. By platonist realism (platonism for short) about numbers is meant the claim that there are platonist numbers, that is there are numbers construed as having no spatiotemporal location and especially no causal powers. One point of agreement with Field is stressed, namely the intention to deny the existence of the causally irrelevant. If numbers prove to be so, then they should be dispensed with; and any claims for their indispensability, such as have been made by Quine and Putnam, should be resisted. Field also describes his position as 'nominalist', and some of the literature discusses its legitimacy. This paper tends to use 'nominalist' and 'physicalist' interchangeably, since as Field notes there isn't much in the word, and the methodology can be applied to specific issues such as the reality of numbers without forcing other concepts under the umbrella. Strictly speaking, Field holds up as undesirable several marks of platonist entities such as numbers: as well as acausality, there are the related problems of unreferentiality and unknowability, especially inability to explain the reliability of mathematicians' beliefs. One can raise questions about which of these imply which, for example whether acausality implies unknowability. But while such issues are important, they are not the project of this paper. Hence we set them aside, and it is taken that it is undesirable to have any one of these marks.

This paper is in two parts. In the present Part 1, Sections II-IV, Field's nominalist strategy is considered and disputed. It is argued that there is a simpler strategy available to the antirealist. Field has considered this strategy and rejected it, and the main aim of Part 1 is to rehabilitate it. It is also argued that any antirealist or antiplatonist strategy, either Field's or the alternative of this paper, must address the additional question of whether it is desirable or correct to eliminate the additional entities. In Part 2, this question is taken up; and it is argued, in line with recent work in Australia, that numbers can be seen as causally and epistemically virtuous. This is particularly evident in spacetime theory, where several antirealist stories are considered and rejected. Thus, far from being able to eliminate numbers, one would want to have them in one's physical theory.

II. Conservativeness

Field's principal claim about the specifically mathematical parts of basic physical theory, is that it is conservative. A clear statement can be found in Field's *Realism, Mathematics and Modality* [1989]:

> (1) A mathematical theory S is conservative if, for any nominalistic assertion A and any body of such assertions N, A is not a consequence of N+S unless A is a consequence of N alone. (p. 125).

In other words, if A follows from N+S then A follows from N alone. A nominalistic assertion is one 'whose variables are all explicitly restricted to non-mathematical entities' (p. 125). Notice in passing the identification of the mathematical as the target, rather than the specifically platonist–mathematical. Field then argues, using what he calls a representation theorem, that the mathematical part of Newtonian gravitation theory is conservative with respect to a certain non-mathematical theory (written in a language of betweennesses, congruences, part–whole etc.), which he claims can be regarded as expressing the whole non-platonist content of gravitation theory. From this it will follow that mathematics is dispensable from Newtonian gravitation theory, and that the Quine–Putnam indispensability claim is undermined. Field expresses the opinion that there should be available similar representation theorems/conservativeness results for other basic physical theories such as GR gravitation theory. Brent Mundy subsequently developed representation theorems covering the conditions in which a quantity-space with weaker structural properties than a continuous metrical comparison between quantities, can be embedded in a continuously metrically ordered quantity space (see e.g. (1987a), (1987b)). Here he and Field were extending the well-entrenched program of measurement theory (see e.g. Adams (1979)).

If Field is right overall, this in turn will open the way for the *instrumentalist* claim that mathematics is false but useful: false on the grounds that we have no reason to believe in its existential claims; but useful because of its conservativeness, in that mathematics can be used as a deductive shortcut without fear that we will generate any more nominalistic statements than appear in the approved nominalist theory anyway. Thus, universities can still justify funding mathematics departments.

Some controversy has taken place over the issue of whether the consequence relation alluded to in the definition of conservativeness is proof-theoretic or semantic. These two are co-extensive (though not identical) in the case of classical first order logic, but not co-extensive in classical second order logic. Field clearly holds this issue to be important and discusses it at length in various places. His original preference was for a second order semantic account, since his representation theorem is model-theoretic in character. In the end he seems to come down on the view that consequence is a primitive modal relation, whose meaning is explicated by a collection of (effective) rules for deducibility and nondeducibility (i.e. consistency, a semantic notion which is not effective at the first order level). However, I will seek to avoid this controversy also, by arguing that a sufficient conservativeness result can be obtained for any of these.

III. Is Mathematics Trivially Conservative?

It might be wondered whether mathematics can be trivially shown to be conservative over non-mathematical physical theory. In order to develop this, it will help to define the concept of a conservative extension of a theory. Suppose we have theories Th1 and Th2, where the language of Th2 is a subset of the language of Th1. Then:

(2) Th1 is a *conservative extension* of Th2 iff any consequence of Th1 in the (sub-) language of Th2 is also a consequence of Th2.

Now we have a simple result, alluded to by Field, (1989, p. 129).

(3) Any mathematical-and-physical theory has a purely physical subtheory of which it is a conservative extension.[1]

This means that no sentence of the nominalist sublanguage is in Th1 without already being in Th2. Any deduction (according to whatever canons of deducibility, semantic, proof-theoretic, first-order or second-order, intuitive or modal) of nominalistic/ physicalistic conclusions from premisses involving a mix of mathematical and physical predicates, will already be deducible in the preferred nominalistic subtheory. Hence, it is open to us to hold that the existence claims found in mathematics are false, while at the same time being able to explain why they might be useful to simplify calculations without adding anything to the "objective" sublanguage. Mathematics can be used where it simplifies computations, in the sure knowledge that there is a purely nominalist subtheory which reflects all and only the nominalist truth. Isn't this 'false but useful'?

Field is aware of this result and rejects it. To appreciate his reply, it will be useful to restate his claim that mathematics is conservative in the above terms. To say that a mathematical theory S is conservative, is to say that for any nominalist theory N, if any A in the language of N is a consequence of N+S then A is a consequence of N alone. That is, S is conservative iff for any nominalist theory N, N+S is a conservative extension of N. So Field's claim that the mathematics of physical theory is generally conservative, amounts to the claim:

[1] Proof: Let Th1 be any theory (i.e. closed under consequences) in language L1, where L1 contains both mathematical and physical predicates. L1 has a sublanguage L2, consisting of all sentences and predicates constructible solely from physical predicates and relations, and let Th2 be the intersection of Th1 and L2. Clearly, Th1 is a conservative extension of Th2, since any A provable in Th1 in the language of L2 is therefore in Th2. It remains only to show that Th2 is a theory, i.e. that Th2 is closed under consequences. Suppose that A1 ... An are in Th2 and that A1 ... An turnstile B where B is in L2. Since Th2 is a subset of Th1, all of A1 ... An are in Th1; so since Th1 is a theory, B is in Th1. But B is in L2, so B is in Th2. That is, Th2 is a theory, and thus a subtheory of Th1. Note that this strategy is essentially that of Craigean transcriptionism. Indeed, in light of Craig's Theorem it is also true that if Th1 is axiomatisable so is Th2.

(4) Any (consistent) mathematical theory when added to any nominalist/physical theory produces a resulting theory which is a conservative extension of the nominalist theory.

Thus, mathematics doesn't generate any more nominalist consequences no matter what nominalist theory it is added to. This is certainly worth calling a dispensability thesis. Contrast it with (3) above, which is that any theory of mathematics-plus-physics has a purely physical subtheory of which it is a conservative extension.

Field contrasts these two propositions, and argues against the usefulness of (3) for the nominalist program. In *Science Without Numbers* he describes such nominalist theories as 'bizarre trickery', 'obviously uninteresting since they do nothing whatever toward explaining the phenomena in question in terms of a small number of principles', and contrasting with theories which are 'reasonably attractive' (p. 8, see also p. 41, 47). In *Realism, Mathematics and Modality*, he says:

> ... the conservativeness of mathematics tells you what happens when you add mathematics to nominalistic theories: it doesn't say anything about the availability of sufficiently interesting nominalistic theories (p. 129)

also:

> [(3) above] or any strengthening thereof is an assertion about the existence of a sufficiently wide variety of nominalistic theories, and this is something that the assertion that mathematical theories are conservative does not claim. (p. 129).

He also refers to such nominalist subtheories as 'unnatural' (p. 133). Geoffrey Hellman (1989) similarly argues that the existence of the nominalist subtheory is 'not sufficient for "good systematisation", not to mention other aspects of "attractiveness"' (p. 135).

I must say that I fail to see the force of these arguments. It is true that there is a difference between (3) and (4). But Field's arguments do not, I suggest, give a clear reason for following his particular route through conservativeness to dispensability. One obvious point in reply is that while one does not have an *a priori* guarantee that the nominalist subtheory guaranteed by (3) is theoretically attractive, such attractiveness is not *ruled out*. That would take further case-by-case arguments which Field does not supply.

Second, let us note that the nominalist subtheory guaranteed by (3) has a feature of *maximality*, in that it is *all* the nominalist consequences of the platonist theory (see proof in footnote 1). But now, any time Field succeeds with his own strategy of finding an attractive nominalist subtheory of which platonist mathematical physics is a conservative extension, then, if Field's attractive subtheory happens to be identical with the maximal nominalist subtheory, we have precisely a counterexample to Field's unattractiveness claim. And if Field's whole program were to succeed on this basis, then what he would have shown is that the maximal nominalist subtheories are attractive throughout the whole range of his chosen platonist theories.

Still, it must be conceded that it is not *guaranteed* by (3) that the maximal nominalist subtheory is theoretically attractive. So one should leave open the possibility that the nominalist subtheory so obtained is theoretically unattractive. But one should also ask whether that matters. If Field's nominalist subtheory is attractive, then it is distinct from the maximal subtheory, and thus a proper subtheory of *that* also. But now we would be in the situation where Field's version of nominalism *fails to decide* where the maximal subtheory decides. At least that makes the maximal subtheory *more testable*. Even if the test should fail in empirical prediction, there is every reason to include the result in the maximal physical subtheory of a revised platonist mathematical physics, as is current practice. Remember, too, that the maximal subtheory is all the nominalist consequences of platonist mathematical physics, which can thus be expected to be well-confirmed. Hence, if the two programs diverge, it seems like Field's version is at a disadvantage in that it under-describes reality.

One final argument is this. If the properties and relations described in the preferred nominalist sublanguage amount to the sum total of the causally relevant relations, if these are the sum total of things which *make a difference* to laws and especially observations, then does it matter overmuch if our pure theory of them is complex and unattractive? One wouldn't care overmuch if the larger *platonist* theory is simpler, needless to say: Field and I agree that this is no reason to opt for platonism. In choosing between nominalist subtheories, an overriding consideration is surely what we take to be *true*; and if our best confirmed platonist theory says that certain nominalist statements are true, then so much the worse for scruples about unattractiveness. What would have been shown at most is that reality is unattractive,[2] whatever that might be. But there is no metaphysical minus in that. Principles of theoretical attractiveness such as simplicity are not guaranteed by the simplicity of the universe. They are desirable on epistemic grounds, because it is sensible to begin with the simplest theories and only complicate them if one is forced to do so. Of course, it is conceded on all sides that there are *epistemically* attractive presentations of nominalist theories, namely their platonist supertheories.

Thus the alternative strategy of appeal to the maximal nominalist subtheory takes note of the point that for the program of showing that mathematics is dispensable and false but useful, there is no evident reason why one should do any more than separate mathematics from true physical theory. Why would it matter if mathematics were indispensable from false physical theories? From some it might be, from some it might not. Field's necessitarianism has led him to overlook the contingency of his thesis, I think. But then, if true physical theory is all that one must eliminate mathematics conservatively from, why not claim that true physical theory is simply the maximal physical subtheory of our best platonist mathematical physics? It is what useful mathematics has revealed to us, though it is not itself mathematics. It has a remarkably simple, attractive and general presentation of course, namely modern mathematical physics, wherein answers to properly presented nominalist questions can be deduced. And finally, there is the advantage that such a nominalist argument can be made out generally, exploiting (3)

[2] This is what distinguishes the present issue of mathematical realism versus mathematical instrumentalism, from the issue of realism about the external world versus phenomenalism and Craigean transcriptionism. In the latter issue, the additional entities, space, time, matter and its properties, are causally and explanatory relevant: vary their distributions and observations vary in systematic ways. So one is more loath to lower the boom on them compared with acausal, acounterfactual, eternal numbers. Physical objects do amount to an expansion of ontology over phenomena, but more than make up for it in lawlike generality.

above, instead of proceeding on Field's program to extend his result to theories stronger than Newtonian gravitation theory (see also below, section VII).

This strategy appears so simple that there must be a catch. And there is, of sorts. The catch arises from the fact that this trick obviously can be pulled for *any* sublanguage, and so what licenses one to claim that a certain sublanguage is the preferred nominalist one? Well, clearly more has to be done, namely that an argument is owed about why the preferred class of predicates and relations exhausts the class of causally relevant predicates and relations. But this is something which Field cannot escape either. That is, I am saying that there can be arguments about whether Field's own chosen set of nominalistically-preferable relations express all the physical aspects of reality, and that indeed for *any* dispensability strategy to succeed such arguments must be supplied.

IV. Three Problems and Replies

In (1996), James Hawthorne points out a problem directed at Craigean transcriptionism by Van Fraassen, which Hawthorne applies in three ways to the case before us of mathematical instrumentalism. In this section we consider these problems, and see that replies are available.

First, we could have a theory in mathematical physics which, in quantifying over mathematical entities, implies that there are entities which *fail* to be at any spacetime point p. The latter is a statement which can be made in the purely physical vocabulary: (Ex)(for all points p)(- x is at p). One would therefore expect it to be in the maximal physical subtheory. However, a physicalist, who believes that physical descriptions exhaust reality, ought to reject the statement, since it asserts that there are entities not in spacetime.

As far as this goes it isn't very worrying. After all, mathematicians do not bother to insert into mathematical theories assertions to the effect that mathematical entities are *not located*. This latter claim is something that metaphysicians typically assert for completeness when observing that mathematical theories simply *lack attributions of location* to mathematical entities. So in real life mathematical theories it isn't obvious that the physicalist has any such problematic consequences to worry about.

Even so, it remains a *potential* problem that there might be such unintended and unacceptable consequences stateable in the physicalist sub-language. Hawthorne points out a second, similar problem which needs somewhat different treatment. The pure physical theory may contain the universal closures of its laws, for example the assertion (for all x)(Ep)(x is at p), or the assertion 'Everything attracts everything else with a gravitational force'. But then, what sort of mathematical theory would be a conservative extension of this? It would attribute position and gravitational attraction to numbers and sets.

The remedy suggested by Hawthorne, following a strategy in *Science Without Numbers*, is plausible. To begin with, one must have a sense of which of one's predicates count as physical and which mathematical. It is to be expected that anyone taking an instumentalist position would have a position on which concepts are to count as objectionable, for instance. Then one needs to require that one's theory contains information pertaining to this distinction. The easiest way to achieve this is to suppose that, or to make it that, *all quantifiers are relativised*. That is, let there be predicates "Px" for 'x is physical' and 'Mx' for 'x is mathematical'. Then let no universal quantifier occur except in the contexts '(x)(Px→...' and '(x)(Mx→...' and '(x)((Px∨Mx)→...'; and let no existential quantifier occur except in the contexts '(Ex)(Px&...' and '(Ex)(Mx&...' and '(Ex)((Px∨Mx)&...'. Then, the maximal physical subtheory will not contain the assertion (Ex)(Mx&(p)(- x is at p)), for that does not appear in the supertheory with relativised quantifiers. One can then if one wishes add to the physicalist subset *the postulate of physicalism*, (x)Px. This will produce a theory with unrelativised quantifiers once more, sufficient to gladden the heart of the sternest physicalist. There is a minor technical hitch here, in that the strategy of relativising all quantifiers has the consequence that neither the mixed superset of sentences nor its physicalist cut-down are *theories*, i.e. deductively closed sets of sentences, since all unrelativised consequences will fail to appear. But, as Hawthorne shows, this is avoidable by recourse to the device of a many-sorted language. The details are omitted here.

A third problem raised by Hawthorne concerns the strategy of conjoining or adding two theories to one another, which naturally enough scientific unity requires. But if the two mathematical superstructures are different, then one must allow the possibility that they might produce unintended physical consequences if conjoined. A worst case might be where incompatible mathematical theories were used, for then conjoining the theories would produce inconsistency.

Hawthorne shows that there are formal precautions one can take to avoid this. An intuitive strategy one can adopt is as follows. Work out the two physicalist subtheories first, then conjoin *them* and close under consequences (since mere set-theoretic union of theories is not automatically a theory). There is surely no problem for the mathematical opportunist in this. No advantages of ease of calculation have been lost, though some physical theorems might have to be proved without detour through mathematics. Furthermore, it is wise advice to proceed like this. If one doesn't believe in one's mathematical superstructure, then one should beware of extending it beyond the context in which it has proved useful.

I conclude with Hawthorne, then, that these problems can be met by the mathematical instrumentalist.

Part 2: Numbers as Causal

V. Numbers and Quantities as Causally Virtuous

The bulk of Field's work on numbers took place before a later Antipodean development on the metaphysics of numbers, namely numbers as relations between quantity-universals, an idea which can be attributed to Newton (see Newton (1728), also Forrest and Armstrong (1987), Bigelow (1988), Bigelow and Pargetter (1988), Mortensen (1987), Cheyne and Pigden (1996).) While the position is doubtless familiar to many Australasian philosophers, it is as well to re-present it for others. The basic idea is that quantities are the primary reality as revealed in our most successful accounts of physical laws. Consider Newton's law of gravitation: any two bodies of mass m_1, m_2 at distance r attract one another with a force $F = G.m_1.m_2/r^2$. Each of these variables has a dimension: mass is in some system of units such as gm, r in distance-units such as cm, F in force units such as dynes, and G is a dimensioned constant, in these units equal to 6.67 times 10^{-8} dyne cm^2

gm^{-2}. Quantities of any kind are multiplied and divided, and quantities of the same quantity-kind are added and subtracted. The unit dynes is specified by saying that one dyne is identical with the force necessary to accelerate one gram by one centimetre per second per second, or 1 dyne = 1 gm cm sec^{-2}.

Quantities 'contain' a number, so a quantity might be thought to be a kind of composite, of a number and a dimension. Similarly, multiplication and division might seem to be achieved by two operations, multiply and divide the numbers inside the quantities and then multiply and divide the units to get the dimension of the result. This might tempt us into Quine's view, which also seems to be Field's, namely that when one says that *a*'s mass is 5 gm, this is to be construed as a relation (actually a function) between *a*'s mass-in-gram and a number: *a*'s mass-in-gram = 5. However, there are a number of difficulties for the Quinean view. One is that implausible causal consequences ensue. Consider a particle having mass 5 gm and charge 5 coulomb. Then on the offending view, *a*'s mass-in-gram = 5 and *a*'s charge-in-coulomb = 5. It follows that *a*'s mass-in-grams is identical with *a*'s charge-in-coulomb. But *a* might be involved in an interaction in which only one of these aspects is causally operative, as when in a nonzero gravitational field but no electromagnetic field. It is not easy to see how to reconcile these last two propositions. A second difficulty is that the offending view has no easy explanation of the fact that *a*'s mass-in-gram = 2,000 iff *a*'s mass-in-kilogram = 2. The natural view is that this holds because '2,000 gm' and '2 kg' are names for the very same thing, a determinate quantity within the quantity kind. A third difficulty is that the offending view threatens to produce too much arithmetic. If *a*'s mass-in-gram = 5 = *a*'s charge-in-coulomb, then 10 = 5+5 = *a*'s mass-in-gram + *a*'s charge-in-coulomb. But the latter sum really would be a mathematical artifact, something which appears to have no point in existing theory since addition and subtraction operate only within the one quantity–kind.

On the other hand, the natural view of basic physical theory is that when we say that *a*'s mass is 5 gm, we are asserting an identity between quantities, *a*'s mass and the mass 5 gm, i.e. *a*'s mass = 5 gm. The existential consequences of this are that *a*'s mass and 5 gm exist. On this view, the arithmetical operations are in the first place operations on quantities, and at most derivatively operations on numbers. Quantities are universal-like in that they satisfy two intuitions often expressed in the theory of universals. First, some predications of the same predicate to different objects hold in virtue of the fact that one and the same thing, a universal, is in common to those objects. Second, some predications of different predicates of the same object hold in virtue of different but universal-like aspects of that object. Perhaps there remain dissimilarities between quantities and universals classically conceived; but these two intuitions, particularly the former, are enough to justify calling them universal-like.

But now the emergence of *something like* dimensionless numbers is virtually inevitable, once the operation of division is permitted between quantities: 10gm/4gm = 25cm/10cm = 2.5 . A dimensionless ratio can thus be thought of as a relation of comparison or proportion between quantities, a comparison moreover which is common between those quantity-kinds which are continuously distributed. This seems to have been Newton's view. (I say that dimensionless ratios are 'something like' numbers because there remain potential differences, particularly in respect of whether analogues of *all* the classically-envisaged numbers exist, or whether for example the dimensionless ratios are 'gappy', such as a gap at the number zero, or a largest finite number; but that is another story, and dimensionless ratios are certainly number-like.) There are further important issues here, for example the specific nature of the relations or functions between quantities or magnitudes such as addressed by Mundy (1988), which are not addressed here though the project seems to be compatible with the present one.

VI. Epistemological Considerations

We can now take epistemological stock. Cheyne and Pigden (1996) develop the epistemological argument against the platonist in the form of a dilemma. Either numbers have causal powers or they do not: if they do then number platonism is false; and if they do not then then the number platonist has to show how numbers can be explanatorily and predictively useful and indispensable, since their presence or absence would seem to make no difference to scientific observations. I am sympathetic to the conclusion of their argument, but one can imagine a reply which leads to something of an impasse. According to our most successful basic physical theories, quantities are thoroughly part of the causal furniture of the world. Our best laws relate quantities, and thus imply their existence. Our best laws also imply that the quantities of things are causal, in the sense that if the quantities of things were differently distributed, the observable future of the universe can be predicted to be different. Furthermore, our best laws, which certainly employ arithmetical operations between quantities, imply the existence of dimensionless ratios or numbers, platonist or not, as derivative from the existence of quantities of which they are relations of comparison. Thus at least the epistemological problem is solved; but even more strongly it might be claimed that if numbers are delivered by best physical theory, then they simply are causal, or as causal as you will get.

The trouble is that this reply is really just Quine and Putnam's argument, and plainly it can be adapted by anyone seeking a way out of the epistemic bind. So it needs emphasising that the epistemic objection to platonist numbers links their problematic epistemic status to their acausality, or at the very least to the idea that platonist numbers make no difference to which contingent laws and predictions hold in our universe. So the number causalist has to show that numbers so conceived do play an explanatory role, and do make a difference to our laws and observations. There are several points on which this can be defended, I suggest.

First, there is the role of the count-nouns 'one', 'two' etc. It is well known that counting does not by itself imply the existence of numbers as referents of the count nouns. A statement such as 'There are just two cells in the dish' is reducible to 'There are an x and a y which are nonidentical cells in the dish, and any cell in the dish is identical with either x or y.' This is a statement about cells, the dish, and identities and disidentities between them, nothing more. Thus, if all that science needed numbers for was to count, then there would be no reason to believe in their existence. But then, if the world were so structured that there were only discrete countable things in existence, say for example if space were quantized, which does seem conceivable, and if further there were a maximum number of things in existence, say 10^{79} which has been suggested for the number of electrons in the universe, then it would seem to be possible in principle to completely describe the world by a finite list of the things in it and their properties. In such a universe, there would be no reason to believe that numbers exist. Further, the jury is still out on whether our own universe is like this. If successful physical theory could decide this issue, which most seem to think is a possibility, then the non-existence of numbers must be regarded as having empirical consequences. Of course, there are true whole number *statements*, but these are made true by complex facts involving conjunctions, identities and non-identities, not number-universals.

Second, considering universes in which there are real number quantity ratios, it should be asked what effects changes in these ratios would have. One thing to note is that one and the same universal can be referred to in different ways, by '*a*'s mass' and by '5 gm'. The former mode of reference identifies longtitudinally as it were, in terms of the continuing identity of one of the bearers; while the latter identifies latitudinally, or better across all spacetime. Now there is no question that ratio comparisons between *a*'s mass and other universals certainly make a difference to which observations we can make. If *a*'s mass were twice what it in fact is, *a* would behave differently. This is even true if *a*'s quantity remains in constant ratio with a universal standard: the upshot of the 'universal nocturnal doubling' debate seems to have been that such a thing is consistent and would have detectable effects. This effect is even more marked if the background space is non-Euclidean with positive curvature, as Nerlich (1991), points out: in this case universal doubling would change the shapes of things. Hence, in this sense number ratios between universals are causally efficacious: the ratios between the quantity universals possessed by objects at a time, and other universals from the same quantity kind, make a difference to how the body will behave at that time.[3]

[3] Perhaps this is the most we can expect for the causal efficacy of number relations, since causal efficacy presumably requires properties to be subject to temporal change. Still, one can also ask whether there is any (other) sense in which relations between universals identified *latitudinally* can be regarded as making a difference to how bodies behave. After all, it might be argued that the identity of the universal 5 gm is *constituted* by the system of ratio relations to other quantities from the same quantity-kind. Indeed, precise ratio comparisons are the mark of genuine universals or genuine samenesses between things. Hence if we try to imagine such relations changing or being different, we are only imagining a change in the distribution of quantities, not a change in the ratio comparison relations between the quantities themselves. Now even here there is a qualification, namely that the pattern of numerical relations does not constitute the whole of the identity of the quantity, since this would omit the quantity-kind. Further, there is the speculation that ratio relations between quantities such as distances might change in systematic global nonlinear ways. For example, the metric of space might change from $ds^2 = dx^2 + dy^2 + dz^2$ to $ds^2 = (dx^2 + dy^2 + dz^2)/z^2$. This changes the geometry of space from Euclidean to non-Euclidean (see Grunbaum (1964) ch. 1), which would readily be detectable by its effects on the behaviour of free particles if laws of gravitation remained the same. What would also change would be the global pattern of congruences between intervals, including in particular congruences between the lengths of bodies. Again, one might suppose a systematic non-linear change in all masses, e.g. from x gm to x^2 gm or to log(x) gm, which would have differential effects on accelerations if forces such as gravity remained dependent on masses in the same way as now. Now these changes might be regarded as changes in the distribution of quantities, but they might also be regarded as changes in the ratios between existing latitudinally identified quantities. Further, this might be the simplest description of the change if it were global, since the alternative would seem to be local changes everywhere in the masses of individual bodies. Variation in ratio relations between quantities, as opposed to variation in individual quantities, thus has effects which are *global* in a way that they are not for variations in individually located quantities. A change in a body's mass is local, but a change in relative size of *whole quantities* affects all bearers of those quantities. It would seem then that if we had reason to believe that such global changes were taking place, we should suspect this possibility. Other examples of relations between universals are laws of nature, as Armstrong persuasively argued, which are also global in reach.

It is concluded, then, that there is every reason to think that variation in the possession and distribution of real number relations would have detectable effects on bodies, and thus be causal in that sense.[4]

VII. Can Field's Result be Strengthened?

We return to the question of the success or otherwise of the dispensability program. It has been argued that the *motivation* for using such a strategy has been weakened if numbers do not have the unattractive features of acausality, unreliability, etc. But a number realist of this kind must still come to terms with the consequence of Part 1, that there are always resources available to ensure the conservativeness of platonist mathematics over nominalist theories. Thus Ockham's Razor threatens the numbers unless their causal and epistemic virtues are secured. We have also noted that Field's version of the program was not carried beyond Newtonian gravitation theory. So the question arises whether his result can be extended to GR gravitation theory. After all, it is surely *necessary* for nominalism that platonist numbers be dispensable from true physical theory. It is argued in the remaining two sections that there are difficulties for making the extension, in Field's way or indeed any other.

Recall that it was argued that no matter whose strategy one adopts, it must be shown that the preferred nominalist sublanguage exhausts the nominalistically or physicalistically acceptable features of the world. GR theory as it stands is involved in numbers at least by virtue of postulating a metrical structure on spacetime, determined by the field equation of GR. A metrical structure amounts to a definition of distance along curves, so numbers are involved as soon as they emerge from distance comparisons. Thus one might get a clue to what could count as an acceptable number-free sublanguage for GR by asking what it would take to dispense altogether with a metrical structure for spacetime.

This puts us on familiar territory, because it is notable that Adolf Grunbaum was influential in denying the existence of an intrinsic metric structure for continuous spaces (see his (1964)). Grunbaum's arguments for his Thesis of Metrical Amorphousness need not concern us here, but it should be noted that Grunbaum was not in the business of

4 Jody Azzouni (1997) has recently argued for a sharp epistemic distinction between subatomic particles and numbers in terms of a distinction between *thick* and *thin* epistemic access. Thick has four significant aspects: (1) robustness—observation operates largely independently of what we believe, seeing is robust over a wide range of circumstances; (2) refinability—we have ways of adjusting our observational means of access to the thing being seen, which aims at increasing robustness; (3) trackability—sensory awareness enables us to track properties of a thing, to form an episodic history of its behaviour; (4) epistemic relevance—we can connect certain properties of things seen with our capacity to know about these properties. Sensory observation is evidently thick. Thin epistemic access, on the other hand, is what a posit enjoys merely by being the value of a bound variable in a theory with the Quinean virtues of simplicity, familiarity, scope, fecundity and success under testing. Thick implies thin, but evidently not vice versa, the numbers being an exception. Electrons and quarks fall into the thick side, by virtue of our having *instrumental* access to them, which is to say access by means of the construction of instruments to whose behaviour we have observational access. In this sense, our knowledge of them is robust, refinable etc. On the other hand, it is almost definitional that we lack thick access to mathematical items, which should be regarded as 'effects of language' (p. 483). Mathematics, as Galileo said, is the 'language of physics'. Azzouni's argument definitely advances the antiplatonist cause, I would say, but not inevitably the antirealist cause. Our access to which quantities a thing possesses, its mass, size, velocity, charge etc., is plainly thick. The thickness is obviously explained by the causal role of quantities, how they appear in basic laws. To the extent that we take quantities to lead to numbers, as argued above, then we have a strengthened epistemic role for the latter as well.

denying numbers, nor of eliminating metrical *terminology* from spacetime theory. Rather, metrical language may be used but need be not taken seriously because it contains conventions which can be varied without violating any facts of the matter. There are many equally correct metric tensor fields which yield different collections of congruences between bodies. Applying this to the problem at hand, it could be said that while number-ratios still emerge as relations between quantities (4 sec/5 sec = 0.8), nonetheless metrical language, and numbers in particular, do not reflect any objective relations between bodies. This is not quite science without numbers; but it is metric geometry where the quantification over quantities and numbers is 'semantically idle'. The universe lacks the structure it would need to have for distances and their associated numbers to have any explanatory force.

But if so, why mislead the youth with the established formalism of metric tensor fields, which *look* as if they are intended to do descriptive and explanatory work in connection with spatiotemporal relations between observable objects and events? At least, show that metric tensor fields are dispensable from physics, preferably with a conservativeness result. Now in a sense this problem was addressed earlier this century by the French mathematician Elie Cartan, who showed how to describe the affine structure of spacetime without resort to a metric tensor field. The affine structure is the structure of geodesics and curvatures, and so it is affine structure that is responsible for inertial behaviour in a gravitational field. Thus, as far as GR gravitation theory is concerned, it is open to claim that spacetime is metrically amorphous, at least up to an affine transformation. This would support the thesis that numbers arising from *metric* geometry do not reflect facts about the gravitational behaviour of bodies, hence are explanatorily idle and apt for dispensibility. Cartan's affine sublanguage would thus supply the preferred nominalist subtheory for Field's conservativeness result.

I don't say that Field is a metric conventionalist. It is only being claimed that the thesis that spacetime lacks a metric structure serves as an example of the required justification for the choice of a metric-free sublanguage, which would be adequate for gravitation theory and yet nominalist as far as *distance* quantities are concerned. Both distances and their pure number ratios would be explanatorily idle, and thus apt for dispensability by whatever version of conservativeness is to one's taste. Thus it is perhaps significant that Field chose to eliminate numbers from *Newtonian* gravitation theory, since Newtonian spacetime has an affine structure but no invariant spacetime interval. Scratch an American and you find a pragmatist, and a metric conventionalist too. (Just kidding, folks.)

Unfortunately, this interesting line of argument does not succeed in the end. We could rehearse existing persuasive criticisms of Grunbaum's arguments for his thesis of metrical amorphousness (e.g. Nerlich (1976)). It can also be pointed out that affine transformations remain ratio-preserving for distances *along the same geodesic*, so that ratios still emerge at the level of affine structure. But there are strong positive arguments from spacetime theory as well. Here the authority of Misner, Thorne and Wheeler's *Gravitation* (1973) is appealed to. In discussing Cartan, they offer a number of reasons why physics needs a full metrical structure for spacetime, rather than the weaker affine structure (see pp. 304-5). Their principal reason is that a full metric tensor field is necessary to guarantee that spacetime is locally SR as ours is, rather than locally Euclidean. In short, there is a difference between a universe in which SR holds and a universe which is Newtonian; but both are flat (zero curvature), and it is the signature of the metric tensor field which distinguishes them. This is a constraint beyond gravitation theory, since it concerns the geometrical structure necessary to describe the behaviour of light, and the structure of time-like (i.e. causal) curves.

It is concluded, then, that the case for an elimination of numbers in GR gravitation theory, via an elimination of metrical structure in favour of affine structure, is doubtful.

VIII. Differential Structure

However, there is an even deeper level of structure in spacetime theory where quantities and numbers intrude, namely its differential structure. Indeed, this level of analysis is quite general in that much physics, not just gravitation theory, can be written in these terms. So it is as well to consider it.

One supposes that there is a manifold, such as a continuous 4-dimensional collection of spacetime points. This is structured so that it is locally isomorphic to (but not identical with) the subspaces of the space of quadruples of real numbers R^4. One further supposes that there is a scalar field defined on all of the manifold. A scalar field consists of a quantity from the same quantity-kind (e.g. temperature) at every point, with the further stipulation that close points have close temperatures. The latter stipulation is sufficient to define the differential structure, namely directional derivatives, tangent vectors and tangent plane of the scalar field at the point. This amount of structure is weaker than affine structure, which requires a further concept of parallelism between vectors, or 'parallel transport', via the stipulation of an affine connection on the space. (The concept of an affine connection has some interesting implications for the theory of universals which we do not take up here; see Forrest (1996).)

We can already see, however, that numerical comparisons between the *scalar* quantities at different points are again available by virtue of the ratios, since the numerical operations are certainly permitted on scalars when defining differential structure. So it looks very much as if (nonplatonist) numbers are delivered right from the outset in basic physical theory.

Looks can be deceptive, of course. Here, we can try out one last antirealist strategy. The rational numbers hold a prospect for motivating the elimination of both the scalar quantities and their associated real number ratios. It is known that fixing all the rational ratios in a continuously ordered collection of quantities from the same quantity-kind (quantity-space) suffices to fix all the real number ratios too. Focussing for simplicity on the case of a single variable, all equational laws of nature can be written in the form $f(x)=0$ where x is a variable ranging over a quantity kind. If $f(x)$ is a continuous function then it has the property that if $f(x)=0$ for all rational x then $f(x)=0$ for all real x also. Hence all laws and relationships holding in the rational language are preserved in the real language, and there are no new laws in language of the cut-down rational theory that are not already there in the cut-down in the first place. Hence one has conservativeness of a kind, in that one could compute with real-quantity-valued continuous functions on a quantity space while aiming for rational-valued results, and be assured that all rational-ratio results so obtained would be obtainable from rational-valued premisses alone.

But now the point of focussing on the rational sub-language is that statements about rational numbers can be reductively analysed away without reference to numbers. The idea is first to translate out the ratio into the obvious comparison between counting facts involving quantities or relations of parts, and then note that counting facts have the usual translation which does not commit one to numbers. For example, the sentence 'x's mass/y's mass = 2/3' can be analysed as,

> There are equimassive bodies r, s, t, u, v such that r and s are non-overlapping, t, u and v are non-overlapping, the mereological sum of r and s has the same mass as x, and the mereological sum of t, u and v has the same mass as y.

Field appears to want to further eliminate quantities in favour of relations between objects; but that is not necessary to make the point that numbers disappear in this story, so that one might again claim conservativeness of the real-number-theoretic supertheory.

Of course, there remains what to do about the further (atomic) statements in the nominalist subtheory of the form 'x's mass = 0.6gm'. These would have to be declared literally false, though someone who believes in relations as universals doesn't seem to have a strong a priori objection to someone who believes in monadic quantities as universals. At any rate, this is a familiar dispute, and one where the naturalness of the existing formulation of physics taken at face value weighs in favour of quantities and their ratios. But there is another cost also, namely that all propositions asserting the existence of *real*-number-valued quantities are definitely false. Indeed, one can understand this strategy as making the empirical claim that the world has less structure than might be thought if one took the existing formalism at face value. Perhaps this empirical claim is right, but it would need independent defence. Without that, it is an arbitrary and unwarranted claim. And well-entrenched theory is certainly against it. Thus for example it would be false that the length of the hypotenuse of a right angled triangle of side 1cm is $\sqrt{2}$cm. It would be false that a circle constructed on a diameter of 1cm has a circumference.

It is concluded, then, that there does not seem to be available a good way to avoid the existence of numbers as ratio-comparisons between quantities, given the role of the latter in fundamental laws, especially spacetime theory.[5]

[5] It should not be thought, of course, that all of mathematics has thus been rendered applicable. For example, set theory remains unjustified, and for good reason. One point of agreement with Field, is on the platonist nature of sets, and their dispensability. Field's chosen physicalist sublanguage is mereology, the theory of the part-whole relation. This is clearly the right way to go, undoubtedly some existing things are parts of some other existing things. The difference between sets and wholes boils down two extra principles in set theory: (i) singleton formation, i.e. given any thing, *a*, there exists its singleton set {*a*} which is different from *a*, and (ii) the existence of the null set {}. As many have noted, neither of these principles is attractive from a physicalist perspective: the addition of entities {*a*}, {{*a*}}, would seem to add nothing to the causal powers of the universe; while the special case of the null set, abstract brackets around *nothing*, is particularly otiose. Set theory has the appearance of a failed and over-baroque attempt at a theory of collections, which mereology does in a way which is much more physicalistically acceptable. Needless to say, this has nothing to say about the usefulness of set theory; nor about problems about the possible second-order nature of comprehension axioms for mereology and set theory. It must be acknowledged at this point that Penelope Maddy (1990) has ingeniously defended a physicalist account of sets. If sets can be made out causally, then I am all for them; but Maddy also seems to suggest that both (i) and (ii) above are false, which seems to me to be mereology. At any rate, it is platonist sets which I deny. Of course, number theory can be *constructed* within set theory. But that is just one theory of number-like structures. Who says that numbers are that way, or another? There was no mention of sets in the above story about ratios. If numbers are causal but sets are not, then one should accept the former but not the latter. Conversely, also, there can be emergentist constructions of sets, such as David Lewis's *megethology* which constructs a set-theoretic-like-structure within mereology. However, the present project, like Field's, concerns the status of mathematical items as basic ontological categories.

IX. Conclusion

We can now see a role for numbers. To begin with, there is a fairly obvious computational role for numbers. The conservativeness of mathematics derives, in the end, from crossing off the dimension in the calculation because what matters is the calculation. This shouldn't blind us to the fact that the primary reality is the dimensioned quantity: if there were no quantities there would be nothing for numbers to compare. However it also true that numbers as common ratios across quantity-kinds do not themselves have a dimension. Thus the same computation works for each quantity-kind which is structured by the same ratio comparisons.

It further follows that what numbers do, that multiple congruence relations do not do, is to provide a *cross-dimensional* comparison of relative size. 15cm/10cm = 30gm/20gm = 6sec/4sec = 1.5 . This is a special case of the fact that stable number ratio comparisons provide a platform for objective congruences between *derivative quantities*, such as centimetres per second. This is importantly part of our sense of the stability of nature. If we operationally define a unit in the dimension, say a metre as the same length as the standard metre bar, then surely one wants the ratio between the standard and a sample to be an objective fact of nature, something which might be unchanging over a period of time for example. And if the values of two independent variables from different quantity-kinds bear no comparative relationships to their respective units, what would one make of a law of linear form, for example $y = k.x$? How could the law $y/x = k$, where k is a constant quantity, explain or even describe any constancies in nature, if the relative sizes of y and x bear no objective relation to each other?

To sum up, we have been considering the prospects for the success of Field's program. It was suggested in Part 1 that there is an easier way than Field's to promote conservativeness, but that in any case any conservativeness claim has to be supported by an argument to the effect that the preferred nominalist sublanguage exhausts physical reality. In Part 2, some strategies along these lines were considered, such as metric conventionalism and a general rational-quantity ontology, and it was argued that these are unsatisfactory. As a consequence, there exists a reasonable physicalist/causalist account of number-like entities as relations between independently well-motivated quantity-universals; and we saw their distinctive role in providing cross-dimensional size comparisons.[6]

REFERENCES

1. Adams, Ernest W., 'Measurement Theory', in Asquith and Kyburg (eds.), *Current Research in Philosophy of Science* (Michigan: Philosophy of Science Association, 1979).
2. Azzouni, Jody, 'Thick Epistemic Access: Distinguishing the Mathematical from the Empirical', *The Journal of Philosophy* XCIV No. 9 (1997), pp. 472–484.
3. Bigelow, John, *The Reality of Numbers: a Physicalist's Philosophy of Mathematics* (Oxford: The Clarendon Press, 1988).
4. Bigelow, John and Robert Pargetter, 'Quantities', *Philosophical Studies* 54 (1988), pp. 287–304.
5. Cheyne, Colin and Charles Pigden, 'Pythagorean Powers and a Challenge to Platonism', *Australasian Journal of Philosophy* 74 (1996), pp. 639–645.

6 Thanks to Greg Currie, Ian Hunt, Hartry Field, Graham Nerlich, Jack Smart and two referees of this iournal for useful comments on an earlier draft.

6. Field, Hartry, *Science without Numbers* (Oxford: Blackwell, 1980).
7. Field, Hartry, *Realism, Mathematics and Modality* (Oxford: Basil Blackwell, 1989).
8. Field, Hartry, 'Mathematics without Truth (A Reply to Maddy)', *Pacific Philosophical Quarterly* 71 (1990), pp. 206–222.
9. Forrest, Peter, 'Space Curvature and Repeatable Properties: Mormann's Perspectival Theory', *The Australasian Journal of Philosophy* 74 (1996), pp. 319–323.
10. Forrest, Peter and David Armstrong, 'The Nature of Number', *Philosophical Papers* 16 (1987), pp. 165–186.
11. Grunbaum, Adolph, *Philosophical Problems of Space and Time* (London: Routledge and Kegan Paul, 1964).
12. Hawthorne, James, 'Mathematical Instrumentalism Meets the Conjunction Objection', *Journal of Philosophical Logic* 25 (1996), pp. 363–397.
13. Hellman, Geoffrey, *Mathematics Without Numbers* (Oxford: The Clarendon Press, 1989).
14. Maddy, Penelope, *Realism in Mathematics* (Oxford, The Clarendon Press, 1990).
15. Misner, C., K. Thorne and J. Wheeler, *Gravitation* (San Francisco: W.H. Freeman, 1973).
16. Mortensen, Chris, 'Arguing for Universals', *Revue Internationale de Philosophie* 160 (1987), pp. 97–111.
17. Mundy, Brent, (1987a) 'Faithful Representation, Physical Extensive Measurement Theory and Archimedean Axioms', *Synthese* 70 (1987), pp. 373–399.
18. ——(1987b), 'The Metaphysics of Quantity', *Philosophical Studies* 51 (1987), pp. 29–54.
19. ——(1988), 'Extensive Measurement and Ratio Functions', *Synthese* 75 (1988), pp. 123.
20. Nerlich, Graham, *The Shape of Space* (Cambridge: Cambridge University Press, 1976).
21. ——, 'How Euclidean Geometry Has Mislead Metaphysics', *Journal of Philosophy* 88 (1991), pp. 168–189.
22. Newton, Isaac, *Universal Arithmetick* (2nd ed.) (London: Longman, 1728).

Change

Change is so pervasive in our lives that it almost defeats description and analysis. One can think of it in a very general way as alteration. But alteration in a thing raises subtle problems. One of the most perplexing is the problem of the consistency of change: how can one thing have incompatible properties and yet remain the same thing? Some have held that change is a consistent process, and rendered so by the existence of time. Others have held that the only way to make sense of change is as an inconsistency. This entry surveys the history of this problem and cognate issues, and concludes that the case for change as inconsistency cannot be dismissed so easily.

1. Introduction

The most general conception of change is simply difference or nonidentity. Thus we speak of the change of temperature from place-to-place along a body, or the change in atmospheric pressures from place-to-place as recorded by isobars, or the change of height of the surface of the earth as recorded by a contour map. Contour lines record sameness in quantities (such as 100 metres) from the same quantity-kind (such as height), and the differences recorded by different contour lines are quantity-differences (100 metres as opposed to 200 metres). The philosophical question here is how to construe such statements of identity and nonidentity, and it seems that the problem of universals is the main issue.

A narrower usage of “change” is exemplified by change in the properties of a body over time, that is temporal change. This essay will focus on temporal change. We begin by separating the concept of change from several cognate concepts, specifically cause, time and motion. Then we briefly survey attempts by such thinkers as Parmenides and McTaggart to deny change. There follows an account of the problem of the instant of change, where it is concluded that the problem is too general to admit a single solution, but requires specification of further metaphysical principles envisaged as constraints on a type of solution. The final three sections, the bulk of the essay, consider the question of the consistency or inconsistency of change, which in one way or another looms over all our discussions. It emerges that the case for change as an inconsistent process is stronger than might be expected.

2. Change, Cause, Time, Motion

Our interest in this essay will be on the special case of temporal change. So construed, the notion of change is obviously bound up with notions of cause, time and motion. Now a distinction between change and cause can certainly be drawn. It is clear that uncaused change is conceptually possible, and arguably actual in such things as radioactive decay. Conversely, the operation of a sustaining cause results in no change in a thing, if the thing would otherwise be undergoing a change which the sustaining cause prevents. Hence, the operation of a cause on a thing is neither necessary nor sufficient for change in that thing. Accordingly, we put the topic of cause in the background when discussing change.

Time cannot be so backgrounded. The thesis that time could pass without change in anything at all has proved controversial, and we have adopted the usage that change in a thing implies the passage of time. Aristotle nonetheless argued that change is distinct from time because change occurs at different rates, whereas time does not (Physics IV,10). This essay focusses on the topic of change, while not denying that the topic of time is inseparable from it. Motion, as change in place, will figure prominently in our discussion.

One well-known idea is that of Cambridge change. This can be arrived at by following the well-tried analytical technique of re-casting philosophically important discussions and concepts in the meta-language. Thus a Cambridge change in a thing is a change in the descriptions (truly) borne by the thing. The phrase “Cambridge change” seems to be due to Geach (1969, 71-2), who so named it to mark its employment by great Cambridge philosophers such as Russell and McTaggart. It is apparent that Cambridge change includes all cases ordinarily thought of as change, such as change of colour, from “red” to “non-red.” But it also includes changes in the relational predicates of a thing, such as when I change from having “non-brother” true of me to having “brother” true of me, just when my mother gives birth to a second son. It might seem faintly paradoxical that there need be no (other) changes in me (height, weight, colouring, memories, character, thoughts) in this circumstance, but it is simply a consequence of the above piece of metalinguistic ascent. It does point up, though, that in attempting to capture the object-language concept, one should take note of the distinction between the monadic or internal or intrinsic properties of a thing, and its relations or external or extrinsic features. Thus the natural view of change is that real, metaphysical change in a thing would be change in the monadic or internal or intrinsic properties of the thing. We will return to this point in Section 5.

3. Denying Change

It is on the face of it extremely implausible to deny change, but extreme implausibility has not always deterred philosophers. The Eleatics (C5th BCE), particularly Parmenides, appear to have been the first to do so. Parmenides maintained that whatever one speaks about or thinks about must in some sense exist; if it did not exist then it could not exist, and thus could not even be thought about. From this Meinongian-sounding thesis, it is deduced that the existing thing cannot have come into existence, because to say that it could would be to speak of a time when it did not exist. By similar reasoning, existing things are eternal because they cannot go out of existence. It is now a small step to conclude that change is an illusion, on the grounds that a change in a thing implies that there was a time when the thing-as-changed did not exist. However, this argument is not persuasive: the premiss that what does not exist cannot exist is dubious, as is the premiss that the non-existent cannot be thought or spoken about.

Parmenides' disciples Melissus and Zeno developed this theme. Melissus argued that motion implies empty space to move into, but empty space is a nothing and so cannot exist, so that motion is impossible since it implies a contradiction. This argument requires the dubious premisses (1) that empty space is a nothing (which is denied by realists from Newton to Nerlich), and (2) that motion would have to be change relative to space. Even those who have held that empty space is a nothing (relationists from Leibniz to Mach and onward) have not generally denied motion, proposing instead that motion of a thing is change in the spatial relations between that thing and other things.

Zeno's brilliant paradoxes are generally accounted as attempts to defend Parmenides. We will not look at these in detail, but his paradox of the arrow is relevant to what follows. This is the argument that an arrow in flight could not really be moving because at any given instant it would be at a place identical with itself (and not another place); but something at just one (self-identical) place could not be described as moving. Discussion of this subtle argument is deferred until discussion in a later section of Graham Priest's position, which turns on similar premisses.

McTaggart's well-known argument (1908) that time is unreal applies equally to the unreality of (temporal) change it seems. McTaggart distinguished between two ways of attributing temporal characteristics to events. The *A*-series of events is given by the descriptions “past,” “present” and “future,” while the *B*-series is strictly in terms of the relational concepts “earlier,” “simultaneous” and “later.” Now the *B*-series is insufficient to define change, because *B*-series relations apply unchangingly if they apply at all; whatever is earlier than something is always earlier than it. Moreover, the *B*-series presupposes the *A*-series since if *X* precedes *Y* then there must be a time when *X* is past and *Y* present. This step in the argument is not at all absurd: the discovery of spacetime, the relativistic realisation of the *B*-series, has impelled many from Minkowski on to describe it as a “static” conception of time. A genuinely dynamic conception of change would thus need to have things coming into and going out of existence with the passage of time, whereas spacetime invites quantification over it all “at once” as it were.

Thus according to McTaggart the source of time and change must be found in the *A*-series. But the *A*-series implies a vicious regress. Any event must have all three properties, pastness, presentness and futurity, but this is a contradiction. The only way out of the contradiction is to say that the event is past, present and future at different times; but the same question arises about the temporal instants themselves, which would force us to appeal to a further time series to avoid the contradiction.

Two millennia of philosophical history show in the greater sophistication of McTaggart's argument over those of the Greeks. Whatever we make of it, and much has been written about it, it highlights the baffling nature of the apparent passage of time. On the other hand, any denial of temporal change such as McTaggart's is surely required to explain the overwhelming fact of its apparent existence. There are problems either way. However, one thing can be said about all the above denials of change: they all argue against change on the ground that it implies a contradiction. But the assumption of the consistency of change has been denied by a number of influential figures, as we will see.

4. The Instant of Change

Consider a car moving off from rest at exactly noon. What is its state of motion at the instant of change? If it is in motion, when did it start? And if it is motionless, when could it ever begin? This problem was explored by Medlin (1963), Hamblin (1969), and others. Put this way, a solution for at least some special cases is readily available. Locate the time origin $t = 0$ at noon. If the car's position function f is given by, say, $f(t) = t^2$, then its speed is $2t$. If motion is defined as having non-zero speed, then the car is motionless at $t=0$. On the other hand, at all $t > 0$ it is in motion, so there is surely no puzzle about when it could ever begin: there is no first instant of motion.

However, there are more troublesome special cases. Suppose that the car's position function is given by: $f(t)=0$ for all $t < 0$, else $f(t)=t$. Then speed is zero for all $t < 0$, and speed is 1 for all $t > 0$. But what of $t = 0$? One should avoid “arbitrary” solutions, which privilege one possibility (such as that it is in motion) over another (that it is not), but do not give a reason for so privileging. There is of course at least one simple solution that is non-arbitrary, namely that it is neither in motion nor motionless, since its speed is indeterminate at $t = 0$. This solution derives from the fact that according to classical calculus there is no derivative of such a function at $t = 0$.

But can we do no better? The present author (1985) proposed to set aside the problem until more is said about various possible constraints on the solution. Unless we had some reason to think that such functions really did describe the world, we might well feel that a solution was less than imperative and less than unique. For example, the world might be described wholly with C-infinity functions (n-th derivatives exist for all n, e.g., cos, sin, log, exponential functions). The above function is not among these, since its derivative is discontinuous. But then it isn't clear what we might say of it if the example is counterfactual. There might be different things to say depending on what further principles describe the possible world. Hence we would need to supplement the original statement of the problem with an argument to the effect that we might expect such functions to describe the real world, or alternatively supply additional metaphysical principles to be regarded as constraints on the solution.

A related problem is the fracture problem, described by Medlin. Imagine fracturing a material body such as a piece of wood, regarded as a plenum (full of matter). What is the state of the two new surfaces after the fracture? Unless matter is to be created or destroyed, we seem to have to say that the break is half-open, with one new matter-surface being topologically closed and the other being topologically open. But which surface is which? There seems to be no principle to determine which. In response, it can be asked how seriously we have to take the postulation of a plenum. If for example matter is as Boscovich suggested, punctate and surrounded by fields, then there are no plena, and the problem is no more than hypothetical. Or again, there might be plena but other principles might apply. For example, mass-density functions might drop smoothly to zero at the boundaries between matter and empty space, which would mean that all surfaces were open. On the other hand, it might be instead that as a matter of fact all surfaces are topologically closed. This would need an inconsistent solution (see below, sections 5-7).

5. Consistent and Inconsistent Change

If a changing thing has different and incompatible properties then a contradiction is threatened. The obvious move to make when confronted with the fact that things change, is to say with Kant (1781) that they change in relation to time, which avoids the inconsistency. But then another problem emerges. In what sense can one thing persist through change? Identity across time and space is the mark of universals, but we also account particulars such as billiard balls and persons as having self-identity across time.

Aristotle's views on the persistence of things are worth noting here. At the risk of gross oversimplification of what is treated thoroughly elsewhere in this Encyclopedia (see Cohen (2001)), it can be said that early on he took the the view that what persists over time and through change, the *substrate*, can be identified with *matter*, and that it is the *form* of matter which is acquired or lost. (Physics I, 5-7). By the Categories, it is *substance* which is said to be susceptible of contrary attributions; and as such substance itself has no opposites. (Categories 4a10). In the Metaphysics Z, a more complex doctrine of substance, that which *is*, is worked out. Substance is not the substrate, matter, since that lacks particularity. *Its* substance, what it is to be *that thing*, that without which it does not exist, is its *essence*. Aristotle then links essence with his theory of causes, being identified variously with its final cause and with its formal cause.

Although Aristotle's views about change -- in particular, his distinction between essence and accident

-- have sometimes been thought to contain a solution to the problem of persistent identity through change, it seems to this author that they do not really get a grip on the problem in its most fundamental form. This is perhaps clearest in the Categories, where the ability of substance to admit incompatible accidental features is more-or-less definitional.

The problem can be made sharper by reflection upon the law of the indiscernability of identicals. If a thing-at-t_1 were identical with a thing-at-t_2, then they should share all their properties. What sort of identity is it, if not that? But if the properties at different times are incompatible, then a contradiction follows. Because they emphatically took the view that contradictions are never true, the great Buddhist logicians Dharmakirti (C7th CE) and his commentator Dharmottara (C8-9th CE), who had certainly read their Aristotle, deduced that identity over time does not exist (see Scherbatsky (1930) vol 2). This is the Buddhist doctrine of moments, essentially an ontology of instantaneous temporal slices. The doctrine of the momentariness of existence is felicitously in accord with the core Buddhist doctrine of the impermanence of all things. The doctrine of moments might seem to be an unnecessarily strong application of impermanence, certainly unnecessary for soteriological purposes, were it not for the evident strength of the argument in its favour, not to mention its accord with the spacetime ontology of modern physics. On the other hand, it is of course psychologically very difficult to believe that one's own self, as something genuinely self-identical, has not endured from moment-to-moment in the past. Even so, the thesis of the momentariness of human existence has had a recent defender in Derek Parfit (1984), who asks what sort of principle could unify the temporal stages sufficiently closely to be worth calling identity. He argues that none could, and proposes that internalising the momentariness of our lives has a beneficial effect on how we should face our deaths.

This theme is echoed in a recent debate on the topic of ‘temporary intrinsics’, which also connects with the earlier-mentioned concept of Cambridge change. Cambridge change in a thing is still change in *something or other*, but it is not always change in *the thing itself.* Thus we might seek to isolate change in the thing itself by change in its *intrinsic* properties. But then we have the problem of in what sense it continues to be just one thing through a change in its intrinsic properties. Now obviously this raises the question of how to define the concept of intrinsicality. We do not address that here, since it is discussed elsewhere in this *Encyclopedia*, see Weatherson (2002). So assuming a *prima facie* distinction between the intrinsic and extrinsic properties of a thing, how does a thing persist through changes in its intrinsic properties? David Lewis and others debated this question, e.g., Lewis (1986),(1988). Several options for a solution were canvassed, three of which were as follows.

(1) The basic existents are things indexed by times, that is time-slices. What primarily exist are things-at-a-time: “*a* is red at *t*” is rendered “*a*-at-*t* is red”. Things that persist over time are then wholes made up of such parts, and one says that persisting things *perdure* rather than *endure*. This is the solution favoured by Lewis, by the present author, and by space-time theory.

(2) A second option is to say that, instead of indexing times, one indexes properties: “*a* is red at *t*” is rendered as “*a* is red-at-*t*”. This option does not seem to have had any defenders, perhaps because those properties which are universals are supposed to be wholly in each of their instances, which the indexing apparently denies.

(3) A third option takes as its basic minimal idea that the index modifies the whole event: (*a*'s being red) holds at *t*. A variant is to take the index as modifying the exemplification ‘relation’: *a* exemplifies-at-*t* redness. Versions of this position were urged by several contributors: Johnston (1987), Lowe (1987), (1988), Haslanger (1989). However, the problem for adverbial-style analyses anywhere is to provide enough semantics, enough logical structure for the event, to account for the logical implications of the sentences under analysis, as Davidson (1967) pointed out. So for example one has things like: $(((Fa)\text{ at }t)\ \&\ a=b)$ implies $((Fb)\text{ at }t)$; or $(((Fa)\text{ at }t_1)\ \&\ ((Ga)\text{ at }t_2)\ \&\ (F$ is incompatible with $G))$ implies not $t_1=t_2$; or $(((Fa)\text{ at }t)\ \&\ ((Gb)\text{ at }t)\ \&\ (F$ is incompatible with $G))$

implies not $a=b$. One thus cannot rest with a minimalist position. At least Lewis' has the merit of providing a viable semantics, a direct parallel with counterpart theory in modal semantics. Of course, the basic ontology of Lewis' favoured position was Dharmakirti's though Lewis did not note that fact. More to the point, Dharmakirti's strategy did not depend on the intrinsic/extrinsic distinction. The problem of contradictory attributions occurs even if the attributions are extrinsic, and Dharmakirti's argument is a straightforward application of Leibniz' law to things-at-a-time. If time-slices are admitted at all, and it is hard not to do so if they are sanctioned by relativity theory, then Dharmakirti's argument goes through.

Others have taken a different course on the issue of the consistency of change. Herakleitos (C6th BCE) wrote in a suggestive fashion, with his doctrine of the unity of opposites. However, his few surviving sentences are too obscure and fragmentary to give much confidence in interpretation. He spoke of the same river having different waters at different times, but there is no development of the observation. Similarly he spoke of the sea as being at one time both life-preserving (to fish) and death-dealing (to humans), and “the path up and the path down are one and the same.” These examples hardly force one to believe in true contradictions, however.

There is also in Herakleitos the idea that everything is in a state of flux, always changing, and that it is the struggle between opposites (opposed tendencies) which drives change. This can be seen as an early version of the Marxist dynamic of dialectical materialism. But without a separate argument for the inconsistency of change, there is no reason to think that it remains anything but a formally consistent theory.

Hegel was more explicit. In *The Science of Logic* he said that only insofar as something has contradiction in itself does it move, have impulse or activity. Indeed, movement is existing contradiction itself. “Something moves not because at one moment of time it is here and at another there, but because at one and the same moment it is here and not here.” (Hegel (1812) p. 440).

There is something appealing in this argument. As Priest and Routley put it, “in change… there is at each stage a moment when the changing item is both in a given state, because it has just reached that state, but also not in that state, because it is not stationary but moving through and beyond that state” (Priest, Routley and Norman, 1989, p. 7). Think of a body coming to rest at a given time, and compare it with the same body proceeding on to further motion. There must be something about the body at that instant which distinguishes the two scenarios, or there could be nothing at the time to count as continuing change. Cause cannot do it, for a body can continue in its state of motion without being impressed by an external force, as Newton taught us. Nor can mere velocity do it, since velocity is a relation to surrounding points. Indeed, there is no difference in velocity between a body momentarily at rest, and a body at rest for a period around the instant; yet one is changing and the other not.

We will look more closely at this argument in the next section. However, here we can remind ourselves of Hegel's idealism. Just about everyone agrees that contradictions within ideas are easier to swallow than contradictions in the external world. In the special case of the phenomenology of motion, it is not such an absurd speculation that what distinguishes the direct perception of motion from the mere static memory of difference in position, is that nearby small variations in the stimulus are read into a kind of buffer where they are not compared as static memory does so much as overlapped or superimposed in the way that contradictions are. After all, we are not at all good at discriminating small intervals of time, as the success of 25 frames per second makes apparent. Thus, the mind constructs a kind of contradictory theory which undergoes constant update. Indeed, this may well be the source of the troublesome intuition we noted earlier, that it is one and the same thing which endures through change, even though it is acknowledged that it has different properties at different (nearby) times. If this is right, then if one thinks with Hegel that the world is a kind of idea, then the contradictoriness of ideas such as motion is apt to spill over to the contradictoriness of their realisations in the world. Indeed, even without the assumption of full-blown idealism, there is always the caution that if a theory (consistent or not) can be made out which describes an epistemic state, i.e., a cognitive state, then how can we be entirely confident that the world simply *could not* be that way?

Taking a far less ambitious view than Hegel, Von Wright (1968) nonetheless proposed an interesting account of conditions in which change would have to be regarded as inconsistent. The account requires two conditions. The first condition is that time is regarded as structured as nested intervals rather than an assemblage of atomic point-instants. This is an attractive proposal, if only because no-one has ever seen a temporal or spatial point. Of course, standard relativity theory proposes that spacetime is punctate, as does the usual mathematics of the continuum. But a successful non-punctate mathematics using intervals instead can be worked out, albeit with considerably extra complexity. (see e.g., Weyl 1960). Now in the ontology of intervals, since there are no atomic points to attach a unique proposition to, the most one can say is that a proposition holds somewhere in the interval, with the limiting case that it holds throughout the interval.

Von Wright's second condition was then to suppose that an interval might be so structured that a given proposition p and its negation $\neg p$ are dense in each other throughout the interval. This means that no subinterval, no matter how small, can be found in which just p holds throughout that subinterval, and no subinterval can be found in which just $\neg p$ holds throughout the subinterval: every subinterval in which one holds, the other holds as well. From an external point of view admitting instants, we can see that this is a genuine consistent possibility, if for example we think of p as the proposition that a rational number of seconds has passed, and $\neg p$ as the proposition that an irrational number of seconds has passed. These are dense in each other on the classical real line regarded as time. Thus, there is no subinterval which is purely p throughout and no subinterval which is purely $\neg p$ throughout.

This was von Wright's proposed account of a continuous change in an ontology of intervals. The state $\neg p$ changes continuously to p if there is a preceding interval which is $\neg p$ throughout, then an interval with $\neg p$ and p dense in each other, then a succeeding interval with p holding throughout. Von Wright described this as a kind of inconsistency. Unfortunately it is not clear from his written words whether he had in mind that the situation was inconsistent or only possibly inconsistent. His argument seems to be this. In an ontology of intervals we begin with descriptions like “It rained here yesterday” which means that it rained sometime here yesterday. The basic description is thus “p holds (somewhere) in the interval I.” The special case where p holds throughout I is noted, where to hold throughout is for there to be no subinterval in which $\neg p$ holds. Now p's holding in I is of course compatible with $\neg p$'s holding in I. But there is no contradiction here, as long as there is a partition of I into subintervals such that p holds throughout the subinterval or $\neg p$ holds throughout the subinterval. Thus if we take in that a disjunction holds in an interval just in case there is a partition in which each of the disjuncts holds throughout its subintervals, we can say that if there is such a partition for p, then the law of excluded middle $p \vee \neg p$ holds throughout the interval. Von Wright introduced the modal operator Np for “Necessarily p.” If we define “Np holds in I” to mean that p holds throughout I, we can say that if there is no continuous change in the above sense, then Excluded Middle LEM holds necessarily, $N(p \vee \neg p)$. However, defining the modal “Possibly” in the usual way as $M =df \neg N \neg$ and assuming de Morgan's Laws, Double Negation and Commutativity, we get the result that in an interval in which there is continuous change, $M(p \,\&\, \neg p)$ holds, ie. a contradiction is possible. Presumably it further follows that in a subinterval which has continuous change throughout, $N(p \,\&\, \neg p)$ holds. Needless to say this implies that a contradiction is true in that subinterval. We might note that the result that continuous change is a true contradiction follows without the detour through modal logic, since if LEM is false then $\neg(p \vee \neg p)$ holds for some p, and so by de Morgan and Double Negation, $p \,\&\, \neg p$ holds (throughout).

This ingenious construction has its problems. It is certainly dangerous to assume De Morgan's Laws and Double Negation when the logic of intervals is the case in point. They both fail for open set logic, which is to say intuitionism, just as they both fail for its topological dual, closed set logic. On the other hand, what is one to say if the world is structured as intervals, non-punctate, and if there are subintervals in which propositions and their negations are dense in each other, interspersed with intervals where one of the propositions holds throughout? The latter are clearly periods of non-change, and the former are reasonably described as intervals of change. And yet it would seem that the best one can do is to say that $p \& \neg p$ holds in the transition periods: there appears to be no consistent way of describing what is happening in the situation which adheres to intervals and eschews points.

6. Inconsistent Motion

Many of the above themes come together in Graham Priest's inconsistent account of motion in In Contradiction (1987). Priest sets up the opposing consistent account of change as what he calls the cinematic view of change. This is the view that an object in motion does no more than simply occupy different points of space at different times, like a succession of stills in a film only continuously connected. He attributes the view to Russell and Hume. It is an extrinsic view of change, in the sense that change is seen as a matter of a relation to states at nearby instants of time. The best-worked-out version of this view is the usual mathematical description of change of position by a suitable function of time; and then motion as velocity, that is rate of change of position, is given by the first derivative, which is a relation to nearby intervals.

Priest wishes instead to have an intrinsic account of change, in which it is a matter of the features of the object solely at the instant whether it is changing at the instant. He offers three arguments against the extrinsic account. First there is the “abutment” argument (p. 203). Taking the usual view of time as a continuously distributed collection of point-instants, in any change there must be an interval throughout which p holds abutting an interval throughout which $\neg p$ holds. It makes no difference whether there is a last instant for p and no first instant for $\neg p$, or no last instant for p and a first instant for $\neg p$; either way there is no room for a time at which the system is changing. For example, if we said that the change was at the boundary point, then there would be nothing about that point to distinguish it from the situation where there was no change at all because the abutting intervals had the same proposition holding throughout each. Hence there is no change at all in the cinematic view: for change there would have to be a time when change was occurring, and that is absent in this case.

Priest's second argument (p. 217) appeals to causation. It is at least imaginable that the universe is “Laplacean,” by which he means that the state at any time is determined by the states at prior times. But if change is cinematic, then there is no sense to saying that the instantaneous state of the world at the prior time determines its state at subsequent times: for example, not even velocity is determined by the intrinsic instantaneous state of a body. Now a Laplacean universe is possible, but the cinematic view makes Laplacean change *a priori* false.

Priest's third argument (p. 218) is his version of Zeno's arrow argument mentioned earlier. In the cinematic view of change, there is nothing about the arrow at any instant to contribute to its motion: it is indistinguishable from an arrow at rest. But then there is nothing to constitute its motion: an infinite number of zero motions does not add up to anything but zero motion. In response to the reply that according to measure theory a (nondenumerably) infinite number of points of measure zero can have a non-zero measure, Priest argues that this is just mathematics: “…it does not ease the discomfort … when one tries to understand how the arrow actually achieves its motion. At any point in its motion it advances not at all. Yet in some apparently magical way, in a collection of these it advances. Now a sum of nothings, even infinitely many nothings, is nothing. So how does it do it?” (pp. 218-9)

Setting aside questions about the strength of these arguments for the present, how then are we to give an acceptable intrinsic account of motion? According to Priest, the only acceptable answer is Hegel's: that motion is inconsistent. Support comes from Leibniz' Continuity Condition (LCC). This is essentially the thesis, suitably qualified, that whatever holds up to a limit, holds at the limit. Priest's argument for the LCC appeals to causality. He describes change violating the LCC as “capricious” (p. 210). Humeans might be able to accept it, but for them there are no connections, nothing to constitute past states' determination of future states. He also argues if the LCC fails, change would occur, but “at no time” (p. 210): for a proposition switching values discontinuously at a boundary there would be no instant identifiable by its intrinsic properties alone as the one at which the change occurred.

Priest's qualification to the LCC is that it applies only to atomic sentences and their negations: otherwise we would have to admit the case where a disjunction $p \vee q$ held right up to a limit in virtue of p holding at the rational points and q holding at the irrational points: this would be capricious behaviour in which we can make no sense of the past determining the future. We would also admit problems if we allowed the LCC to apply to tense operators: Future-p can obviously hold up to a limit without holding at the limit.

But now we observe that the LCC so qualified implies that continuous change is contradictory. For consider any particle with equation of motion $x = f(t)$. Then at $t = a$ its position $x = f(a)$. However if it is in motion then in the neighbourhood we have $\neg(x = f(a))$, so by the LCC at the limit also $\neg(x = f(a))$, along with of course $x = f(a)$ as well. Priest amplifies this account by proposing that no moving body can be consistently localised. Rather, in moving at time t it inconsistently occupies a small finite (Planck length) lozenge of space, which is made up of the positions it takes in the corresponding lozenge of time surrounding t. This gives a natural intrinsic account of motionlessness at t, namely that there is no contradiction in its position at t. One can propose an account of velocity, as varying with the length of the lozenge or spread of position in the direction of motion. There are applications in Quantum Theory, too. The Heisenberg uncertainty of position may simply be the size of the spread or smeared position. Moreover, there is a possibility for backward causation implicit in the advanced wave front of inconsistency affecting earlier states in the inconsistently identified smear of spatial positions; and backward causation may be the way to go with quantum nonlocality, as Huw Price (1996) has argued.

One quick objection does not succeed. One might argue that since motion and rest are not relativistically invariant, neither could the contradictoriness in motion be part of the absolute character of reality. This may be so, but it does not prevent the concept being of use in the analysis of phenomena by means of frames: frame-relative inconsistencies would still be a (relational) part of the world. More importantly, the concept may find its natural home in QM rather than GR. It is well-known that there are deep incompatibilities between them as they now stand, but the jury is still out on how to resolve them, and it may well be that absolute motion is a part of the solution.

In asking how strong are the arguments in favour of this well-crafted position, we return to Priest's three arguments against the rival, consistent, extrinsic, cinematic view. We recall the first argument was the “abutment” argument: consistent change cannot allow that there is a (single) time at which the change takes place. This will not sway the opposition, who will reply that it is the nature of change, even change at a point, that it is relational in that it requires comparison with nearby points; hence the demand for an intrinsic conception of change is a mistake.

The second argument was that the cinematic view is incompatible with the Laplacean view that the past determines the present. The way Priest puts it is not so plausible: he says that Laplaceanism is possible, whereas the cinematic view rules it out “*a priori*” (p. 217). But this is a modal fallacy: the cinematic view is only ruled out when one adopts the Laplacean view, and so that is only relatively apriori.

The third argument, Zeno's arrow, has greater force though. How can any number, even an infinite number, of zeros add up to a nonzero? The mathematics of measure theory may say that intervals have a non-zero measure whereas individual points are zero, but so what? What is needed is a story which makes its application intelligible and non-arbitrary. If this is not forthcoming, there is the strong counter-intuition that zero marks the absence of existence; and no number of absent or non-existent things or quantities makes a present, existent thing or quantity.

So Zeno's argument after all seems to be the most resilient. But the Laplacean universe also has appeal. Many philosophers have felt uneasy about Hume's views on causation: if the past does not determine the future then the universe is indeed capricious. Still, we can note at least one consequence of Priest's position which many will find implausible. It can be debated whether if a body is changing its position, i.e., if it is in motion, then it has a non-zero speed (consider the case of having non-zero acceleration but zero speed). Most philosophers would however agree with the converse: if a body has non-zero speed then it is in motion. It must be noted, though, that Priest is denying that a non-zero speed is sufficient for motion: the point of the cinematic analogy is to say that a sequence of stills of changing positions does not amount to motion, yet it is precisely this sequence of stills which is used to define the concept of speed. Against Laplaceanism, it will be argued that the concept of speed is present in everyone's story, and the counter-thesis that non-zero speed is sufficient for motion is simpler in that it does not involve appeal to a further, mysterious, intrinsic feature. On the other hand, this has to be weighed against the force of the above counter-arguments, especially Zeno's arrow.

7. Discontinuous Change and the Leibniz Continuity Condition

If the LCC is to have a chance of being applicable, then it needs further restriction, beyond atomic sentences and their negations. This is because it has implausible consequences when applied to certain atomic sentences. Consider any increasing function $f(t)$. Then sentences of the form $f(t) < f(a)$ will hold for $t < a$. By the LCC then, $f(a) < f(a)$. This is surely a gratuitous conclusion even before the contradicting sentence $\text{-}f(a) < f(a)$ is taken into account. The present author (1997) therefore proposed to restrict the application to the atomic sentences of equational theories, that is to sentences of the form $f(t) = 0$. This is not so unreasonable on independent grounds, since the basic laws of nature are expressed in equational form.

So restricted, we can note that far from being unreasonable, it turns out that the LCC is satisfied in a large class of reasonable models, specifically the C-infinity worlds mentioned earlier, in which every function is continuous. These include all those of GR. Now a C-infinity world gives us a kind of half-way house for cause. It might be that all correlations are coincidences, but at least if functions are continuous then causation is a distinctive correlation in that it is transmitted locally. This can be applied beneficially to produce not a general account of inconsistent change, but a particular account of certain inconsistent changes, as follows.

Quantum measurement has long been problematic, for more than one reason. One reason has been that it represents an irreducibly different kind of process from Schrodinger evolution. Another is that it is change which is discontinuous and yet causal: one can make things to happen with measurement, even though one cannot determine the exact outcome. A third reason is nonlocality itself: the nonlocal is *ipso facto* the discontinuous, and yet the nonlocal is governed by a kind of statistical causality. But now, to settle at least some of these issues, it has been proposed to utilise the theory of inconsistent continuous functions. These arise when a function is classically discontinuous, but we inconsistently identify the limit of the function (assuming it has a limit) with its value at the limit. Such functions, by virtue of being continuous, can be shown to satisfy the LCC. But granted that the formal details exist, what reason is there to apply them? It is precisely that we want to preserve a degree of causality, that is LCC-causality, while yet retaining the essential discontinuity and unpredictability of the process. Thus the slogan “nonlocality is inconsistent locality,” which is intended not to apply to change in general but to discontinuous change which we nonetheless have reason to think of as causal.

8. Conclusion

There remain many loose ends from our discussion. Still, it emerges that the connection between change and inconsistency is deep, and that the case for inconsistencies in motion and other change is surprisingly robust.

Bibliography

- Cohen, S.Marc, 2001, Aristotle: Metaphysics, Stanford Encyclopedia of Philosophy.
- Dainton, Barry, 2001, Time and Space, Chesham: Acumen.
- Davidson, Donald, 1967, “The Logical Form of Action Sentences ”, in N.Rescher (ed.) The Logic of Decision and Action, U. of Pittsburgh Press.
- Dharmakirti, 1930, A System of Logic (with Commentary by Dharmottara) in F. Th. Scherbatsky Buddhist Logic, New York: Dover ed. 1962.
- Geach, P.T., 1969, God and the Soul, London: Routledge and Kegan Paul.
- Haslanger, Sally, 1989, “ Endurance and Temporary Intrinsics ”, Analysis 49: 119-125.
- Herakleitos, Fragments, 1987, tr. T.H.Robinson, Toronto: University of Toronto Press.
- Hamblin, Charles, 1969, “Starting and Stopping,” The Monist 53: 410-425.
- Hegel, G., 1812, Wissenschaft der Logik, see A.Miller (tr) Hegel's Science of Logic, London: Allen and Unwin, 1969.
- Johnston, Mark, 1987, “Is There a Problem About Persistence? ”, Proceedings of the Aristotelian Society (Supp): 107-35.
- Kant, Immanuel, Critique of Pure Reason (The Transendental Aesthetic, Section 5), 1781, tr. N. Kemp Smith, London: McMillan, 1933.
- Lewis, David, 1986, On the Plurality of Worlds, Oxford, Blackwell.
- Lewis, David, 1988, “Rearrangement of Particles: Reply to Lowe ”, Analysis 48:65-72.
- Lowe, E.J., 1987, “Lewis on Perdurance versus Endurance ”, Analysis 47: 152-154.
- Lowe, E.J., 1988, “The Problems of Intrinsic Change: Rejoinder to Lewis”, Analysis 48:72-77.
- McTaggart, J.E., 1908, “The Unreality of Time,” Mind 17: 457-74.
- Medlin, Brian, 1963, “The Origin of Motion,” Mind 72: 155-175.
- Mellor, Hugh, 1981, Real Time, Cambridge: Cambridge University Press.
- Mortensen, Chris, 1985, “The Limits of Change,” Australasian Journal of Philosophy 63: 1-10.
- Mortensen, Chris, 1997, “The Leibniz Continuity Condition, Inconsistency and Quantum Dynamics,” The Journal of Philosophical Logic 26: 377-389.
- Nerlich, Graham, 1976, The Shape of Space, Cambridge: Cambridge University Press.
- Parfit, Derek, 1984, Reasons and Persons, Oxford: The Clarendon Press.
- Price, Huw, 1996, Time's Arrow and Archimedes' Point, Oxford: Oxford University Press.
- Priest, Graham, 1987, In Contradiction, Dordrecht: Nijhoff.
- Priest,G, R.Routley and J.Norman (eds), 1989, Paraconsistent Logic, Munich: Philosophia Verlag.
- Von Wright, G.H., 1968, Time, Change and Contradiction [1968], Cambridge: Cambridge University Press .
- Weatherson, Brian, 2002, “Intrinsic vs. Extrinsic Properties”, Stanford Encyclopedia of Philosophy.
- Weyl, H., 1960, Das Kontinuum und Andere Monographien, New York: Chelsea.

Other Internet Resources

[Please contact the author with suggestions]

Related Entries

Aristotle: metaphysics | Hegel, Georg Wilhelm Friedrich | Heraclitus | intrinsic vs. extrinsic properties | logic: paraconsistent | Mach, Ernst | mathematics: inconsistent | McTaggart, John M. E. | Newton, Isaac: and Newtonianism | Parmenides | time | Zeno's paradoxes

Chris Mortensen
Chris.Mortensen@adelaide.edu.au

A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z

Stanford Encyclopedia of Philosophy

CHRIS MORTENSEN and JULIANNA CSAVAS

IN THE BEGINNING

ABSTRACT. In this paper, a survey is made of some of the contributions to the interpretation of Hartle and Hawking's theory of the wave function of the universe and its beginning. It is argued that there are considerable difficulties with the interpretation of the theory, but that there is at least one interpretation hitherto not found in the literature which survives existing philosophical objections.

1. INTRODUCTION

It is widely held that Hartle and Hawking (1983) described a revolutionary account of how time might have come to begin in a smooth way, not in a singularity as the current theories of the Big Bang imply. In order to make the singularity vanish, their proposal was (in part) to transform time t in the very early universe into imaginary time, that is $\mathbf{i}t$, where $\mathbf{i}$ is the square root of minus 1. This has given rise to a large literature about how to understand their suggestion. In the present paper we will survey a subset of the literature, focussing mostly but not exclusively on contributions by philosophers. This is because the philosophers' contributions have largely been the clearest and most rigorous, as we will see, and therefore most worth addressing. Nonetheless we aim to show how inconclusive the discussion has been on all sides. We will conclude with an alternative suggestion of our own, and argue that it escapes the major problems of existing interpretations.

2. THE HARTLE-HAWKING THEORY

One of the more accessible sources is Hawking's own account in *A Brief History of Time* (1988, pp. 145–148). The singularity at the origin of the universe according to classical GR arises because at $t = 0$ the universe has infinite curvature. The universe can thus be represented as the surface of a cone with its sharp point being the first instant. Since in GR curvature is related to mass density by the equation Curvature $= 8\pi$.Density, this implies infinite density at the point also. Hawking describes the situation

Erkenntnis **59:** 141–156, 2003.

as undesirable due to our inability to make a prediction when quantities assume infinite values.

We note a quick digression: there is of course no problem over the mathematics of the infinite. There are well-known rigorous theories of infinite cardinals and ordinals. The theory of the infinite most suitable for Hawking's purposes would be that of Robinson's *Non-Standard Analysis* (1966), since that theory has all the theorems of classical standard analysis, and is even more simple computationally. There is no question that this developed theory permits easy manipulation of differential equations applying to infinite quantities and their reciprocals, the infinitesimals. However, it seems that the infinity at the point of the origin of time remains "absolute infinity", the reciprocal of zero. Unfortunately, despite much effort, there is no theory of the reciprocal of zero, consistent or inconsistent, which is even remotely functional enough to permit calculations. Hence, while infinite numbers are highly effective in the right context, the kind of infinity in question in the present discussion is definitely very undesirable. (Even so, it must be allowed that absolutely infinite curvature is possible in some sense, precisely because a consistent geometrical object such as a cone can exist.)

Returning to our main theme: to avoid an impasse over predictability, Hawking proposes that the absolute probability of the universe's being in a given state is the sum over all histories (earlier spacetimes) with that state as outcome. To make such histories well-defined (converge), it is necessary to sum over histories with a locally Euclidean spacetime metric. This is obtained by transforming the time variable t by multiplication by **i**. A simplified example, adapted from Hartle and Hawking (1983), is a state function ϕ given by $\phi = \int e^{it} dt$. If the integral is taken over the whole of the positive real line, this is not defined since it is the integral of a periodically oscillating function. Under the transformation which maps t to $\mathbf{i}t$ however, it becomes $\int e^{-t} dt$, which is 1 over the positive real line.

The transformation from t to $\mathbf{i}t$ evidently also changes the form of the metric, from:

$$ds^2 = dx^2 + dy^2 + dz^2 - dt^2, \text{ to}$$
$$ds^2 = dx^2 + dy^2 + dz^2 + dt'^2,$$

where $t' = \mathbf{i}t$.

That is, the transformation changes the usual (+, +, +, −) signature Minkowski metric to a (+. +, +, +) signature Euclidean metric. Now while this transformation ensures that certain important integrals exist, it does not by itself transform away the singularity, since multiplying infinity by **i** is still infinite. There is thus required a further move, and so Hawking then *postulates* that the universe did not begin with a singularity. Such non-singular Euclidean metrics take the form of the surface of a closed sphere in 4 dimensions, a 4-sphere. Hence, the universe is explained as arising smoothly without singularity from a very early (half-spherical) Euclidean region which undergoes a change to the more familiar Minkowski metric for spacetime, and then expands in the accustomed inflationary manner as in the standard model of the big bang.

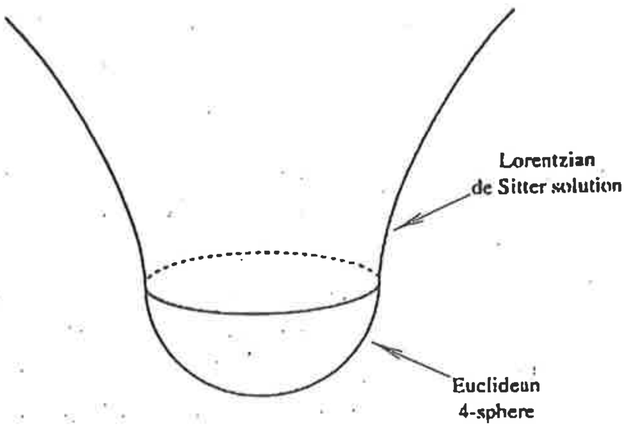
This is unquestionably a very interesting idea. But we see that there is a veritable thicket of complexities over interpretation.

3. NO BOUNDARY IS NO BIG DEAL

Hawking describes his theory as the No Boundary Condition: "the boundary condition of the universe is that it has no boundary". What does this mean? After all, when we compare Hawking's smooth-bottomed bucket (half-sphere) with the pointy-bottomed cone, we see that they are *homeomorphic*, that is there is no *topological* difference between them. But it is obvious that the surface of a sphere has no boundary. It follows then that the surface of the cone likewise has no boundary. Why then does Hawking describe the half-sphere as being preferential to the cone in this respect?

The answer must obviously lie in the *difference* between a cone and a half-sphere. The difference of course is that at its point a cone has infinite curvature, whereas the half-sphere nowhere has infinite curvature. But surely this makes no difference to the existence of a boundary? Curvature is an affine concept, and strictly less general than a topological concept like a boundary. The answer is in the (simple linear) relation between curvature and mass-density. As we saw, if curvature is infinite, so must be density. But it is (absolutely) infinite mass-density which is the real singularity here, not the point of the cone itself which is only the geometric "indicator" as it were. So far so good, but why does infinite mass-density make for a boundary? The reason can only be that infinite mass-energy is never realised in the universe. Thus, if the universe were like a cone, then it would be a cone with the end point nipped off, a *nonexistent* vertex. Now a cone with its end-point removed is homeomorphic to a half-sphere with a point removed. In fact it is homeomorphic to a half-sphere with a disc removed, and that certainly has a boundary, namely the circle at the edge of the disc. Moving back to the cone then, removing the point creates a boundary, the "absent" point (which can conveniently be thought of as existing in a space containing the cone). That is, to say that the universe has no boundary is to deny that the universe has a hole in it at the point of creation. But this in turn is simply to deny that a singularity of mass-density ever existed. *Thus, to say that the universe has no boundary is simply to refuse to countenance singularities in the distribution of matter. It is not to solve the problem, it is simply to declare the hope that a reasonable constraint on a solution will be satisfied.*

4. THE EUCLIDEAN REGION: SPACE OR SPACETIME?



The above picture, taken from Hawking and Penrose (1996, p. 86), shows two spacetime regions, a Euclidean region like a smooth half-bucket, joined to a Minkowski region like a cone with a small part of its pointy end nipped off. It should be noted that space and time in the picture are confined to the surface, not the interior. Now there is an important distinction to be noticed. Just about all contributors to the debate speak of the Euclidean region as one where the distinction between time and space disappears, where time becomes "spatialised". The motivation for this way of talking is clear enough: the shift from a $(+, +, +, -)$ metric to a $(+, +, +, +)$ metric removes a distinctive *metrical* role for time. Furthermore, this has significant further implications. For example, in Minkowski spacetime the difference between spacelike, timelike and null separation of events derives from the sign of the spacetime distance between them, so that with the shift in the metrics the whole light-cone structure is removed.

This then raises the question of the ontological status of *time in the Euclidean region*. We find in the literature two discernibly different ways of talking. On the one hand one can take the above talk of disappearance literally at face value, as the view that the Euclidean region simply has *no time at all*, just 4 *spatial* dimensions. This might be described as an *eliminationist* interpretation of time in the Euclidean region. However, equally ubiquitous is the continued use of temporal terminology to describe the

Euclidean region, as we see below. To explain that, one might understand the eliminationist talk as a slightly loose way of speaking of a phenomenon which is *still time*, just with *somewhat different properties*. This might be described as a *reductionist* interpretation. Perhaps the clearest statement of the latter interpretation is where Davies says in *About Time* (1995, pp. 191–192): "Time is always time, it doesn't really turn into anything . . . what we call time may once have had some of the properties we normally associate with space".

Various of the difficulties we discuss below apply to both interpretations. However, the eliminationist interpretation has further problems of its own, we think. Chief among these is the question of in what sense we have an explanation *at all* of the Minkowski epoch of the universe, if there is no causal process leading up to it. Having said that, we signal that in a late section we will take this up again and endeavour to make sense of it. For the time being, we propose to adopt the view that the Euclidean region possesses time in some sense. There remain problems aplenty to discuss. In defence of this interpretation, it can be said that the loose talk of "no difference" may well just mean, in the context of metrics, no more than "no *metrical* difference". From the fact that there is no metrical difference between time and space, it does not follow that there is no difference at all. After all, there is no metrical difference between the signatures of the 3 spatial dimensions either, but it is unquestionable that spacetime has 3 spatial dimensions not 1. The disappearence of the Minkowski causal structure is inevitable for Euclidean spacetime, but then Minkowski spacetime is not a necessary truth. Naturally causality has to be carried by a different physical structure, but that is to be expected, not feared. Misner, Thorne and Wheeler in Gravitation (1973, 51) condemn the old practice of using Euclidean coordinates in their section "Farewell to '*ict*' "; but in the context it is clear that they are objecting to the use of such coordinates to describe universes which are still fundamentally Minkowski. That is not so here. It is not as if one were simply redescribing a universe which is in reality Minkowski: rather there would be a fundamentally different set of laws of nature.

5. THE JOIN PROBLEM: A CONTRADICTION?

In short, we proceed for the present with the interpretation that the Euclidean region has time of some sort in it. So the theory involves the postulation of two regions, one Euclidean and one Minkowskian. This leads one to inquire in more detail about the relation between the two regions. Hawking draws them as joined together. But what is the nature of the relation? This is called *the join problem* by Deltete and Guy (1996). For example, in describing the theory, Hawking speaks of the Minkowski phase as being "later than" the Euclidean phase (pp. 145–146). Again, Paul Davies in *About Time* (1995) stresses that the join is not a sharp discontinuous change but is smeared out, then speaks of a continuous sequence where time "starts out" as space and gradually "turns into" time as we know it. One can see here two variants of the join problem. One is the problem of how the Euclidean region can be earlier than the Minkowski region; the other is the problem of what the nature of join, in particular is it abrupt or fuzzy? We briefly consider the first of these here, then turn to the second in the next section.

In (1993), Quentin Smith seizes on Hawking's talk about how "once" quantum smearing effects subside, 4-D Euclidean space joins onto Minkowski spacetime. If in the Euclidean region there was no real time, how could it stand in the relation of being earlier than? " ... it is not really earlier than, later than or simultaneous with the four-dimensional spacetime manifold. Accordingly it is false that the 4-sphere joins ... *once* .... Nor can this "once" refer to imaginary time, which would imply that the real spacetime is later in imaginary time than the 4-sphere, which it is not" (p. 318). The idea here seems to be an appeal to the analogy of complex numbers with rotation through 90 degrees, since a line at right angles couldn't amount to being earlier. Smith also argues that "later" is not an appropriate description of imaginary time at all, since "imaginary time is instead like a spatial dimension, in which there is no direction" (p. 318). The way the latter point is made appears to be a difficulty directed at the eliminationist interpretation.

Similarly, Deltete and Guy (1996) maintain that there is a contradiction implicit in this way of talking about the join. They argue that if real time is later than imaginary time, then it is later either in real time or in imaginary time. However, either answer implies that the one time existed in the other time's region. This aspect of the join problem is certainly problematic. We will need to keep it in mind as an important constraint on any solution to the join problem. We will see eventually that there is (at least) one way of talking about the join problem which avoids the difficulty, but that will have to wait until a much later section.

6. THE JOIN PROBLEM: VAGUENESS

We recall Davies' insistence on the gradualness of the join. Deltete and Guy reject a fuzzy transition as "facile" (1996, p. 190), and contend that the above problem of contradictoriness arises as much for a vague or smeared-out join. In addition, a further problem of consistency apparently arises: a succession of stages implies a succession of times during which time gradually came into existence, which is contradictory. This is clearly a problem which is distinctive to a vague join: it does not arise if the join is sudden, as if the universe "flips over" into our spacetime.

However, in a later paper (1997), Smith offers a defence of a vague join. He argues that Deltete and Guy assume that a metrically amorphous join is impossible, that having a topological structure at all requires metrical relations of some sort. However, it is well-known that topological structure is more general than metric. It does not logically imply a metric and can therefore logically exist without it.

Smith is correct about the logic here. Spacetime has several levels of structure, specifically metrical, affine, and topological. Each is more general than its predecessors. Therefore, the postulation of a metrically indeterminate topological structure is consistent. Hence, if that is what a vague join implies, then it cannot be convicted of contradiction.

Smith goes on to argue *for* the join region being metrically indeterminate. Though the details remain to be worked out with exactitude, it seems reasonable to say that the transition from Euclidean to Minkowski regimes is described by a superposition of quantum states described by the respective characteristic metrics. Around the join, the contributions of the two components of the wave function are approximately of the same size. Thus they cannot deliver a unique metric the way they do elsewhere. This is metrical indefiniteness, if anything is.

Again we agree. The only alternative is that the change from one metrical structure to another takes place discontinuously. While this is unquestionably a consistent possibility, in Quantum Theory such changes take place only with a measurement, and we should be avoiding measurement at all costs if we are talking about a wave function of the universe. Hence the universe as a whole is described by an analog of Schrödinger evolution, the Wheeler–DeWitt equation. The evolution is smooth and thus the change must take place gradually. Even so, perhaps there is scope for some amount of discontinuous change. It might be that the wave function contains a measure of lower-level complexity due to space at the very small scale bubbling with virtual pairs, well below the Planck length. The whole wave function of the universe would thus contain this small scale bubbling, which would occasionally manifest itself at the larger scale, and precipitate very rapid changes. This does not need measurement, only small-scale complexity.

We are not confident that this is a viable description of literally discontinuous change. However, the alternative of a vague join, i.e. Davies and

Smith's position, is certainly defensible from a charge of contradiction. A fuzzy join is not a contradiction, and may well be required by QM. Having said this, we register that Deltete and Guy's original consistency problem remains: there is a problem about how one region could be *before* the other.

7. IRREALISM

It is time to track another theme in this whole debate, namely that of the various brands of irrealism, instrumentalism, realism etc. which are on display here. Hawking's own mad-dog instrumentalism is well-known. He disparages his own clever suggestion by describing it as merely a mathematical device or trick to enable us to calculate answers about real spacetime. "So it is meaningless to ask: Which is real, 'real' or 'imaginary time'? It is simply a matter of which is the more useful description" (1988, p. 148).

Such hard-core positivism has largely departed the philosophical scene, it is pleasing to note. It is interesting to ask how Hawking envisages one description could be more useful and enabling us to calculate answers about real spacetime, if it is a mere fiction. What confidence could one have that one's descriptions had anything to say about the real world, if the real world is nothing like the way one's descriptions describe it to be? Hawking was bothered about his inability to make predictions about the history of the universe, but how could he seriously make predictions with instruments he holds not to accord, even approximately, with the world?

Deltete and Guy argue against Quentin Smith's more subtle brand of irrealism . In (1993) Smith proposes what he calls "quasi-realism". This is the thesis that only the universe after the join is real. Before that, he denies any structure corresponding to the Euclidean metric. It is true that this avoids the objection that to say that the after-join universe is later than the before-join universe is contradictory. This is done by the expedient of denying the existence of the before-join universe altogether. But this is bought at a large cost. If there is nothing before the join, then one aim of the game is given up: explanation of the very early epoch. There could be no sense of deducing the way things are from some description which one proposes is close to the way the world is. It would be just another story, that's all, and talk is cheap.

In the later paper (1997), Smith gives up this mixed irrealism in favour of a full realism of both regions. Even so, he is unable to resist a residual irrealism, we suspect. Commenting on Davies' gradualism, he flirts with the contradiction, describing space as "gradually" becoming time-like, "journey", "approached", "distant from" one region to another; and then defends the language as "definable entirely in co-ordinate and topological terms ... no fact of the matter as to whether there is a first temporal interval of each actual length" (p. 175). Now it is correct that one speaks of approaching a limit along an arbitrary curve of co-ordinates. This is at the level of abstract differential topological structure, and need have nothing to do with time *per se*. So such language should not be seized on too quickly as evidence of self-contradiction. But "no fact of the matter" is conventionalist-irrealist language, unquestionably, and in such a case one wants to know just where the facts of the matter reside.

Another version of irrealism is conjured up with the idea that the descriptions of the world in the extreme quantum region are only "symbolically" true, not literally true, "metaphorical" perhaps. In (1993), Chris Isham speaks of "the origination of the universe" in an imaginary space-time, from which the Minkowski "process ... emerges ... well away from the originating 4-sphere" (p. 74). These appear to be temporal and metrical concepts. However, Isham denies this: they are to be understood "in a symbolic sense" (p. 74).

This too is criticised by Deltete and Guy. They suspect it of instrumentalism: "In what sense is the 4-sphere an 'originating' sphere if the Lorentzian spacetime does not succeed it temporally? And in what sense is the Lorentzian real-time region 'well away from' the four-space if 'well away from' does not mean 'temporally distant from'? Finally, if the words 'emerge' and 'process' are being used symbolically, what do they symbolise?" (p. 192). Isham indeed disavows a realist interpretation: we are talking about features of 'mathematical models' of the Universe and not to features of the Universe itself (p. 193). However, Deltete and Guy's argument is the appropriate one against anyone who is tempted by metaphors: in that case, what is the cash value of the metaphor? One cannot hide behind metaphors forever. Eventually the audience wants to have some sort of indication of what the symbols are supposed to signify.

8. AN ARTIFACT OF CO-ORDINATES?

One technical version of irrealism which has attracted support is the claim that the singularity is an artifact of co-ordinates. Davies (1992, p. 67) describes the singularity at $t = 0$ as "just the coordinate singularity". Even more explicitly, John Gribbin in *Schrödinger's Kittens* (1996) writes that "mathematicians are free to choose many aspects of the co-ordinate systems they use ... only a change in choice of mathematical co-ordinates" (p. 211). Gribbin, however, also attributes to Hawking the thesis that "our everyday understanding of time is wrong, and that a better model of the way the Universe works is...imaginary time" (p. 210), and that the latter alternative is "*more* physically reasonable" (p. 213). In that case, one might ask what account of the wrongness of our ordinary understanding Gribbin has in mind. In passing, we note that Gribbin also describes Hawking's 4-sphere universe without talking about a join: Gribbin seems to take it that the entire history of the universe might really be Euclidean. We find this altogether too unbelievable, and we do not pursue it further.

However, the suggestion that somehow a co-ordinate change can avoid a real singularity is an interesting one, and worth exploring. This kind of strategy has had its successes in the history of physics. One of the better known is the use of Kruskal/Szeckeres co-ordinates to describe black holes. Early theorising about black holes made it look like they had two singularities, one at the light horizon and one right at the centre where mass-density becomes absolutely infinite. However, Kruskal and Szeckeres showed independently that the former infinity does not appear if a new set of co-ordinates is used. These are well-behaved although the transformation cannot perforce be continuous everywhere, for that would only reproduce the singularity. This device had the consequence that an observer freely falling into the black hole would not be crushed out of existence at the light horizon, but would continue on into the centre, where it would really be crushed, at the real singularity.

There is some realism here, however. To conclude that an observer would *not* be crushed at the horizon is to argue that the K/S co-ordinates are in reality *topologically correct*, and the other co-ordinates *incorrect*. In particular, the singularity in the other co-ordinates *does not exist in physical space*. So for the co-ordinate change strategy to work anywhere, it must be accompanied by the hypothesis that the world really is that way and not the other. We certainly don't mean to imply that one should follow Gribbin all the way into total irrealism about our Minkowski world, of course. Pace the early Smith, realism about both regions is possible and more reasonable. But realism about the Euclidean regime must be part of that strategy if we think that a co-ordinate change will gain us any mileage at all. Thus, mere co-ordinate change is not so mere: unless it is realist, it does not escape the difficulties of irrealism; and so far as we have got, realism hasn't escaped the join problem.

9. SPACE AND TIME AS MACROPHENOMENA

In (1999a and 1999b), Butterfield and Isham make a significant advance. They promote the idea that spacetime emerges in the large scale "from something else" (p. 25) of smaller scale. The Minkowski metrical structure is not intrinsic to spacetime, but rather is a relation between spacetime and matter: GR metric, the physics of the large, is emergent. The important point is that they contend that this is not a temporal process, and that hence all objections about contradictoriness fall down, for these rely on interpreting the metaphors temporally (p. 57).

We are very sympathetic to this suggestion as far as it goes. Indeed, it would seem to be inevitable for any theory in which there is a *distinctive* physics of the very small. To put our own spin on it, Minkowski spacetime emerges from a fuzzy Euclidean background of phenomena below the Planck length, as the scale becomes too big to be affected by quantum phenomena. This is precisely the phenomenon of quantum decoherence. It is like this: the Universe starts out smoothly as a quantum fuzz. The radius grows and grows until things begin to be too big for quantum effects. The Minkowski metric emerges as decoherence cancels out the effects of the very small, particularly on the behaviour of light. Very soon after, inflation as we know it takes over. Of course the "join region" will have an indeterminate metric, but we have seen that this is not so problematic. Butterfield and Isham rather spoil it with more irrealism, by asserting that the Euclidean region has no more physical meaning than does an imaginary spacetime trajectory in normal physics. This "does *not* mean that this solution has any ontological status in the quantum theory" (p. 59). Of course, this might be a simple caution not to be too trusting of physicists' tricks with models, which we certainly endorse but which leaves open realism somewhere. Setting aside the irrealism then, it is an attractive position. It appeals to a familiar phenomenon of scale as the mechanism at work. It unifies static and dynamic phenomena, synchronic and diachronic, as a *single process* of decoherence.

But there is still a fallacy lurking here. To say that there is one mechanism is not to say that there aren't two different processes. What we wanted to understand, was the origin of the universe, the temporal process. That remains as a *special case* of scale. Thus, there remains the special feature of a temporal join. The threatened contradiction concerns the temporal order between parts of this process, and that has not been dismissed by the simplification afforded by seeing it as part of a more general phenomenon. Hence it can't yet be said that the join problem has been avoided.

10. IMAGINARY SPACE

So, the problem of the temporal order between parts around the join is still with us. One way remains for solving the problem. This applies whether the join is fuzzy and gradual or not.

We recall the point that the substitution of $t' = \mathbf{i}t$ turns the metrical form from a signature of $(+, +, +, -)$ to $(+, +, +, +)$. This was the point about abolishing a metrical distinction between space and time. However, it invited the question of in what sense $\mathbf{i}t$ could be earlier than t. Now one thing which seems not to have been noticed is that the Minkowski signature is often written $(-, -, -, +)$ or $(+, -, -, -)$. Such a signature is regarded as an equally legitimate description of SR spacetime. It follows that there is another way to describe the transition. The transformation of the spatial co-ordinates (x, y, z) to $\mathbf{i}(x, y, z)$, that is $(\mathbf{i}x, \mathbf{i}y, \mathbf{i}z)$, will equally transform a Minkowski metrical form into a Euclidean form. That is, to transform a SR metric to Euclidean, it suffices to take imaginary space, not imaginary time. *But in this model there is no problem about which times are earlier or later*, because the time variable t is univocal throughout. The join region is a superposition of *spatially* distinct regions, with a common time. The contradiction problem dissolves. So does the problem of how the postulation of the Euclidean region can explain the Minkowski region and yield reliable calculations about it. The Euclidean region is earlier than the Minkowski region and thus causally prior to it. Indeed, we go so far as to suggest that even the problem of the fuzziness of the join looks more tractable. Vague spatial states, that is a superposition of spatial states, are fairly routine quantum weirdness, even if superposing Euclidean and Minkowski times were hard to comprehend.

This is a better solution interpretationally, we think, especially when combined with Butterfield and Isham's application of the principle of scale. All it would seem to require is the capacity to transform three sets of spatial variables through mutually orthogonal directions. The result is presumably another 3-D space at right angles to our own. This is not so absurd a thought. In fact, it's the one you would naturally come to if you thought of the change as a process, something taking place over (the one kind of) time.

Moreover, there is apparently no change in the way the calculations are conducted. Indeed, inspection of Hartle and Hawking (1983) indicates that in the way they first set up the analysis (p. 2960), the integral which is made to converge in the Euclidean regime is only indirectly a function of time through the functional dependence of space upon time in conditions of an expanding universe. The state function ψ is given as $\int e^{iS}dx$, where S is a function of x which is in turn a function of t. The transformation of t to $\mathbf{i}t$ then turns the integral into $\int e^{-S}dx$. That this has the desired effect for suitable S, x and t can be seen simply by setting $S = x = t$, which reduces the example to the simplified case given in Section 2. The functional dependence of this integral on t, that is the fact that the state function ψ so determined is a function of time, means that one can in turn proceed to calculate the probability that a given ψ occur at t by summing over "histories" which lead up to ψ at t, that is by integrating with respect to t in the usual fashion in QM. In short, it seems that there are no adverse consequences for the calculations of amplitudes and probabilities, which is thus some indirect support for the viability of our suggestion.

11. EUCLIDEAN TIME AS SPACE

In an early section we distinguished two interpretations of the ontological status of time in the Euclidean region. We have mostly been discussing the reductionist interpretation, namely that there is after all time of a sort in this region, having registering further distinctive difficulties for the eliminationist approach. It is time to see if nothing better can be said in favour of eliminationism. One place where a ghost of an eliminationist suggestion emerges is in Gribbin (1996, pp. 211–212) where he interprets Hawking's thesis to think of the expanding universe "... not in terms of an expanding bubble of spacetime that appears out of a mathematical point (the singularity) and grows, but in terms of lines of latitude drawn on the surface of a sphere *which stays a constant size*" (emphasis ours). This looks like the denial of time in the Euclidean region, even if the next few sentences take it back: "A tiny circle drawn around the north pole of the sphere represents the Universe *when young* – all of space is represented by the line that makes up the circle. As the Universe *expands*, it is represented by lines drawn further from the pole and closer to the equator" (emphasis ours)

Is it possible to give sense to the idea that it begins as space alone? Of course, our suggestion in the last section is not available, since that proposes that time exists in a fairly robust sense throughout. Still, it is not so absurd to talk this way, perhaps. Imagine that, looking backward, instead of there being a pointy temporal singularity there is a bubble of space. The bubble of space doesn't change in size. It is simply the origin of the universe, but it is not itself extended in time. *It is surely not a necessary truth that whatever exists, is temporally extended.* Another analogy that we think is useful is of a mercury thermometer. It has a (static) bubble at the bottom and out of that rises a column of mercury. The dynamic aspect of time is captured by the motion of the column as it rises. Now ask Hawking's question: how can one calculate the probability that our universe arose from a region where there is no singularity of mass-density? Would it not be a reasonable suggestion to apply the analogy of summing over all possible non-singular "histories" which join to a small but incontrovertibly Minkowski state, save that one sums over all possible non-singular spatial bubbles from which the Minkowski state initiates? We don't think that the word "initiates" gives anything away here. It is the Minkowski universe which initiates, not the spatial bubble which initiates it. Time begins as Minkowski, there is no before, but there is a finite space at the origin nonetheless.

We do not wish to endorse this way of talking too strongly. We still think that there is no clear sense in which the Minkowski region would be explained by the Euclidean region, if the latter has no time. If that is so, then what we have here is at most *description* but not *explanation*. In which case, the hypothesis offers nothing as explanation and so there are no grounds for its retention. That is, if one takes the eliminationist position about time in the Euclidean region then in turn one should eliminate the Euclidean region altogether. This of course was Smith's conclusion in arguing for his original quasi-realism. Moreover, on this suggestion the status of the calculation of the probability that the Minkowski region exists, on analogy with summing over histories in ordinary QM, is problematic. These are not "histories" in any sense. Hence it is not at all clear what could justify Hawking's claim that he has calculated the probability that the universe arises out of "... a zero three-geometry, i.e., a single point. In other words, the ground state is the amplitude for the Universe to appear from nothing" (1983, p. 2961). Further clarification of these matters is to be awaited with interest.

12. CONCLUSION

To sum up, we believe that we have sketched a consistent and coherent interpretation of the Hartle–Hawking theory of the very early epoch of the Universe. The present account appeals to a simple re-interpretation of the change of metric which is founded in actual practice in presenting the relativistic metric. It also adopts as an independent principle the Butterfield–Isham application of scale or decoherence. The result is a theory which, we propose, avoids the major problems that bedevil other interpretations, allows for a reasonable concept of explanation, and retains the calculatory advantages of the original theory.

There are other interpretational issues in this area which we do not propose to take up, since it would seem that they are not affected by the details of our argument. One issue concerns the difference between two rival accounts of the change in the literature, the Hartle–Hawking theory and Vilenkin's quantum tunneling approach. Another is the debate between Smith and Craig over Smith's contention that the sum-over-histories methodology implies an anti-theistic argument, by yielding a non-zero probability that the universe arose from nothing at all.

Instead, we will content ourselves by recalling Hawking's words in *A Brief History of Time* (p. 185):

> Up to now, most scientists have been too occupied with the development of new theories that describe what the universe is, to ask the question why. On the other hand, the people whose business is to ask why, the philosophers, have not been able to keep up with the advance of scientific theories. In the seventeenth century, philosophers considered the whole of human knowledge, including science, to be their field and discussed questions such as: Did the universe have a beginning? However, in the nineteenth and twentieth centuries, science became too technical and mathematical for the philosophers, or anyone else but a few specialists. Philosophers reduced the scope of their inquiries so much that Wittgenstein, the most famous philosopher of this century, said 'The sole remaining task for philosophy is the analysis of language'. What a comedown from the great tradition of philosophy from Aristotle to Kant!

This is unworthy of Hawking. Wittgenstein died in 1951, and did not see the rise of a generation of scientifically-literate philosophers who do regard the whole of human knowledge as their field. The list of twentieth century philosophers who have been technically well-informed about modern physics is very long. If one had to name the three most influential philosophers of the twentieth century one might guess Russell, Wittgenstein and Quine, though we doubt that Wittgenstein would be the most famous. Of these, Russell and Quine display strong mathematics and physics. Even Wittgenstein, trained as an engineer, showed no fear of technical mathematics. Wittgenstein's role in twentieth-century philosophy is a complex one, though it can be said that in his later period he held the view that philosophical puzzles arise solely from confusion over our ordinary use of language. This is understandable for its time, but its limitations were seen some decades ago. Finally, one can accuse the physicists of collaborating in the confusion by formulating their theories in ways which are infected with bad philosophy. Salient examples are the formulations of both SR and QM in terms of observers. These are positivist accretions which persist even today among physicists when better versions in terms of frames and interactions respectively exist, in no small part due to the efforts of philosophers.

We take it that the present paper, along with those we have been discussing, constitutes a counter-example to Hawking's opinions on philosophers.

ACKNOWLEDGEMENTS

We wish to thank a referee of this journal for very helpful comments, as well as Graham Nerlich, Stephanie Lewis, John Bigelow, and members of the audience at a meeting of the Australasian Association of Philosophy.

REFERENCES

Butterfield, J. and C. Isham: 1999a, 'Spacetime and the Philosophical Challenge of Quantum Gravity', e-print archives at uk.arXiv.org, gr-qc/9903072.

Butterfield, J. and C. Isham: 1999b, 'On the Emergence of Time in Quantum Gravity', gr-qc/9901024.

Gribbin, John: 1996, *Schrödinger's Kittens*, Phoenix, London.

Davies, Paul: 1992, *The Mind of God*, Penguin, London.

Davies, Paul: 1995, *About Time*, Penguin, London.

Deltete and Guy: 1996, 'Emerging from Imaginary Time', *Synthese* **108**, 185–203.

Hartle, J. and S. Hawking: 1983, 'The Wave Function of the Universe', *Physical Review* D **28**(2), 2960–2973.

Hawking, Stephen: 1988, *A Brief History of Time*, Bantam Doubleday Dell Books, London.

Hawking, Stephen and Roger Penrose: 1996, *The Nature of Space and Time*, Princeton University Press, Princeton.

Isham, Chris: 1988, 'Creation of the Universe as a Quantum Process', in R. Russell et al. (eds.), *Physics, Philosophy and Theology: A Common Quest for Understanding*, The Vatican Observatory Foundation, Vatican City.

Isham, Chris: 1988, 'Quantum Theories of the Creation of the Universe', in R. Russell and N. Murphy (eds.), *Quantum Cosmology and the Laws of Nature*, The Vatican Observatory Foundation, Vatican City.

Misner, Charles, Kip Thorne, and J.A. Wheeler: 1973, *Gravitation*, Freeman, San Francisco.

Robinson, Abraham: 1979, *Non-Standard Analysis*, North-Holland, Amsterdam.

Smith, Quentin: 1993, 'The Wave Function of a Godless Universe', in W. Craig and Q. Smith (eds.), *Theism, Atheism and the Big Bang*, Clarendon Press, Oxford.

Smith, Quentin: 1993, 'The Ontological Interpretation of the Wave Function of the Universe', *The Monist* **80**, 160–185.

Department of Philosophy
The University of Adelaide
North Tce, Adelaide, SA 5005
Australia
E-mails: Chris.Mortensen@adelaide.edu.au
jcsavas@arts.adelaide.edu.au

Manuscript submitted 5 September 2002
Final version received 21 February 2003

DHARMAKĪRTI AND PRIEST ON CHANGE

Chris Mortensen
Department of Philosophy, University of Adelaide

This essay looks at the contrasting accounts of change and motion given by the Buddhist logician Dharmakīrti (seventh century C.E.) (along with his commentator Dharmottara [eighth to ninth centuries C.E.]) and the contemporary analytical philosopher Graham Priest. In what are otherwise strikingly similar positions, they take opposite views on the Law of Non-contradiction—the former appealing to it and the latter denying it. The question of who is right is raised, and a qualified endorsement of Dharmakīrti is entered.

Priest's Inconsistent Account of Motion

Graham Priest (1987) argues that motion must be an inconsistent process. The idea of motion and change as inconsistent has a long history, from Heracleitus to Hegel. As a major premise in his proposed account, Priest, following Hegel, argues that an account of motion at a given time ought to be *intrinsic* rather than *in relation to other times.* He offers two main arguments in favor of this premise. First, he proposes that it must be possible to distinguish *from its instantaneous state alone whether a body is in motion or at rest.* An instantaneous snapshot ought to reveal which components of a process are in motion and which are not. Now, one might initially think that one would have *velocity vectors,* which intrinsically constitute the instantaneous velocity of a particle. However, Priest points out that velocity, as the derivative of position, is not intrinsic. It is a relation to nearby times; it is the limit of average velocities for shorter and shorter times. As such, motion considered as velocity could not be a property internal to the instantaneous state itself; there must be an additional ingredient to have motion.

Priest's second argument is a challenge: if you believe that Laplacean determinism is at least *possible,* then you allow that there is a possible world in which any instantaneous state in the future is determined by the present state. But this can only be so if change is intrinsic: the instantaneous non-relational state cannot determine whether a particle is in the same place or different places at *other* times. But if it were *logically necessary* that change were a relation between times, then Laplacean determinism would not even be coherent.

In order to set aside this second argument, let us be clear about what Priest is denying. He is denying the proposition that velocity being nonzero suffices for motion; that is, $v \neq 0 \rightarrow$ motion. (The converse implication might also be disputed, but that need not concern us here). Now, I suggest, one should not go so far as to say that Laplacean determinism is *impossible* or *incoherent.* It is *possible* for motion to be as Priest and Hegel say it is. On the other hand, Priest and Hegel must admit that

Philosophy East & West Volume 54, Number 1 January 2004 20–28

the relational facts of the matter are present even in their story; that is, $v \neq 0$. Hence, their view is at least vulnerable to the charge of multiplying entities beyond necessity. The alternative, that nonzero velocity is sufficient for motion, is simpler and explanatorily clearer. Moreover, there is also the difficulty that Laplacean determinism appears not to be true, because of the probabilistic nature of quantum measurement.

However, I am more interested in what Priest does with this account. He contends that a plausible intrinsic account can be given if we think of a moving particle as occupying a *small spread* or interval of spatial points at any given time. He argues that the classical position function can be thought of as the leading edge of this interval, and that speed can be thought of as a measure of the extent of the spread, with the limiting case of zero being the absence of motion.

So far so good. Indeed, so far so consistent: the particle is smeared out; it occupies spatial extension. Now, spatially extended things are no news, and hardly inconsistent. What makes the theory inconsistent is the further premise that *no single body can be in distinct positions at the same time.* More precisely, a body's having position *p1* at *t* excludes that body having position *p2* at *t*, if *p1* and *p2* are different positions. Moreover, this premise applies *even when* both positions belong to the postulated spread of positions at *t*, a point that might seem questionable if the particle is thought of as spatially extended. Keep the premise in mind. Priest does not offer any further argument for it that I can find, but it is essential to his account. We will see that Dharmakīrti is likewise committed to a parallel assumption.

Dharmakīrti's Account of Change

Dharmakīrti's central argument is found in *A Short Treatise on Logic (Nyāya Bindu),* along with Dharmottara's Commentary (see Stcherbatsky [1930] 1962, especially 1 : 103–105, 414; and 2 : 8, 94, 196–197). What follows is a modernized reconstruction, but it essentially follows Stcherbatsky's own twentieth-century reconstruction.

Consider a thing that changes from having a property F (such as blue) at *t1* to having an incompatible property G (such as yellow) at *t2*. Then the body-at-*t1* has an incompatible property from the body-at-*t2*. Thus, they cannot be identical: what it means to have incompatible properties is that the one excludes the other; one and the same thing cannot possess both. Hence, all things are made up of temporal atoms or time slices that are entirely distinct; no two are identical. No one thing changes; there are merely differences between distinct time slices. Thus the Buddhist doctrine of momentariness is deduced.

Stcherbatsky describes this as an application of the "Law of Contradiction," which is the traditional name for what we now commonly call the "Law of Noncontradiction" (LNC): ~(P & ~P). LNC is indeed involved in the reconstructed argument below, and Dharmakīrti and Dharmottara certainly spend considerable time defending it (secs. 12 ff.). But one might at least implicate Leibniz' Law (LL), as well: if two things are identical then they have the same properties. The argument is of the form:

1. First, $x = y$ & Fx & Gy & $(z)(Fz \rightarrow \sim Gz)$ implies $x = y$ & Fx & $\sim Fy$ (by first-order-logic principles).
2. But, $x = y$ & Fx & $\sim Fy$ implies Fx & $\sim Fx$ (by LL).
3. Now, $\sim(Fx$ & $\sim Fx)$ (LNC).
4. So, $\sim(x = y$ & Fx & $\sim Fy)$ (2, 3, modus tollens).
5. However, Fx & Gy (by observation).
6. Hence, Fx & $\sim Fy$ (since F and G are incompatible).
7. So, $\sim(x = y)$ (4, 6, first-order logic).

Here, x and y are the body-at-*t1* and body-at-*t2*, respectively, and $(z)(Fz \rightarrow \sim Gz)$ expresses the incompatibility of F and G, the mutual exclusion of F and G.

Notice that Priest is likewise committed to (1) and (2) and indeed their antecedents: body b-at-*t1*, in motion, has the properties of being at place *p1* and at place *p2*, and, as we have seen, being at *p1* excludes being at *p2*. (Dharmakīrti's move from [5] to [6] is parallel.) Now, for Priest, there is no question as to whether b-at-*t1* is self-identical. Since these facts are certain, there is no question of their being overridden by LNC. Hence, being at *p1* and not being at *p1* both apply to b-at-*t1*, and motion is inconsistent. One person's modus ponens is another's modus tollens.

Leibniz' Law

In passing, it is worth noticing that there are different versions of Leibniz' Law in operation here. Three equivalents can be distinguished. The whole inference above is given by:

4.1 $(x)(y)((Fx \,\&\, Gy \,\&\, (z)(Fz \rightarrow \sim Gz)) \rightarrow \sim(x = y))$

Substituting $\sim F$ for G satisfies $(z)(Fz \rightarrow \sim Gz)$, so this implies:

4.2 $(x)(y)((Fx \,\&\, \sim Fy) \rightarrow \sim(x = y))$

The latter is obviously equivalent to LL in its familiar form:

4.3 $(x)(y)(x = y \rightarrow (Fx \rightarrow Fy))$

Conversely, assume (4.2) and the antecedent of (4.1). From the latter, we may deduce $(z)(Gz \rightarrow \sim Fz)$, and hence $Gy \rightarrow \sim Fy$. Applying modus ponens gives $\sim Fy$, and conjoining gives Fx & $\sim Fy$. This is the antecedent of (4.2), so we obtain $\sim(x = y)$. Hence, (4.1) is equivalent also.

Space-time Theory

Time Slices

Who is right here—Priest or Dharmakīrti? I confess a definite inclination to side with Dharmakīrti and the Buddhists, with some qualifications.

The first and most obvious point is that twentieth-century space-time theory sides with Dharmakīrti. Space-time is a whole whose atomic parts are space-time points. These stack into slices that make up space-time frames. Continuing objects are thus made up of time slices, which are themselves spatially complex, in accordance with Buddhist theory. The basic items of reality are space-time points, and everything else is wholes of these. Indeed, it is tempting to believe that Buddhist thought on this point constitutes the first coherent attempt at a space-time ontology, although, of course, its vindication from speculative status awaited much more sophisticated science.

Relativistic Invariance

Furthermore, there is a problem for Priest and Hegel, namely that in relativity theory motion is relative to frames, not absolute. Hence Priest's theory is not relativistically invariant: the spread of inconsistent locations vanishes in a frame in which the body is at rest.

Actually this is not such a damaging problem. So what if inconsistent theory is appropriate for the description of space-time frames but not invariance? This locates its sphere of usefulness, but does not destroy it. Even so, inconsistency would be eroded as a basic character of the world. It would enjoy the same status as space and time separately, without enjoying the full reality of the invariants of space-time.

No Atomic Times?

The viability of this account of the basic ontology of the universe turns on whether the universe is, at bottom, atomic, that is, punctate. At the very least it needs that time be punctate. If time had no instants, only nested shorter and shorter intervals, then there would be no ultimate time slices out of which to build continuing existences.

It is certainly true that relativity theory, both special and general, postulates space-time as punctate. Thus, our best theory of the (large-scale) structure of the universe supports the atomicity of space and time (in the small scale, which is hardly surprising!). Quantum theory is much more equivocal. It is true that Hilbert spaces, the phase spaces of QM, are described by sharp-valued coordinates including space and time coordinates. But, as many texts point out, this is something of an illusion: the Heisenberg Uncertainty relations for spatial position P and time T require that both have indeterminacy: $\delta P.\delta Q$ and $\delta T.\delta E$ are both bounded below (Q and E here are, respectively, momentum and energy). There are technical issues here: it is possible to recover exact values for position and duration using a formal device called the Dirac Delta function, but the cost is a considerable complication of the mathematics, so the resulting theory is less simple.

However, it is also true that a few theorists have suggested that it is unnecessary to have exact positions and times (see, e.g., Mortensen and Nerlich 1978). The idea here is simply that there are intervals of ever-diminishing size, but no basic sizeless atoms of space and time. This seems to have been Aristotle's view, and it is certainly a natural view, if only because no one has ever detected a point. It is suggested here that there is no conclusive proof either from evidence or from a priori argument, or both, that reality *could not* be temporally nonatomic. If this right, then support for Dharmakīrti must be tempered with the proviso that reality should turn out to be as our best current space-time theory says it is, rather than the non-punctate alternative.

Differentiating a Scalar Field

Stcherbatsky raises the example of the differential calculus ([1930] 1962, 1 : 106–108). He contends that it shows that time is punctate. As we have already seen, there is no doubt that differential calculus employs functions on sets of points that are isomorphic with the Real Numbers. So anyone wishing to deny the atomicity of reality ought to show how calculus can be done on the kinds of manifolds they are postulating, for it is in terms of differential equations that the basic laws of nature are written. But it should not be thought that this is such an impossible task. There are various approaches that have been taken regarding this problem; for example, on the use of synthetic differential geometry, see Kock 1981 or Bell 1988.

The application of calculus in physical theories actually looks at first glance to make things worse for Stcherbatsky and the Buddhists. For instance, we find the velocity of a particle by considering the position function of *the particle* over time. Now this is already problematic for Buddhist logic, since it postulates a *particular* that continues to be self-identical at different times: a particle is a particular, if anything is. Similarly, wanting to find the particle's change of mass over time, or the dynamics of its temperature, involves us in the *mass function M(t)* and *temperature function T(t)* of the particle. Thus, it seems that the Buddhist doctrine of momentariness is immediately defeated by differentiation, since the latter postulates identity over time.

More careful accounts of taking the derivative of a scalar field avoid this problem for the Buddhists, although further problems emerge. We suppose that the universe (or a region of it) is pervaded by a field, such as a field of gravitational potentials, or a field of temperatures. That is, a scalar quantity-kind, such as temperature, takes a definite value at different points in space (the space in question might not be physical space but phase space). The universals in question are the individual temperatures at points in the field: they are universals because they can take an identical value at different places distant from one another. Thus far, there are no continuing particulars. There is a problem over the identity of the *kind* of field at different points: all temperatures, for example. But if we have universals at all, then presumably there is in principle no problem with higher-order properties of universals and relations between universals: some such account of what binds quantities together into quantity-kinds would have to be right.

Now we take the derivative by taking the average variation in the field over smaller and smaller nested intervals of the independent variable (e.g., distance or duration) in a given direction as determined at the point at which the derivative is being taken. The limit of these averages as the independent variable tends to zero is stipulated to be the rate of change of the field in that direction. Note that one can differentiate at a point in whatever direction and along whatever pathway we choose. There is no preferred direction such as following the track of a particle in phase space. One can choose to do that, of course, but differentiation does not *require* it.

There is *one* preferred direction of the field, the one in which the magnitude of the derivative is the greatest, and this is called the *gradient* of the field. This is the track down which the system will proceed without outside interference. Think of the contours of a hillside: the route down which water will flow is the route of steepest descent. But that doesn't significantly alter the point of this story, which is this: that we don't have to track a particle through time or space. There is just the all-pervasive field with a universal instantiated at every point in it, and we can differentiate in whatever direction we choose.

Differentiating a Vector Field: Direction and Distance at Different Places

For the sake of completeness, it needs to be noted that space-time theory has to do more differentiating than this, and this leads to an increase in the conceptual resources required. In addition to differentiating a scalar field one wants to take the second derivative of position, namely acceleration. Taking the next derivative by differentiating the resulting vector field is not an immediate requirement, however, for there is no natural way of comparing vectors at different points, because vectors have a direction as well as a magnitude, and we have yet to import directions. This is solved in general relativity theory by introducing the further concept of an *affine connection*. An affine connection provides for the idea of *parallelism,* specifically of a *parallel transport* between vectors attached to differing (nearby) points. This, evidently, serves to introduce the idea of *sameness of direction* at different points. In point of fact, something even stronger is introduced into space-time: the idea of a *metric*. A metric provides a comparison of distance between neighboring points, and along a path. The distance here is generally space-time distance rather than spatial or temporal distance. A metric determines a unique affine connection but not vice versa.

The metaphysics of distance and direction have to be carefully worked out. It is apparent that space and space-time cannot adequately be described without systems of *relations,* so realism about spatial universals is realism about relational universals of various kinds. I just want to claim here that this apparatus would seem not to need anything more in principle than the identity of various universals at different points. Useful work on this has been done by Peter Forrest (Forrest and Armstrong 1987; Forrest 1988) and John Bigelow (1988).

Universals and Particulars

Thus, the Buddhist doctrine of momentariness had best be interpreted as the claim of disidentity of atomic particulars. Lurking in the background of this discussion has been the distinction between universals and particulars, which the Buddhist logicians certainly knew about. Now it is *continuing particulars* that are being primarily denied: universals are not the sorts of things of which one can take space-time slices. And it is particulars that are relevant to religious practice and the Buddhist doctrine of impermanence, for it would seem that we ourselves are particulars. In the last section above, we defended Dharmakīrti from the specter of the Moving Particle, by arguing that properly understood differentiation does not require the identity of objects over time. A more modern account of field theory requires scalar universals distributed across space-time. Universals are fully in all of their instances, so there is identity across space-time; but, of course, from the beginning the theory of universals was designed precisely for universals to be self-identical over time and space.

Unfortunately Dharmakīrti's position is made more difficult by the strong tradition of Buddhist nominalism to which he subscribed. This is hardly surprising—the Buddhists appear as the Knee-Jerk Impermanence Police here. But it does mean that the defense given above to the objection from the Moving Particle is unavailable. Notice, too, that this cannot be turned by insisting that quantities are merely relational after all, for relations are equally universals. I suggest, however, that admitting universals is a small price to pay. Still, there are several complications, as we shall see.

First, in Plato's theory universals are eternal, existing independently of their instances. But in Aristotle's rival account, which the Buddhist logicians knew about, universals are immanent; that is, they exist in and only in their instances. This surely opens the door for impermanent universals: a universal ceases to exist when all of its instances do. Hence, adopting Aristotle's account would permit a universal doctrine of impermanence. This battle is not yet settled. In recent decades we have had Michael Tooley (1977) arguing in the name of Plato against David Armstrong's Aristotelian universals (1978). Armstrong and Tooley agree that a (basic) law of nature is a relation between universals; however, Tooley argues that laws can fail to be instantiated (such as a law governing a temperature that as a matter of fact is higher than any realized). In that case, if one wants to say that the counterfactual "if that temperature were realized, then . . ." holds, one would have to look for a truth maker, and the only truth makers are relations between universals that continue to exist when uninstantiated.

Second, for the Buddhist logicians, universals do not exist in the external world, only in the mind (where they are genuinely existing). It is natural to project objective properties onto the world, but these are constructs of an active mind, illusions. All there is in the mind-independent world are "powers" to produce (universal) ideas in us, and it is well known that different powers can produce the same mental result. The doctrine of mental universals was in strong dispute with realist schools such as *Nyāya* and *Vaiśeṣika*. Stcherbatsky ([1930] 1962, 1 : 446) discusses "The Experiment of Dharmakīrti," which was influential as an argument for the mentality of universals. This argument proposes that the phenomenon of inattention shows that perceptions can be present while the understanding fails to be engaged. In turn, this shows that understanding is an active and constructive process. But it is understanding that is employed when general concepts are employed. If there were an objective sameness in the world corresponding to the general term, then an active and constructive understanding would be otiose; knowledge would simply be produced directly in the mind. Against this argument, while we can agree that inattention is certainly a philosophically interesting phenomenon, and while it is unquestionable that the mind constructs and projects so that some of our classifications are constructed, still this does not go to show that there are no objective samenesses in the world. Indeed, the very employment of the language of "powers" is general; we must allow at least the lively possibility that the same power manifests itself in more than one place.

Third, an assumption unexamined so far is that a person is a particular. That is what is needed to salvage the religious relevance of the doctrine of impermanence from the threat of universals. But there also lurks the possibility that the soul is a universal, one in many. This kind of view was adopted some centuries later by Averroës, for example. If it were right, then when the appropriate universal is later instantiated, so is the person. This opens the way for a kind of survival, even eternal survival if Platonic universals are adopted. Lest the position seem too ancient, reflect that Hilary Putnam's 1960 functionalism postulated that minds were systems of relations between logical states, and multiply realizable. This does seem to have the consequence that if identical functional states are realized in two different places (as with copying a disk, for example), then there would be the same mind. Even so, it is very difficult to believe that souls are universals. If we took two identical twins, so identical that they had the same thoughts, then it would surely not follow that they had the same mind. If one dies while the other continues to live, then one mind exists but the other does not. This strong intuition is, I think, at the basis of the idea that a person is not a universal.

Whatever is the case here, the *scientific* arguments for universals of some form are strong. Even before modern field theory and relativity theory, Newton's physics was up to its neck in universals. Consider the law of gravitation, $F = Gm_1m_2/r^2$, which relates the force of gravitational attraction to masses and distances in their respective units. Masses are obviously universal-like, since they can be identical in different bodies and at different times. The whole law postulates a regularity across space and time: wherever such mass universals are instantiated, they will attract one another with the same precise force and undergo the same acceleration. These are not mere powers, but substantive postulates about existing things. (On these matters see, e.g., Forrest and Armstrong 1987; Bigelow 1988; Mortensen 1987.)

It must be conceded to Dharmakīrti's Experiment that it is right to be suspicious of the genuine unity of external universals. For example, the case of colors makes it clear that what is in the world *might be* a power (to produce secondary qualities) and that different powers can produce the same mental effect. Intuitions about what is constructed and projected are thus apt to be unreliable. But this proper caution is bypassed by the results of our best science. Unless a wholesale reconstruction of basic physical theory can be demonstrated, the criterion of rationality here is the usual scientific one, and real generalities and regularities must be accepted.

Conclusion

Thus, Buddhist nominalism is an unnecessary restriction. All that the doctrine of impermanence needs in order to be applicable to people's lives is that people are particulars and that particulars are impermanent. The thesis that people are particulars is a rejection of souls as universals. We noted that this is not quite so obvious but that it is reasonable nonetheless. The thesis that particulars are impermanent is common experience, as the Buddha saw. It would seem from Dharmakīrti's argument, then, that we have reason to believe his conclusion of the momentariness of existence, appropriately construed so as to allow universals, against Priest's inconsistent account. However, we also noticed that there are various complicating factors. Chief among these was the question of the contingent truth or falsity of the thesis that physical reality is atomic. Only if that is true can we make Dharmakīrti's argument work.

References

Armstrong, D. M. 1978. *Universals and Scientific Realism.* 2 vols. Cambridge: Cambridge University Press.

Bell, John L. 1988. "Infinitesimals." *Synthese* 75 : 285–316.

Bigelow, John. 1988. *The Reality of Number: A Physicalist's Philosophy of Mathematics.* Oxford: Clarendon Press.

Forrest, Peter. 1988. "Supervenience: The Grand-Property Hypothesis." *Australasian Journal of Philosophy* 66 : 1–12.

Forrest, Peter, and David Armstrong. 1987. "The Nature of Number." *Philosophical Papers* 16 : 165–186.

Kock, Anders. 1981. *Synthetic Differential Geometry.* Cambridge: Cambridge University Press.

Mortensen, Chris. 1987. "Arguing for Universals." *Revue Internationale de Philosophie* 160 : 97–111.

Mortensen, Chris, and Graham Nerlich. 1978. "Physical Topology." *Journal of Philosophical Logic* 5 : 209–223.

Priest, Graham. 1987. *In Contradiction.* Dordrecht: Nijhoff.

Stcherbatsky, F. Th. [1930] 1962. *Buddhist Logic.* 2 vols., including (vol. 2) *A Short Treatise on Logic,* by Dharmakīrti with Commentary by Dharmottara. New York: Dover.

Tooley, Michael. 1977. "The Nature of Laws." *Canadian Journal of Philosophy* 7 : 667–698.

PART 6

Computer Studies

AUTOMATED REASONING PROJECT

TECHNICAL REPORT TR-ARP-9/89 20 February 1989

COMPUTING DUAL PARACONSISTENT AND INTUITIONIST LOGICS

Chris Mortensen
Steve Leishman

TECHNICAL REPORT SERIES

Research School of Social Sciences
Australian National University

COMPUTING DUAL PARACONSISTENT AND INTUITIONIST LOGIC

Chris Mortensen

and

Steve Leishman

Department of Philosophy

The University of Adelaide

ADELAIDE S.A. 5000

ABSTRACT

Topologies of closed sets yield paraconsistent logics. Hence the closed-open duality transfers over to a topological duality between paraconsistent and intuitionistic logics. Finite topological spaces can be generated from finite posets by a well-known construction. So in this study all nonisomorphic finite posets to size five are generated and nonisomorphic paraconsistent and intuitionist operators are computed for each. The latter computation also required computation of Boolean operators on the power sets up to size five. It is proposed to apply these methods to study the solution of inconsistent sets of linear equations.

Table of Contents

Computing Dual Paraconsistent and Intuitionist Logics

1. Topological Spaces and Logics

It is well known that intuitionism has a natural connection with topology, in that the usual intuitionist propositional calculus can be semantically characterised in terms of Heyting algebras, which in turn arise as algebras of open sets on a topological space. It is also known, though less well known, that algebras of closed sets are inconsistency-tolerating, or *paraconsistent*. By this we mean that theories can be constructed by allowing sentences to take values from elements of an appropriate algebra (in this case an algebra of closed sets) in such a way that the theory is negation-inconsistent while nontrivial, i.e. not every sentence is in the theory.

It is not difficult to see informally how this works, by looking at the intuitionist case first. A Boolean algebra is isomorphic to a field of sets. The sets can be thought of as sets of "worlds". Assigning a proposition to an element of a Boolean algebra, a set of worlds, can be regarded as equivalent to declaring the proposition to be true at just those worlds. If the field of sets is additionally endowed with a topological structure, then the intuitionist point of view can be described as proposing that truth is only ever truth on open sets of worlds, that a proposition is only ever true at the points of an open set. Since intuitionist negation is an operation on intuitionist propositions, the points at which the negation of a proposition are true also form an open set, the obvious candidate being the interior of the set-theoretic (Boolean) complement of the original set. But this leaves the possibility that the set on which either a proposition or its intuitionist complement are true, which is to say the union of the two open sets, will not be the whole space unless the open sets are also clopen. The whole space is, needless to say, an open set. Indeed, it is the set on which, according to Intuitionism, propositions must always be true in order to be semantically valid (logical truths). Every topological space thus gives rise to a logic, as the set of sentences which, when their atomic sentence-letters are assigned to arbitrary open sets, are true at all points in the space. Hence the law of excluded middle $A \vee \neg A$ fails to be valid in such logics.

The dual paraconsistent point of view requires that propositions only be assigned to the closed sets of a topological space. If paraconsistent negation (in this paper written $\ulcorner$) is to be an operation on closed sets, then the obvious candidate is the closure of the Boolean complement. With the natural assumption that conjunction corresponds to set-theoretic intersection, this means that a proposition and its paraconsistent negation will both be true on the boundary of the two sets, itself a closed set and non-null unless the closed sets are also clopen. So, if a theory is determined by the stipulation that its members be those sentences true on any of a suitable collection of closed sets which includes a boundary, then that theory will be negation-inconsistent but in general nontrivial.

These ideas can be presented more formally as follows. Begin with a propositional language $\mathcal{L}$ with binary operators (&,v) and a unary operator $\ulcorner$. A function $v : \mathcal{L} \to C$, where C is any set-theoretic closure algebra, is a ***paraconsistent valuation*** if it satisfies (i) $v(p) \in C$ for any atomic letter p (ii) $v(A \& B) = v(A) \cap v(B)$ (iii) $v(A \vee B) = v(A) \cup v(B)$ (iv) $v(\ulcorner A)$ = the closure of the set-theoretic complement of A. A subset $\mathcal{D}$ of C is a ***semifilter*** iff if $S_1 \in \mathcal{D}$ and $S_1 \subseteq S_2$ then $S_2 \in \mathcal{D}$. A semifilter $\mathcal{D}$ is a ***filter*** iff if $S_1 \in \mathcal{D}$ and $S_2 \in \mathcal{D}$ then $S_1 \cap S_2 \in \mathcal{D}$. A ***semilogic*** L is a set of propositions closed w.r.t. uniform substitution and having a binary consequence relation $\models_L$ satisfying: if $A \in L$ and $A \models_L B$ then $B \in L$. If $A \in L$ we also write $\models_L A$. If a semilogic L also satisfies the condition that if $\models_L A$ and $\models_L B$ then $\models_L A \& B$, then it is a ***logic***. A closure algebra C and a filter $\mathcal{D}$ on it then determine a logic L via the stipulations (i) $L = \{A$: for all paraconsistent valuations v on C, $v(A) \in \mathcal{D}\}$ and (ii) $A \models_L B$ iff for all paraconsistent valuations v on C, $v(A) \subseteq v(B)$. There are other possible definitions of $\models$ here, for example $(\forall v)$ (if $v(A) \in \mathcal{D}$ then $v(B) \in \mathcal{D}$); but since there can be different $\mathcal{D}$'s on the same C, it seems a good idea to have a $\models$ which is independent of any particular $\mathcal{D}$. The set $\mathcal{D}$ is called the ***designated*** *values* of L in C. Note that such a set L is always closed w.r.t. uniform substitutions, conjunctions and the consequence relation $\models_L$. If C is a finite closure algebra, then such an L is *finite-valued.*

An *L-semitheory* relative to a logic L is any set of sentences closed w.r.t. $\models_L$. (L-semitheories are not necessarily closed w.r.t. uniform substitutions.) An L-semitheory also closed w.r.t. conjunctions is an *L-theory*. An L-theory Th is *inconsistent* iff for some A, both $A \in Th$ and $\ulcorner A \in Th$, otherwise *consistent*; and *trivial* iff $Th = \mathcal{L}$, otherwise *nontrivial.* A logic L is a *paraconsistent* logic iff there are inconsistent but nontrivial L-theories.

If L_1, L_2 are logics determined by filters $\mathcal{D}_1$ and $\mathcal{D}_2$ on the same closure algebra C, then if $\mathcal{D}_1 \subset \mathcal{D}_2$ then $L_1 \subseteq L_2$. (The latter $\subseteq$ cannot be strengthened to $\subset$ because of the possibility that no theorem A of L_2 take values in $\mathcal{D}_2 - \mathcal{D}_1$.) So, to produce a weaker logic on the same closure algebra, restrict the designated values. The limiting case of weakness is $\mathcal{D} = \{\vee\}$ where $\vee$ is the maximal element. The other limiting case, $\mathcal{D} = C - \{\wedge\}$ where $\wedge$ is the minimal element, does not in general produce more than a semilogic since such a $\mathcal{D}$ might only be a semifilter.

A set of designated values can be used to determine a theory as well. If v is any paraconsistent valuation on C then choosing $\mathcal{D}$ and setting $Th = \{A : v(A) \in \mathcal{D}\}$ determines a theory of the logic determined by C and $\mathcal{D}$. If logics $L_1 \subset L_2$, then every L_2-semitheory is an L_1-semitheory and every L_2-theory is an L_1-theory, though the converses fail. So a theory determined by a valuation and set of designated values is a theory of any logic determined by that $\mathcal{D}$ or any smaller $\mathcal{D}$. So if $\mathcal{D}_1 \subset \mathcal{D}_2$ and L_2 is a paraconsistent logic, so is L_1. Obviously also, theories determined by smaller sets of designated values are smaller theories.

These definitions can easily be reworked substituting intuitionist negation $\neg$ for paraconsistent negation $\ulcorner$, an open set algebra O for the closure algebra C, and the interior of the set theoretic complement for the value of a negated formula. So there is a natural topological duality between intuitionist logics and theories on the one hand and the above paraconsistent logics and theories on the other. There are other paraconsistent logics (e.g. the relevant logics and the Brazilian logics) to which these remarks do not apply, though the present structures can fairly be characterised as Brazilian-style. The duality runs quite deep, in that the topological duality is reflected in the fact that topos theory gives rise as naturally to paraconsistent logics as to intuitionist logic, as announced in [10] and demonstrated in [11]. In the light of these results, the usual public-relations exercise for intuitionism simply cannot be sustained.

Finite-valued intuitionist logics and their corresponding finite interior algebras have proved useful in model theory. Their dual paraconsistent logics are beginning to be useful as inconsistent mathematics is developed (see [9] also [5] [7] and [8]). So it would seem to be useful to display all such structures, and that is what the present study aims at. Further computational directions and applications are reviewed in the final section.

2. Implication and Deducibility

The present approach to paraconsistent logics seems to be simpler than those of Fitting [2] and Hardegree [4]. However, computational questions about the implication operator $\rightarrow$ on paraconsistent logics are consciously ignored here. There are a couple of reasons for this. In the first place, the present state of inconsistent mathematics is quite extensional or zero degree, as is the bulk of classical mathematics. Insofar as one is interested in questions of deducibility one has the natural metalinguistic $\models$ corresponding to $\leq$ (or $\subseteq$) on the background lattice, so that inserting an $\rightarrow$ into the object language seems fairly unnecessary. Second, the exercise of adding an $\rightarrow$ to closed set algebras and logics is not too difficult. This is contrary to what seems to be suggested by Goodman in [3], though he deserves credit for seeing the connection between closed sets and inconsistency-toleration, also noticed by Fitting, Hardegree and Lawvere.

A number of natural desiderata for $\models$ and $\rightarrow$ can be given. Among these are:

(1) $\models A$, $A \models B$ /∴ $\models B$.

(2) For any theory Th, $A \in Th$, $A \models B$ /∴ $B \in Th$.

(3) $A \models B$ iff $\models A \rightarrow B$.

(4) $\models A$, $\models A \rightarrow B$ /∴ $\models B$.

Now (1) holds for any $\mathcal{D}$ closed upward w.r.t. $\subseteq$ as we have required, since $(\forall v)(v(A) \in \mathcal{D})$ and $(\forall v)(v(A) \subseteq v(B))$ ensure $(\forall v)(v(B) \in \mathcal{D})$. Number (2) holds because it has been stipulated for theories. On (3), one has to fix two variables, a definition of $\rightarrow$ as well as $\mathcal{D}$. If (a): we define $v(A \rightarrow B) = \vee$ (equivalently $= \ulcorner A \mathrm{v} B$) iff $v(A) \subseteq v(B)$, and $v(A \rightarrow B) = \wedge$ otherwise; then (3) holds for any $\mathcal{D} \subset C$ (since if $(\forall v)(v(A) \subseteq v(B))$ then $(\forall v)(v(A \rightarrow B) = \vee)$, while if $(\exists v)(v(A) \not\subseteq v(B))$ then $(\exists v)(v(A \rightarrow B) = \wedge \notin \mathcal{D})$. This makes this definition of $\rightarrow$ suitable for the logic which is the intersection of all closed-set logics, since it is appropriately general. Note that $\models (A \& \ulcorner A) \rightarrow B$ fails in general. Another definition for $\rightarrow$ is (b): $v(A \rightarrow B) = v(\ulcorner A \mathrm{v} B$ if $v(A) \subseteq v(B)$, and $v(A \rightarrow B) = v(B)$ otherwise. This definition is not too far from the $\rightarrow$ of RM and $RM2n+1$, having the first clause in common. Here, (3) fails except if $\mathcal{D}$ is defined as in intuitionism to $= \{\vee\}$, which yields the minimal paraconsistent logic for that C. Indeed, given this $\mathcal{D}$ then another definition for $\rightarrow$, namely (c): the full RM-ish $\rightarrow$ with $v(A \rightarrow B) = v(\ulcorner A \& B)$ if $v(A) \not\subseteq v(B)$, also satisfies (3).

Given (1) and (3), (4) follows. A fifth possible condition on $\models$ and $\rightarrow$, the semantic deduction theorem $A_1, \ldots, A_n \models B$ iff $A_1, \ldots, A_{n-1} \models A_n \rightarrow B$, needs a definition of $A_1, \ldots, A_n \models B$. Given the natural $\mathrm{glb}\{v(A_i) : 1 \leq i \leq n\} \subseteq v(B)$, the SDT fails for all three definitions (a)-(c) of the previous paragraph. This is not necessarily a bad thing since the same happens in the usual relevant logics. Various one-way versions hold.

So there are various interesting implications on these structures with natural and general properties and we do not pursue the matter further.

3. Summary of Algorithm

It is standard theory that topological spaces can be generated from preordered sets, where a preorder is a reflexive transitive relation. The present approach detours via modal logic in what we believe is a naturally intuitive way. Given any subset S of a preordered set thought of as subset of worlds at which a given proposition P is true, the interior of S is that subset of S at which the proposition *necessarily* P would be true, where necessarily P is interpreted as for the modal logic S4. This is $\{x : (\forall y)(x \leq y \supset y \in S)\}$. The closure of S is that superset of S where the proposition *possibly* P would be true. This is $\{x : (\exists y)(x \leq y \& y \in S)\}$. The open sets are then those elements of the power set Boolean algebra which are interiors, and the closed sets are those which are closures. If we begin with a preorder, then each of these satisfy the postulates for open sets and closed sets of a topological space respectively, and moreover the open sets are the Boolean complements of the closed sets, and vice versa. In particular the open sets are closed w.r.t. $\cap$ and $\cup$, as are the closed sets. Intuitionist negations of propositions (i.e. sets of worlds) are then computable, as the interiors of the set-theoretic complements of open sets; and paraconsistent negations are computable as closures of set-theoretic complements of closed sets.

Computing the preorders is of interest. The number of different preorders on, say, 5 worlds, is large, and in particular contains many isomorphs, so considerable pruning was required. First, preorders which contained symmetric pairs of worlds, i.e. for which $x \neq y$ but $x \leq y$ and $y \leq x$, were eliminated. This was done on the grounds that when turned into algebras such pairs of worlds behaved as a unit, both being either in a closed set or out of it, both in an open set or out of it, so could be replaced with an equivalence class and thus be isomorphic to a topological space of smaller size. Thus $\leq$ was required to be antysymmetric as well as transitive and reflexive, turning the preorder into a poset.

Disconnected outrider worlds for which weak connectedness $(\exists y)(x \leq y \vee y \leq x)$ fails unless $x = y$, represent topological spaces for which the singleton $\{x\}$ is a clopen set. Thus the subspace $\{\vee, \{x\}, \vee - \{x\}, \wedge\}$ is as logic, a Boolean algebra. We took the opportunity to eliminate such well-behaved structures as of lesser interest. A reasonably fast isomorphism test for posets was obtained by matching each world of one with the list of those worlds of the other with the same total of arrows in and arrows out. The set of all permutations without repeat from those lists is the set of candidate isomorphisms. These can then be tested according to whether the two arrays are identical under any candidate isomorphism. If even one candidate produces identity, it is an isomorphism; while if none do, the two posets are nonisomorphic. It speeds the isotest considerably if only those orders are considered which are subsets of the order on the natural numbers, since any other order is isomorphic to one of these. This was done early in the poset generation loop.

So the procedure is: generate all five power-set Boolean algebras for universes up to size = 5 (2^n elements for dimension n), together with their set theoretic operations $\cup, \cap, -$. The $\cup$ and $\cap$ double as v and & on the logics. For each size, generate and test all binary relations which are reflexive, antisymmetric, transitive and embeddable in the natural order on that size. We called these Frasts, for Function Reflexive AntiSymmetric and Transitive. (The 'F' arises because we represented these by their characteristic functions.) Eliminate isomorphs, producing isofrasts; and use these as kernels around which to permute isofrasts of greater size. Discard isofrasts failing weak connectedness, producing isoconfrasts. The isotest is the major slowing factor beyond size 5 and parallelism is indicated here. Fix an isoconfrast, and for each element of the power set Boolean algebra, compute its interior (necessitation) and closure (possibilification). The sets of interiors and closures are two lattices on which intuitionist and paraconsistent negations are computed, as described above. Repeat for each isoconfrast. The programme is dynamic, since topological operations and negations are computed as each new isoconfrast arises.

Logics are displayed up to size = 5 (47 logics at that size), with the elements of the power set coded as binary 5-tuples, and recoded in decimal more in keeping with their traditional conception as values of a many-valued logic. Those subsets of the power set Boolean algebra which are interiors then constitute the values of the intuitionist sublogic, while the closures are the values of the paraconsistent sublogic. Inclusion of tables for necessity and possibility means that the results also include computations of finite-valued S4-ish modal and epistemic logics, of 2^n values. We note that the role necessity plays in intuitionist negation qualifies it as an intuitionist operator, while correspondingly possibility is a paraconsistent operator.

Future directions which we hope to pursue include the following. Parallelising the algorithm as noted above, and concomitantly increasing the size, are indicated: run times for the present algorithm on a VAX are pretty limiting (see Table I). Visual display of results can be substantially improved using geometrical structures rather than numbers. Posets fairly obviously lend themselves to geometry. A Boolean algebra size = n (i.e. n baseworlds) is naturally thought of as an n-dimensional structure with two subalgebras, paraconsistent and intuitionist. Add in a third variable, distinguished elements, and one has a case for geometry and colour. A different approach to representing n-dimensional Boolean algebras utilises the interesting and surprisingly difficult equivalent problem of visually representing Venn diagrams for more than three sets, on which see [1]. Finally, application of these results to mathematical theories or finite-valued reasoners, partly inconsistent and partly intuitionist, should be made. Here we have in mind adapting the methods of [5] to study the interaction between functionality and propositional aspects discovered therein.

Thanks for useful suggestions are due to Glen Farrell and Mike Dunn. We also acknowledge as one of our sources Meyer and Slaney's excellent [6].

4. Bibliography

1. Edwards, Anthony, 'Venn Diagrams for Many Sets', *New Scientist*, 7 Jan. 1989, 51-56.

2. Fitting, Melvin, 'Logic Programming on a Topological Bilattice', offprint.

3. Goodman, Nicholas, 'The Logic of Contradiction', *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, Bd. 27 (1981), 119-126.

4. Hardegree, Gary, 'Bitopological Spaces and The Representation of Distributive Lattices', offprint.

5. Meyer, Robert K. and Mortensen, Chris, 'Inconsistent Models for Relevant Arithmetics', *Journal of Symbolic Logic* 49 No.3 (Sept. 1984), 917-929.

6. Meyer, Robert K. and Slaney, John, 'Computing De Morgan Monoids', *Technical Reports on Automated Reasoning*, TR-ARP-1/87, Australian National University, 1987.

7. Mortensen, Chris, 'Inconsistent Nonstandard Arithmetic', *Journal of Symbolic Logic* 52 No.2 (June 1987), 512-518.

8. Mortensen, Chris, 'Inconsistent Number Systems', *Notre Dame Journal of Formal Logic* 29 No.1 (1988), 45-60.

9. Mortensen, Chris, 'Models for Inconsistent and Incomplete Differential Calculus', to appear.

10. Mortensen, Chris, 'On the Logic of Toposes', delivered to the Logic Group, Australian National University, Nov. 1986.

11. Mortensen, Chris and Lavers, Peter, 'Paraconsistency and Topos Logic', to appear.

RESULTS

Table I Runtimes VAX 11/785

Size (no. of baseworlds)	Best CPU time (millisec)	No. of logics
1	130	1
2	40	2
3	120	3
4	690	11
5	9100	47
6	209900 (isoconfrasts only)	255

Table II: Finite-valued paraconsistent, intuitionist and modal logics to size = 5.

SIZE = 1

Boolean tables

&	1	0	~
1	1	0	0
0	0	0	1

v	1	0
1	1	1
0	1	0

Isoconfrast No. = (1,1)

1

begin Logic(1,1)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
1	1	1	1	0	0	1	1	0	0
0	0	0	0	1	1	0	0	1	1

end Logic(1,1)

SIZE = 2

Boolean tables

&	3	2	1	0	~
3	3	2	1	0	0
2	2	2	0	0	1
1	1	0	1	0	2
0	0	0	0	0	3

v	3	2	1	0
3	3	3	3	3
2	3	2	3	2
1	3	3	1	1
0	3	2	1	0

Isoconfrast No. = (2,1)

1 1
0 1

begin Logic(2,1)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
3	11	3	11	0	00	3	11	0	00
2	01	3	11	1	10	2	01	0	00
1	10	1	10	3	11	0	00	2	01
0	00	0	00	3	11	0	00	3	11

end Logic(2,1)

SIZE = 3

Boolean tables

&	7	6	5	4	3	2	1	0
7	7	6	5	4	3	2	1	0
6	6	6	3	2	3	2	0	0
5	5	3	5	1	3	0	1	0
4	4	2	1	4	0	2	1	0
3	3	3	3	0	3	0	0	0
2	2	2	0	2	0	2	0	0
1	1	0	1	1	0	0	1	0
0	0	0	0	0	0	0	0	0

~
0
1
2
3
4
5
6
7

v	7	6	5	4	3	2	1	0
7	7	7	7	7	7	7	7	7
6	7	6	7	7	6	6	7	6
5	7	7	5	7	5	7	5	5
4	7	7	7	4	7	4	4	4
3	7	6	5	7	3	6	5	3
2	7	6	7	4	6	2	4	2
1	7	7	5	4	5	4	1	1
0	7	6	5	4	3	2	1	0

Isoconfrast No. = (3,1)

1 0 1
0 1 1
0 0 1

begin Logic(3,1)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
7	111	7	111	0	000	7	111	0	000
6	011	7	111	1	100	6	011	0	000
5	101	7	111	2	010	5	101	0	000
4	110	4	110	7	111	0	000	3	001
3	001	7	111	4	110	3	001	0	000
2	010	2	010	7	111	0	000	5	101
1	100	1	100	7	111	0	000	6	011
0	000	0	000	7	111	0	000	7	111

end Logic(3,1)

_____Isoconfrast No. = (3,2) ________________________

1 1 1
0 1 1
0 0 1

________________________________ begin Logic(3,2) _______

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
7	111	7	111	0	000	7	111	0	000
6	011	7	111	1	100	6	011	0	000
5	101	7	111	4	110	3	001	0	000
4	110	4	110	7	111	0	000	3	001
3	001	7	111	4	110	3	001	0	000
2	010	4	110	7	111	0	000	3	001
1	100	1	100	7	111	0	000	6	011
0	000	0	000	7	111	0	000	7	111

end Logic(3,2)

_____Isoconfrast No. = (3,3) ________________________

1 1 1
0 1 0
0 0 1

________________________________ begin Logic(3,3) _______

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
7	111	7	111	0	000	7	111	0	000
6	011	7	111	1	100	6	011	0	000
5	101	5	101	4	110	3	001	2	010
4	110	4	110	5	101	2	010	3	001
3	001	5	101	4	110	3	001	2	010
2	010	4	110	5	101	2	010	3	001
1	100	1	100	7	111	0	000	6	011
0	000	0	000	7	111	0	000	7	111

end Logic(3,3)

SIZE = 4

Boolean tables

&	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	~
15	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	0
14	14	14	10	9	8	10	9	8	4	3	2	4	3	2	0	0	1
13	13	10	13	7	6	10	4	3	7	6	1	4	3	0	1	0	2
12	12	9	7	12	5	4	9	2	7	1	5	4	0	2	1	0	3
11	11	8	6	5	11	3	2	8	1	6	5	0	3	2	1	0	4
10	10	10	10	4	3	10	4	3	4	3	0	4	3	0	0	0	5
9	9	9	4	9	2	4	9	2	4	0	2	4	0	2	0	0	6
8	8	8	3	2	8	3	2	8	0	3	2	0	3	2	0	0	7
7	7	4	7	7	1	4	4	0	7	1	1	4	0	0	1	0	8
6	6	3	6	1	6	3	0	3	1	6	1	0	3	0	1	0	9
5	5	2	1	5	5	0	2	2	1	1	5	0	0	2	1	0	10
4	4	4	4	4	0	4	4	0	4	0	0	4	0	0	0	0	11
3	3	3	3	0	3	3	0	3	0	3	0	0	3	0	0	0	12
2	2	2	0	2	2	0	2	2	0	0	2	0	0	2	0	0	13
1	1	0	1	1	1	0	0	0	1	1	1	0	0	0	1	0	14
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15

v	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
14	15	14	15	15	15	14	14	14	15	15	15	14	14	14	15	14
13	15	15	13	15	15	13	15	15	13	13	15	13	13	15	13	13
12	15	15	15	12	15	15	12	15	12	15	12	12	15	12	12	12
11	15	15	15	15	11	15	15	11	15	11	11	15	11	11	11	11
10	15	14	13	15	15	10	14	14	13	13	15	10	10	14	13	10
9	15	14	15	12	15	14	9	14	12	15	12	9	14	9	12	9
8	15	14	15	15	11	14	14	8	15	11	11	14	8	8	11	8
7	15	15	13	12	15	13	12	15	7	13	12	7	13	12	7	7
6	15	15	13	15	11	13	15	11	13	6	11	13	6	11	6	6
5	15	15	15	12	11	15	12	11	12	11	5	12	11	5	5	5
4	15	14	13	12	15	10	9	14	7	13	12	4	10	9	7	4
3	15	14	13	15	11	10	14	8	13	6	11	10	3	8	6	3
2	15	14	15	12	11	14	9	8	12	11	5	9	8	2	5	2
1	15	15	13	12	11	13	12	11	7	6	5	7	6	5	1	1
0	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

Isoconfrast No. = (4,1)

1 1 0 1
0 1 0 1
0 0 1 1
0 0 0 1

begin Logic(4,1)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	15	1111	5	1100	10	0011	0	0000
12	1101	15	1111	3	0010	12	1101	0	0000
11	1110	11	1110	15	1111	0	0000	4	0001
10	0011	15	1111	5	1100	10	0011	0	0000
9	0101	15	1111	6	1010	9	0101	0	0000
8	0110	11	1110	15	1111	0	0000	4	0001
7	1001	15	1111	11	1110	4	0001	0	0000
6	1010	6	1010	15	1111	0	0000	9	0101
5	1100	5	1100	15	1111	0	0000	10	0011
4	0001	15	1111	11	1110	4	0001	0	0000
3	0010	3	0010	15	1111	0	0000	12	1101
2	0100	5	1100	15	1111	0	0000	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,1)

Isoconfrast No. = (4,2)

1 1 0 1
0 1 0 0
0 0 1 1
0 0 0 1

begin Logic(4,2)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	13	1011	5	1100	10	0011	2	0100
12	1101	15	1111	3	0010	12	1101	0	0000
11	1110	11	1110	13	1011	2	0100	4	0001
10	0011	13	1011	5	1100	10	0011	2	0100
9	0101	15	1111	6	1010	9	0101	0	0000
8	0110	11	1110	13	1011	2	0100	4	0001
7	1001	13	1011	11	1110	4	0001	2	0100
6	1010	6	1010	15	1111	0	0000	9	0101
5	1100	5	1100	13	1011	2	0100	10	0011
4	0001	13	1011	11	1110	4	0001	2	0100
3	0010	3	0010	15	1111	0	0000	12	1101
2	0100	5	1100	13	1011	2	0100	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,2)

Isoconfrast No. = (4,3)

1 1 0 0
0 1 0 0
0 0 1 1
0 0 0 1

begin Logic(4,3)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	13	1011	5	1100	10	0011	2	0100
12	1101	15	1111	3	0010	12	1101	0	0000
11	1110	11	1110	10	0011	5	1100	4	0001
10	0011	10	0011	5	1100	10	0011	5	1100
9	0101	15	1111	6	1010	9	0101	0	0000
8	0110	11	1110	13	1011	2	0100	4	0001
7	1001	13	1011	11	1110	4	0001	2	0100
6	1010	6	1010	15	1111	0	0000	9	0101
5	1100	5	1100	10	0011	5	1100	10	0011
4	0001	10	0011	11	1110	4	0001	5	1100
3	0010	3	0010	15	1111	0	0000	12	1101
2	0100	5	1100	13	1011	2	0100	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,3)

Isoconfrast No. = (4,4)

1 1 1 1
0 1 0 1
0 0 1 1
0 0 0 1

begin Logic(4,4)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	15	1111	5	1100	10	0011	0	0000
12	1101	15	1111	6	1010	9	0101	0	0000
11	1110	11	1110	15	1111	0	0000	4	0001
10	0011	15	1111	5	1100	10	0011	0	0000
9	0101	15	1111	6	1010	9	0101	0	0000
8	0110	11	1110	15	1111	0	0000	4	0001
7	1001	15	1111	11	1110	4	0001	0	0000
6	1010	6	1010	15	1111	0	0000	9	0101
5	1100	5	1100	15	1111	0	0000	10	0011
4	0001	15	1111	11	1110	4	0001	0	0000
3	0010	6	1010	15	1111	0	0000	9	0101
2	0100	5	1100	15	1111	0	0000	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,4)

Isoconfrast No. = (4,5)

1 1 1 1
0 1 0 0
0 0 1 1
0 0 0 1

begin Logic(4,5)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	13	1011	5	1100	10	0011	2	0100
12	1101	15	1111	6	1010	9	0101	0	0000
11	1110	11	1110	13	1011	2	0100	4	0001
10	0011	13	1011	5	1100	10	0011	2	0100
9	0101	15	1111	6	1010	9	0101	0	0000
8	0110	11	1110	13	1011	2	0100	4	0001
7	1001	13	1011	11	1110	4	0001	2	0100
6	1010	6	1010	15	1111	0	0000	9	0101
5	1100	5	1100	13	1011	2	0100	10	0011
4	0001	13	1011	11	1110	4	0001	2	0100
3	0010	6	1010	15	1111	0	0000	9	0101
2	0100	5	1100	13	1011	2	0100	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,5)

Isoconfrast No. = (4,6)

1 1 1 1
0 1 0 0
0 0 1 0
0 0 0 1

begin Logic(4,6)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	13	1011	5	1100	10	0011	2	0100
12	1101	12	1101	6	1010	9	0101	3	0010
11	1110	11	1110	7	1001	8	0110	4	0001
10	0011	13	1011	5	1100	10	0011	2	0100
9	0101	12	1101	6	1010	9	0101	3	0010
8	0110	11	1110	7	1001	8	0110	4	0001
7	1001	7	1001	11	1110	4	0001	8	0110
6	1010	6	1010	12	1101	3	0010	9	0101
5	1100	5	1100	13	1011	2	0100	10	0011
4	0001	7	1001	11	1110	4	0001	8	0110
3	0010	6	1010	12	1101	3	0010	9	0101
2	0100	5	1100	13	1011	2	0100	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,6)

Isoconfrast No. = (4,7)

1 1 1 1
0 1 1 1
0 0 1 1
0 0 0 1

begin Logic(4,7)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	15	1111	5	1100	10	0011	0	0000
12	1101	15	1111	11	1110	4	0001	0	0000
11	1110	11	1110	15	1111	0	0000	4	0001
10	0011	15	1111	5	1100	10	0011	0	0000
9	0101	15	1111	11	1110	4	0001	0	0000
8	0110	11	1110	15	1111	0	0000	4	0001
7	1001	15	1111	11	1110	4	0001	0	0000
6	1010	11	1110	15	1111	0	0000	4	0001
5	1100	5	1100	15	1111	0	0000	10	0011
4	0001	15	1111	11	1110	4	0001	0	0000
3	0010	11	1110	15	1111	0	0000	4	0001
2	0100	5	1100	15	1111	0	0000	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,7)

Isoconfrast No. = (4,8)

1 1 1 1
0 1 1 1
0 0 1 0
0 0 0 1

begin Logic(4,8)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	15	1111	5	1100	10	0011	0	0000
12	1101	12	1101	11	1110	4	0001	3	0010
11	1110	11	1110	12	1101	3	0010	4	0001
10	0011	15	1111	5	1100	10	0011	0	0000
9	0101	12	1101	11	1110	4	0001	3	0010
8	0110	11	1110	12	1101	3	0010	4	0001
7	1001	12	1101	11	1110	4	0001	3	0010
6	1010	11	1110	12	1101	3	0010	4	0001
5	1100	5	1100	15	1111	0	0000	10	0011
4	0001	12	1101	11	1110	4	0001	3	0010
3	0010	11	1110	12	1101	3	0010	4	0001
2	0100	5	1100	15	1111	0	0000	10	0011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,8)

Isoconfrast No. = (4,9)

1 0 0 1
0 1 0 1
0 0 1 1
0 0 0 1

begin Logic(4,9)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	15	1111	2	0100	13	1011	0	0000
12	1101	15	1111	3	0010	12	1101	0	0000
11	1110	11	1110	15	1111	0	0000	4	0001
10	0011	15	1111	5	1100	10	0011	0	0000
9	0101	15	1111	6	1010	9	0101	0	0000
8	0110	8	0110	15	1111	0	0000	7	1001
7	1001	15	1111	8	0110	7	1001	0	0000
6	1010	6	1010	15	1111	0	0000	9	0101
5	1100	5	1100	15	1111	0	0000	10	0011
4	0001	15	1111	11	1110	4	0001	0	0000
3	0010	3	0010	15	1111	0	0000	12	1101
2	0100	2	0100	15	1111	0	0000	13	1011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,9)

Isoconfrast No. = (4,10)

1 0 1 1
0 1 1 1
0 0 1 1
0 0 0 1

begin Logic(4,10)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	15	1111	2	0100	13	1011	0	0000
12	1101	15	1111	11	1110	4	0001	0	0000
11	1110	11	1110	15	1111	0	0000	4	0001
10	0011	15	1111	5	1100	10	0011	0	0000
9	0101	15	1111	11	1110	4	0001	0	0000
8	0110	11	1110	15	1111	0	0000	4	0001
7	1001	15	1111	11	1110	4	0001	0	0000
6	1010	11	1110	15	1111	0	0000	4	0001
5	1100	5	1100	15	1111	0	0000	10	0011
4	0001	15	1111	11	1110	4	0001	0	0000
3	0010	11	1110	15	1111	0	0000	4	0001
2	0100	2	0100	15	1111	0	0000	13	1011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,10)

Isoconfrast No. = (4,11)

1 0 1 1
0 1 1 1
0 0 1 0
0 0 0 1

begin Logic(4,11)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
15	1111	15	1111	0	0000	15	1111	0	0000
14	0111	15	1111	1	1000	14	0111	0	0000
13	1011	15	1111	2	0100	13	1011	0	0000
12	1101	12	1101	11	1110	4	0001	3	0010
11	1110	11	1110	12	1101	3	0010	4	0001
10	0011	15	1111	5	1100	10	0011	0	0000
9	0101	12	1101	11	1110	4	0001	3	0010
8	0110	11	1110	12	1101	3	0010	4	0001
7	1001	12	1101	11	1110	4	0001	3	0010
6	1010	11	1110	12	1101	3	0010	4	0001
5	1100	5	1100	15	1111	0	0000	10	0011
4	0001	12	1101	11	1110	4	0001	3	0010
3	0010	11	1110	12	1101	3	0010	4	0001
2	0100	2	0100	15	1111	0	0000	13	1011
1	1000	1	1000	15	1111	0	0000	14	0111
0	0000	0	0000	15	1111	0	0000	15	1111

end Logic(4,11)

SIZE - 5

Boolean tables

&	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	~
31	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	0
30	30	30	25	24	23	22	25	24	23	22	15	14	13	12	11	10	15	14	13	12	11	10	5	4	3	2	5	4	3	2	0	0	1
29	29	25	29	21	20	19	25	15	14	13	21	20	19	9	8	7	15	14	13	5	4	3	9	8	7	1	5	4	3	0	1	0	2
28	28	24	21	28	18	17	15	24	12	11	21	9	8	18	17	6	15	5	4	12	11	2	9	8	1	6	5	4	0	2	1	0	3
27	27	23	20	18	27	16	14	12	23	10	9	20	7	18	6	16	5	14	3	12	2	10	9	1	7	6	5	0	3	2	1	0	4
26	26	22	19	17	16	26	13	11	10	22	8	7	19	6	17	16	4	3	13	2	11	10	1	8	7	6	0	4	3	2	1	0	5
25	25	25	25	15	14	13	25	15	14	13	15	14	13	5	4	3	15	14	13	5	4	3	5	4	3	0	5	4	3	0	0	0	6
24	24	24	15	24	12	11	15	24	12	11	15	5	4	12	11	2	15	5	4	12	11	2	5	4	0	2	5	4	0	2	0	0	7
23	23	23	14	12	23	10	14	12	23	10	5	14	3	12	2	10	5	14	3	12	2	10	5	0	3	2	5	0	3	2	0	0	8
22	22	22	13	11	10	22	13	11	10	22	4	3	13	2	11	10	4	3	13	2	11	10	0	4	3	2	0	4	3	2	0	0	9
21	21	15	21	21	9	8	15	15	5	4	21	9	8	9	8	1	15	5	4	5	4	0	9	8	1	1	5	4	0	0	1	0	10
20	20	14	20	9	20	7	14	5	14	3	9	20	7	9	1	7	5	14	3	5	0	3	9	1	7	1	5	0	3	0	1	0	11
19	19	13	19	8	7	19	13	4	3	13	8	7	19	1	8	7	4	3	13	0	4	3	1	8	7	1	0	4	3	0	1	0	12
18	18	12	9	18	18	6	5	12	12	2	9	9	1	18	6	6	5	5	0	12	2	2	9	1	1	6	5	0	0	2	1	0	13
17	17	11	8	17	6	17	4	11	2	11	8	1	8	6	17	6	4	0	4	2	11	2	1	8	1	6	0	4	0	2	1	0	14
16	16	10	7	6	16	16	3	2	10	10	1	7	7	6	6	16	0	3	3	2	2	10	1	1	7	6	0	0	3	2	1	0	15
15	15	15	15	15	5	4	15	15	5	4	15	5	4	5	4	0	15	5	4	5	4	0	5	4	0	0	5	4	0	0	0	0	16
14	14	14	14	5	14	3	14	5	14	3	5	14	3	5	0	3	5	14	3	5	0	3	5	0	3	0	5	0	3	0	0	0	17
13	13	13	13	4	3	13	13	4	3	13	4	3	13	0	4	3	4	3	13	0	4	3	0	4	3	0	0	4	3	0	0	0	18
12	12	12	5	12	12	2	5	12	12	2	5	5	0	12	2	2	5	5	0	12	2	2	5	0	0	2	5	0	0	2	0	0	19
11	11	11	4	11	2	11	4	11	2	11	4	0	4	2	11	2	4	0	4	2	11	2	0	4	0	2	0	4	0	2	0	0	20
10	10	10	3	2	10	10	3	2	10	10	0	3	3	2	2	10	0	3	3	2	2	10	0	0	3	2	0	0	3	2	0	0	21
9	9	5	9	9	9	1	5	5	5	0	9	9	1	9	1	1	5	5	0	5	0	0	9	1	1	1	5	0	0	0	1	0	22
8	8	4	8	8	1	8	4	4	0	4	8	1	8	1	8	1	4	0	4	0	4	0	1	8	1	1	0	4	0	0	1	0	23
7	7	3	7	1	7	7	3	0	3	3	1	7	7	1	1	7	0	3	3	0	0	3	1	1	7	1	0	0	3	0	1	0	24
6	6	2	1	6	6	6	0	2	2	2	1	1	1	6	6	6	0	0	0	2	2	2	1	1	1	6	0	0	0	2	1	0	25
5	5	5	5	5	5	0	5	5	5	0	5	5	0	5	0	0	5	5	0	5	0	0	5	0	0	0	5	0	0	0	0	0	26
4	4	4	4	4	0	4	4	4	0	4	4	0	4	0	4	0	4	0	4	0	4	0	0	4	0	0	0	4	0	0	0	0	27
3	3	3	3	0	3	3	3	0	3	3	0	3	3	0	0	3	0	3	3	0	0	3	0	0	3	0	0	0	3	0	0	0	28
2	2	2	0	2	2	2	0	2	2	2	0	0	0	2	2	2	0	0	0	2	2	2	0	0	0	2	0	0	0	2	0	0	29
1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	1	0	30
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31

v	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31
30	31	30	31	31	31	31	30	30	30	30	31	31	31	31	31	31	30	30	30	30	30	30	31	31	31	31	30	30	30	30	31	30
29	31	31	29	31	31	31	29	31	31	31	29	29	29	31	31	31	29	29	29	31	31	31	29	29	29	31	29	29	29	31	29	29
28	31	31	31	28	31	31	31	28	31	31	28	31	31	28	28	31	28	31	31	28	28	31	28	28	31	28	28	28	31	28	28	28
27	31	31	31	31	27	31	31	31	27	31	31	27	31	27	31	27	31	27	31	27	31	27	27	31	27	27	27	31	27	27	27	27
26	31	31	31	31	31	26	31	31	31	26	31	31	26	31	26	26	31	31	26	31	26	26	31	26	26	26	31	26	26	26	26	26
25	31	30	29	31	31	31	25	30	30	30	29	29	29	31	31	31	25	25	25	30	30	30	29	29	29	31	25	25	25	30	29	25
24	31	30	31	28	31	31	30	24	30	30	28	31	31	28	28	31	24	30	30	24	24	30	28	28	31	28	24	24	30	24	28	24
23	31	30	31	31	27	31	30	30	23	30	31	27	31	27	31	27	30	23	30	23	30	23	27	31	27	27	23	30	23	23	27	23
22	31	30	31	31	31	26	30	30	30	22	31	31	26	31	26	26	30	30	22	30	22	22	31	26	26	26	30	22	22	22	26	22
21	31	31	29	28	31	31	29	28	31	31	21	29	29	28	28	31	21	29	29	28	28	31	21	21	29	28	21	21	29	28	21	21
20	31	31	29	31	27	31	29	31	27	31	29	20	29	27	31	27	29	20	29	27	31	27	20	29	20	27	20	29	20	27	20	20
19	31	31	29	31	31	26	29	31	31	26	29	29	19	31	26	26	29	29	19	31	26	26	29	19	19	26	29	19	19	26	19	19
18	31	31	31	28	27	31	31	28	27	31	28	27	31	18	28	27	28	27	31	18	28	27	18	28	27	18	18	28	27	18	18	18
17	31	31	31	28	31	26	31	28	31	26	28	31	26	28	17	26	28	31	26	28	17	26	28	17	26	17	28	17	26	17	17	17
16	31	31	31	31	27	26	31	31	27	26	31	27	26	27	26	16	31	27	26	27	26	16	27	26	16	16	27	26	16	16	16	16
15	31	30	29	28	31	31	25	24	30	30	21	29	29	28	28	31	15	25	25	24	24	30	21	21	29	28	15	15	25	24	21	15
14	31	30	29	31	27	31	25	30	23	30	29	20	29	27	31	27	25	14	25	23	30	23	20	29	20	27	14	25	14	23	20	14
13	31	30	29	31	31	26	25	30	30	22	29	29	19	31	26	26	25	25	13	30	22	22	29	19	19	26	25	13	13	22	19	13
12	31	30	31	28	27	31	30	24	23	30	28	27	31	18	28	27	24	23	30	12	24	23	18	28	27	18	12	24	23	12	18	12
11	31	30	31	28	31	26	30	24	30	22	28	31	26	28	17	26	24	30	22	24	11	22	28	17	26	17	24	11	22	11	17	11
10	31	30	31	31	27	26	30	30	23	22	31	27	26	27	26	16	30	23	22	23	22	10	27	26	16	16	23	22	10	10	16	10
9	31	31	29	28	27	31	29	28	27	31	21	20	29	18	28	27	21	20	29	18	28	27	9	21	20	18	9	21	20	18	9	9
8	31	31	29	28	31	26	29	28	31	26	21	29	19	28	17	26	21	29	19	28	17	26	21	8	19	17	21	8	19	17	8	8
7	31	31	29	31	27	26	29	31	27	26	29	20	19	27	26	16	29	20	19	27	26	16	20	19	7	16	20	19	7	16	7	7
6	31	31	31	28	27	26	31	28	27	26	28	27	26	18	17	16	28	27	26	18	17	16	18	17	16	6	18	17	16	6	6	6
5	31	30	29	28	27	31	25	24	23	30	21	20	29	18	28	27	15	14	25	12	24	23	9	21	20	18	5	15	14	12	9	5
4	31	30	29	28	31	26	25	24	30	22	21	29	19	28	17	26	15	25	13	24	11	22	21	8	19	17	15	4	13	11	8	4
3	31	30	29	31	27	26	25	30	23	22	29	20	19	27	26	16	25	14	13	23	22	10	20	19	7	16	14	13	3	10	7	3
2	31	30	31	28	27	26	30	24	23	22	28	27	26	18	17	16	24	23	22	12	11	10	18	17	16	6	12	11	10	2	6	2
1	31	31	29	28	27	26	29	28	27	26	21	20	19	18	17	16	21	20	19	18	17	16	9	8	7	6	9	8	7	6	1	1
0	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

Isoconfrast No. = (5,1)

1 0 1 0 1
0 1 1 0 1
0 0 1 0 1
0 0 0 1 1
0 0 0 0 1

begin Logic(5,1)

No.	Binary Character	Possibility (closure)		Paraconsistent Negation		Necessity (interior)		Intuitionist Negation	
31	11111	31	11111	0	00000	31	11111	0	00000
30	01111	31	11111	1	10000	30	01111	0	00000
29	10111	31	11111	2	01000	29	10111	0	00000
28	11011	31	11111	16	11100	15	00011	0	00000
27	11101	31	11111	4	00010	27	11101	0	00000
26	11110	26	11110	31	11111	0	00000	5	00001
25	00111	31	11111	6	11000	25	00111	0	00000
24	01011	31	11111	16	11100	15	00011	0	00000
23	01101	31	11111	8	10010	23	01101	0	00000
22	01110	26	11110	31	11111	0	00000	5	00001
21	10011	31	11111	16	11100	15	00011	0	00000
20	10101	31	11111	11	01010	20	10101	0	00000
19	10110	26	11110	31	11111	0	00000	5	00001
18	11001	31	11111	26	11110	5	00001	0	00000
17	11010	17	11010	31	11111	0	00000	14	00101
16	11100	16	11100	31	11111	0	00000	15	00011
15	00011	31	11111	16	11100	15	00011	0	00000
14	00101	31	11111	17	11010	14	00101	0	00000
13	00110	26	11110	31	11111	0	00000	5	00001
12	01001	31	11111	26	11110	5	00001	0	00000
11	01010	11	01010	31	11111	0	00000	20	10101
10	01100	16	11100	31	11111	0	00000	15	00011
9	10001	31	11111	26	11110	5	00001	0	00000
8	10010	8	10010	31	11111	0	00000	23	01101
7	10100	16	11100	31	11111	0	00000	15	00011
6	11000	6	11000	31	11111	0	00000	25	00111
5	00001	31	11111	26	11110	5	00001	0	00000
4	00010	4	00010	31	11111	0	00000	27	11101
3	00100	16	11100	31	11111	0	00000	15	00011
2	01000	2	01000	31	11111	0	00000	29	10111
1	10000	1	10000	31	11111	0	00000	30	01111
0	00000	0	00000	31	11111	0	00000	31	11111

end Logic(5,1)

“Inconsistent Control Systems”

co-authored with Steve Leishman

The University of Adelaide, 1997

Inconsistent Control Systems

1. Introduction

Paraconsistent logic is driven by the idea of inconsistency tolerance, particularly the rejection of the principle that from a contradiction everything can be deduced. This suggests a methodology for the broad theory of *fault tolerant* control systems: where a fault or malfunction arises, seek to represent the fault as a contradiction and then exploit the contradiction-containment capacities of paraconsistent logics. This methodology is pursued in the present study. An *inconsistent controller* is constructed in software simulation, which represents malfunctioning by means of an inconsistent *virtual model* of the situation, that incorporates elements from both the expected operation of the system and its observed operation. Results are reported below which indicate that there are conditions under which such an inconsistent controller is capable of returning a system to correct functioning.

2. Inconsistent Systems of Linear Equations.

The inconsistent case of systems of linear equations is well known, but little has been done to analyse its structure. The present section follows Chapter 8 of Mortensen (1995).

Definition 1 A matrix M is *row reduced* if (a) every leading entry of a nonzero row is 1, and (b) every column containing such a leading entry 1 has all other entries zero. M is in *solution form* if additionally (c) each zero row comes below all nonzero rows, and M is in *row echelon form* if additionally (d) leading coefficients begin further to the right as one goes down.

It is known that any matrix can be placed in row echelon form by elementary row operations, and that the two matrices represent systems of linear equations with the same sets of solutions. The weaker solution form suffices for a reasonably tidy presentation of the solution of a system of linear equations with all zero (nonindependent) rows shifted to the bottom, and in this paper we work mainly with that.

Consider a system S of n linear equations in s unknowns $x_1..x_s$, having an nxs coefficient matrix $Mc=[a_{ij}]$ and an $nx(s+1)$ augmented matrix $Ma=[Mc,B]$, where $B=col[b_1..b_n]$ is the column vector of constants. The system of equations which Ma represents can be solved by reducing the Ma to solution form and the solution read off. Reducing Ma to solution form is of course the same as reducing Mc to solution form for a consistent set of equations.

Now if one has an inconsistent system of linear equations, then one can always reduce its augmented matrix to row echelon form. The resulting matrix will contain a lowest nonzero row which has zero in all places except for a 1 in the right hand column. This represents an equation of the form $0.x_1+..+0.x_s=1$, which is an inconsistency. However, the bottom 1 in the right hand column will also have been used to reduce all other entries on the right hand column to zero, which destroys the information necessary to solve the consistent set of equations above that row. Clearly this has no sensitivity in seeking solutions to inconsistent systems of equations.

Definition 2 An augmented matrix $Ma=[Mc,B]$ is in *weak row echelon form* (WREF), if Mc is in solution form.

Clearly we have:

Theorem 3 If S is a consistent set of linear equations, then Ma is in WREF iff Ma is in solution form. That is, Mc is in solution form iff Ma is in solution form.

This suggests a methodology for the inconsistent case. Reduce Ma to WREF, ie. reduce Mc to solution form, but do not otherwise touch the right-hand column of constants except in performing the elementary row operations necessary to bring Ma to WREF. The bottom portion of the WREF Ma consists of a series of rows with zeros everywhere except perhaps in the right hand column. These represent a series of inconsistent identities $0=r_1$, $0=r_2$.. which must be satisfied in the solution. A *solution* to the original inconsistent system of equations consists of a (consistent) set of values for the unknowns, together with a (finite) set of inconsistent identities $\{0=r_i\}$, obtained by reducing the original augmented matrix to WREF. Together, these suffice to make the original set of equations hold simultaneously in an inconsistent space.

There are two complications which are avoided here (see Mortensen 1995 p80) First, multiple inconsistent identities require one to associate each inconsistent identity with one dimension of the phase space, ie. one of the unknowns $x_1,..x_s$. This means in turn that one has to introduce a more complicated geometrical interpretation. For the present application, it is sufficient to remain with the simpler case of just a single inconsistency. Second, multiple WREFS are obtainable from the one inconsistent set of equations. However, this does not render the overall solution indeterminate, since there are only a finite number of WREFs obtainable from the finite number of reorganisations of the initial set of linear equations. That is, any set of equations has in general multiple solutions, as in the consistent case. The simplest way to deal with this here is to make the assumption common in engineering of some preferred ordering, such as reliability. That way, the only time row interchange is used is as part of a final manouvre to shift a row of zeros (except perhaps for the last place) downward to obtain WREF.

3. Control Systems.

A control system operating correctly and stably is represented in the usual way by a linear transformation M operating on a (column) vector of inputs u to produce a (column) vector of outputs y, or $y = M.u$. For simplicity, y and u are assumed here to be the same length, so that M is square. A more detailed analysis incorporating feedback and a state vector x is standardly given by supposing four matrices A,B,C,D with the two relations: $x(t+1) = Ax(t) + Bu(t)$ and $y(t) = Cx(t) + Du(t)$. However, it is not necessary to incorporate these relations here.

An unexpected and persistent change is postulated in the output. This can be regarded as resulting from a change in the physical laws of the plant hitherto described by M. This prompts a distinction between the matrices M*old* and M*new*. M*old* is the original M, and is responsible for the predicted output *ypred* via *ypred*=M*old*.*u* . M*new* is the actual laws of the plant, and is responsible for the observed output *yobs* via *yobs*=M*new*.*u* . M*old* is known from the original specifications of the plant, but M*new* is unknown though its output *yobs* is known. One can now define a plant to be *wellfunctioning* iff *yobs*=*ypred*, otherwise *malfunctioning*.

To represent malfunctioning as an inconsistency, there are a number of options. First note that one represents the equation $y = M.u$ as an augmented matrix whose right hand column is the outputs y , and whose remaining body is the product M.u. Reducing this to solution form provides the solution of the input u which produces that output in the consistent case. To incorporate aspects of what one knows into an inconsistent picture, one option is to form the *augmented checkmatrix*, `Mac`, which consists of a core which is M*old*.*u*, a right hand column which is *yobs*, and a bottom row (the checkrow), each entry in which is the sum of that column, except that the right hand bottom entry is Σ*ypred.* (This worked for the simple cases described here, but a more generally useful entry is $\Sigma(ypred\text{-}yobs)^2$.) When Σ*ypred* is not the same as Σ*yobs*, reducing M*ac* to WREF gives a nonzero entry called *circ* in the bottom right hand corner. The number *circ* is a parameter describing an inconsistent environment which is part of the solution for *u* to the simultaneous equations which inconsistently identify M*old.u* with *yobs*.

The conjecture, then, is that if a controller is built which at each instant computes the WREF for M*ac*, then adjusts the input *u* to a set of values which take the inconsistent environment of *circ* into account, it may be possible to return *yobs* to *ypred* without complete shutdown (zero output). Again, there are a number of options in adjusting *u*. One is to return *u* to *u*mod*circ*, which was used in these studies. An alternative is to return *u* to the largest 0mod*circ* less than its existing value. Yet another alternative is to use whatever *u* it takes to return *yobs* (using M*old*) to the largest 0mod*circ* less than its existing value. In representing this situation, it is also useful to make a distinction commonly made in control theory between the input vector *u* and the state vector *x* which takes into account feedback from the controller and on which M*old* operates directly.

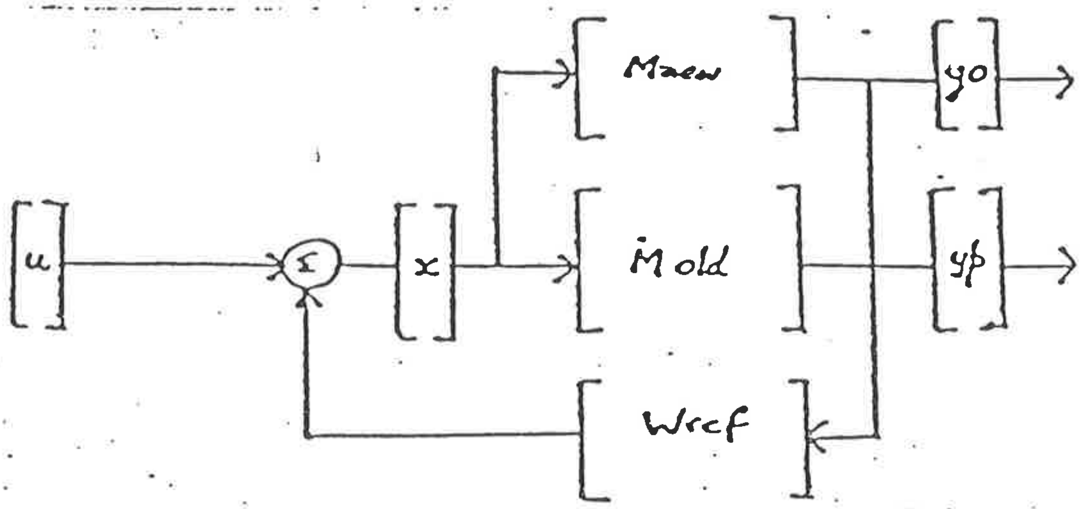


Fig 1 A general inconsistent controller

4. Results.

A desirable behaviour for malfunctioning plants is that, under the action of the controller, the plant eventually becomes wellfunctioning by modifying the input. We see that this proves possible. Among plants which are *not* eventually wellfunctioning, there can be defined several types.

Definition 4.

(a) A plant *cycles* iff *yobs*(t)=*yobs*(t+k) for some k and all t (or all t after an appropriate t_0).

(b) A plant is *persistent (after t)* iff *yobs*(t)=*yobs*(t+k) for all k>0.

(c) A plant is *bounded* iff no component of *yobs*(t) ever gets more than a fixed number k from zero.

(d) A plant may be defined operationally to *explode* iff some component of *yobs* exceeds a predetermined bound (in these studies it was taken as 10^5).

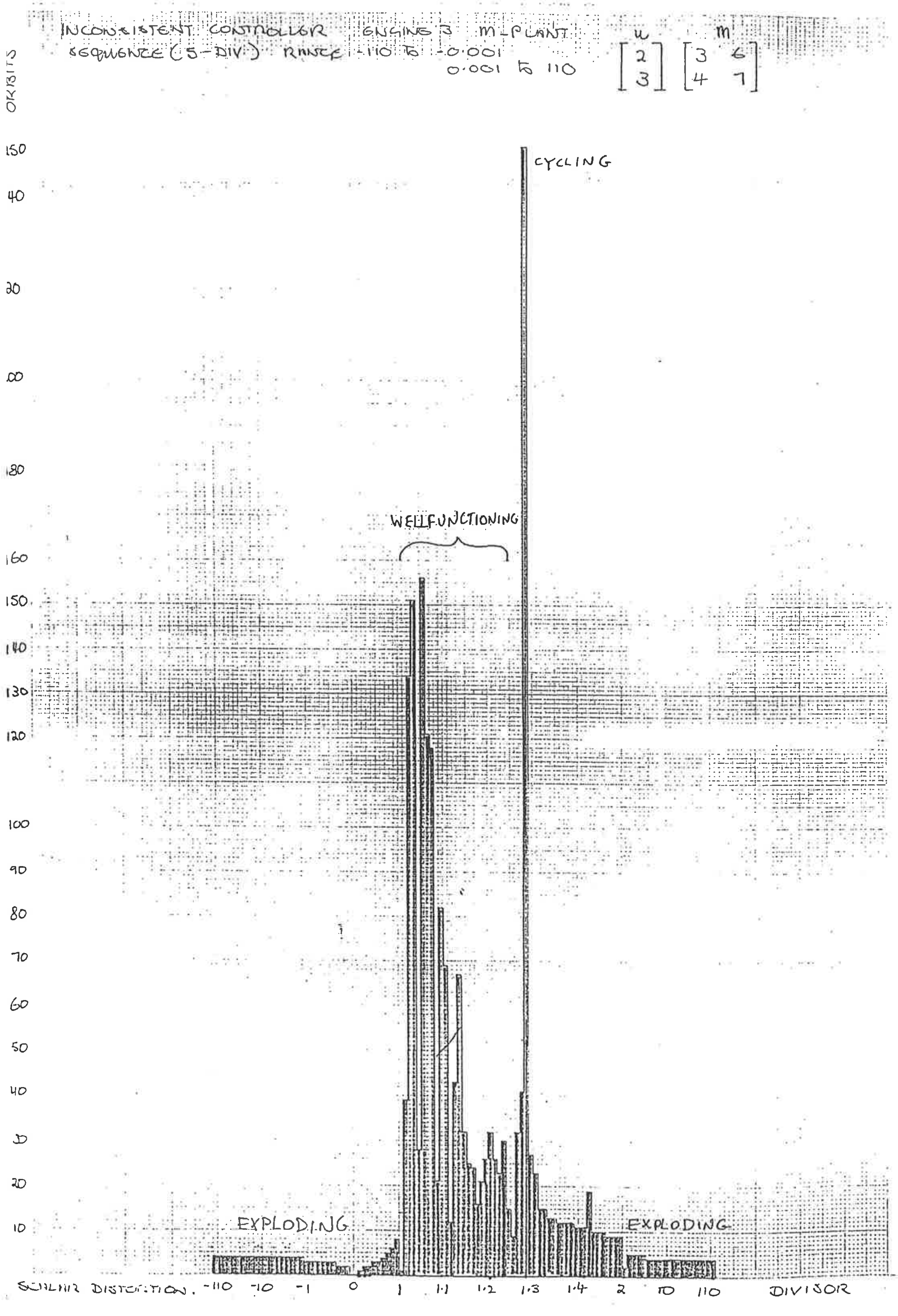
For a plant which does not eventually wellfunction, any of the behaviours (a)-(c) are more desirable than explosion.

Programs were written to simulate the controller described in Section 3. To facilitate inspection of large numbers of runs, M*old* was held fixed and M*new* varied systematically by applying a multiplier to one row. The size of M was kept low, to 2x2 or 3x3. The following data summarise the results of various runs with two different methods of adjustment of the state vector *x*. It is noted that the outcome of eventual wellfunctioning is in several cases achievable within a continuous range of variation of M*new*, which suggests a systematic effect.

In addition, the above behaviours of cycling and persistence have been observed. Boundedness cannot be directly observed, of course, but long runs have been observed without explosion.

Bibliography

Mortensen, Chris, *Inconsistent Mathematics*, Dordrecht, Kluwer Mathematics and Its Applications Series, 1995.

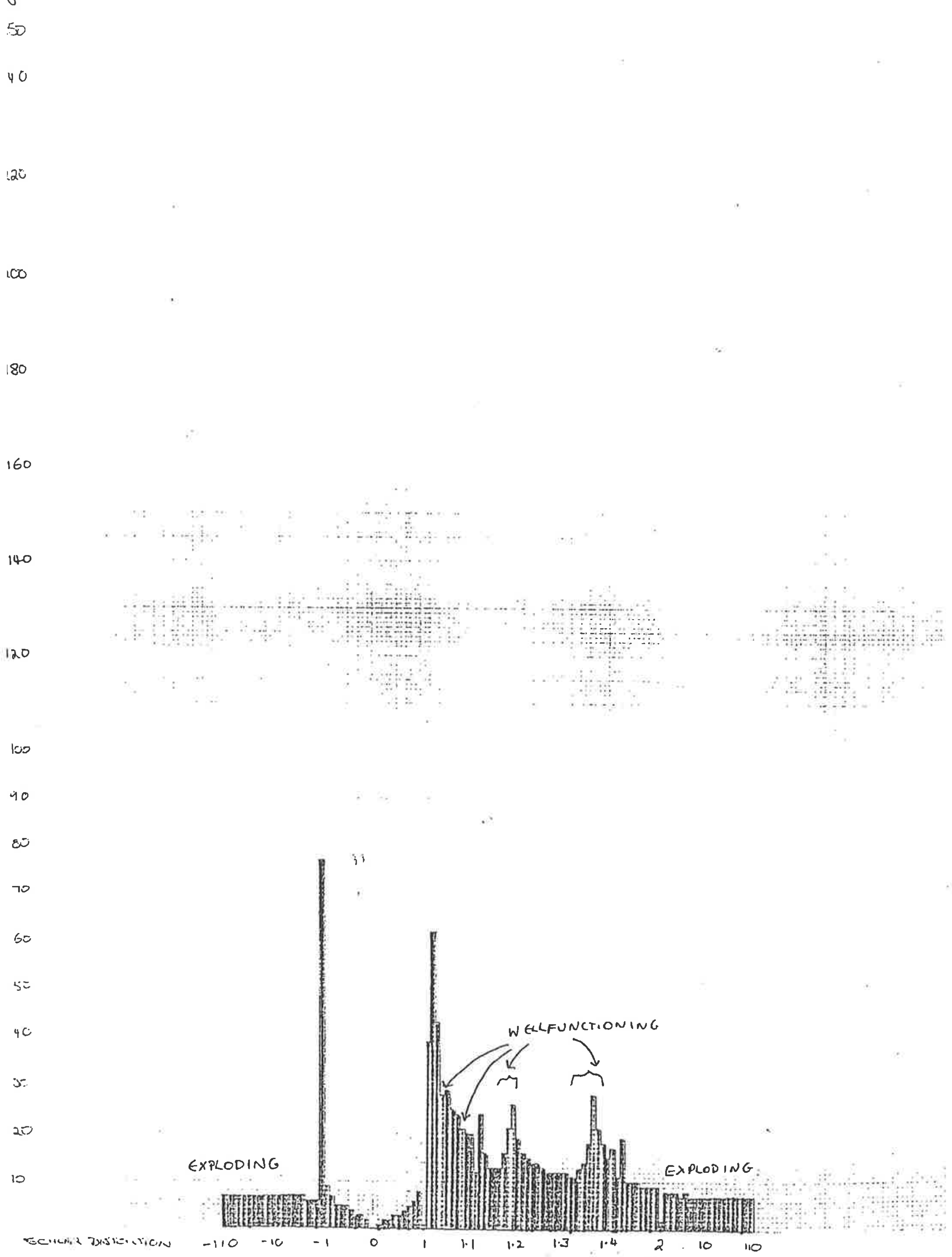


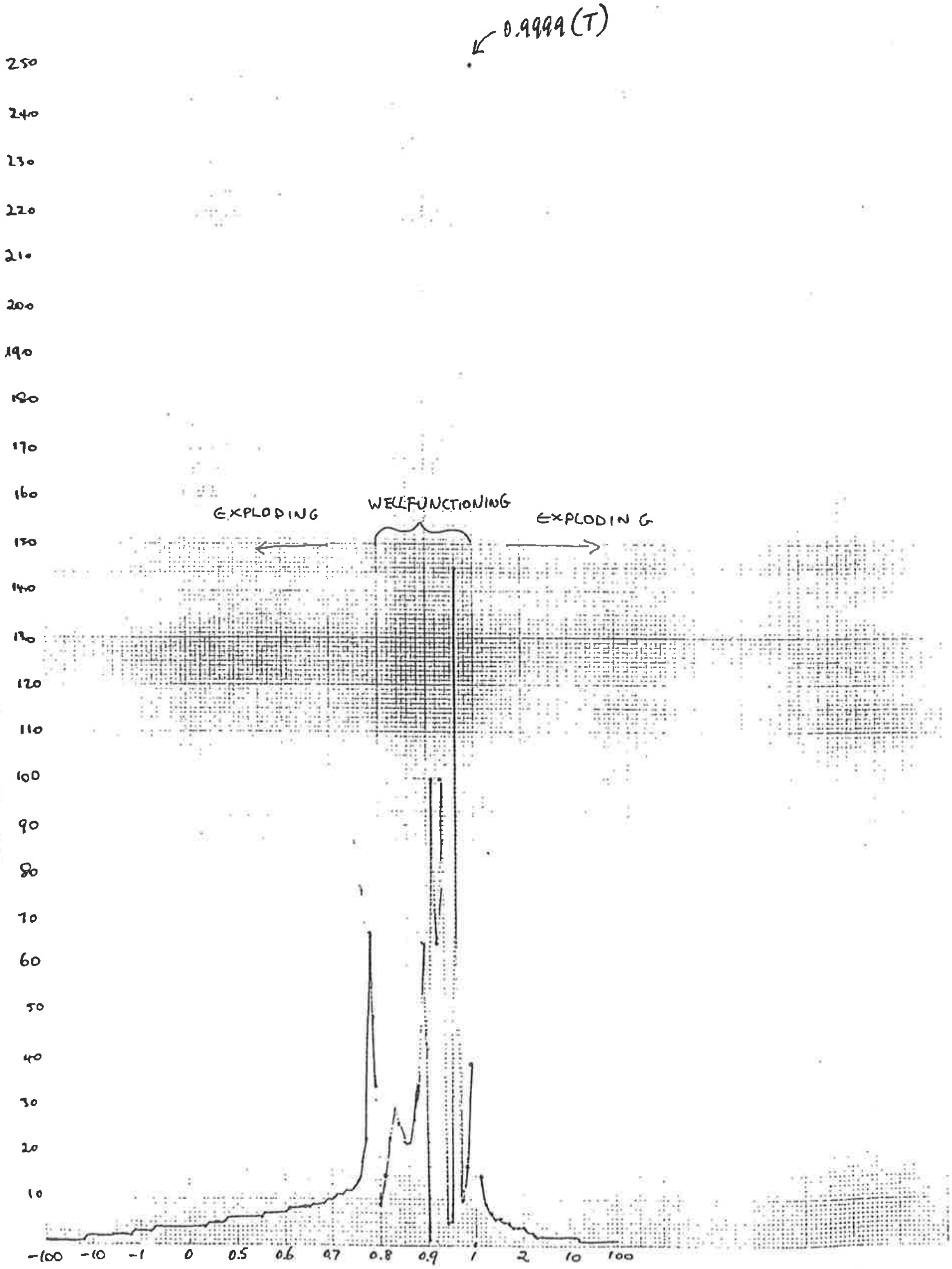
INCONSISTENT CONTROLLER ENGINE 3 M-PLANT
SEQUENCE (5-DIV) RANGE -110 to -0.001
0.001 to 110
u
[2 3]
m
[3 6
4 7]
ORBITS
CYCLING
WELLFUNCTIONING
EXPLODING
EXPLODING
160
150
140
130
120
100
90
80
70
60
50
40
20
10
SCALAR DISTORTION. -110 -10 -1 0 1 1.1 1.2 1.3 1.4 2 10 110
DIVISOR

INCONSISTENT CONTROLLER : SEQUENCE (S + DIV) RANGE −110 to −0.001
0.001 to 110

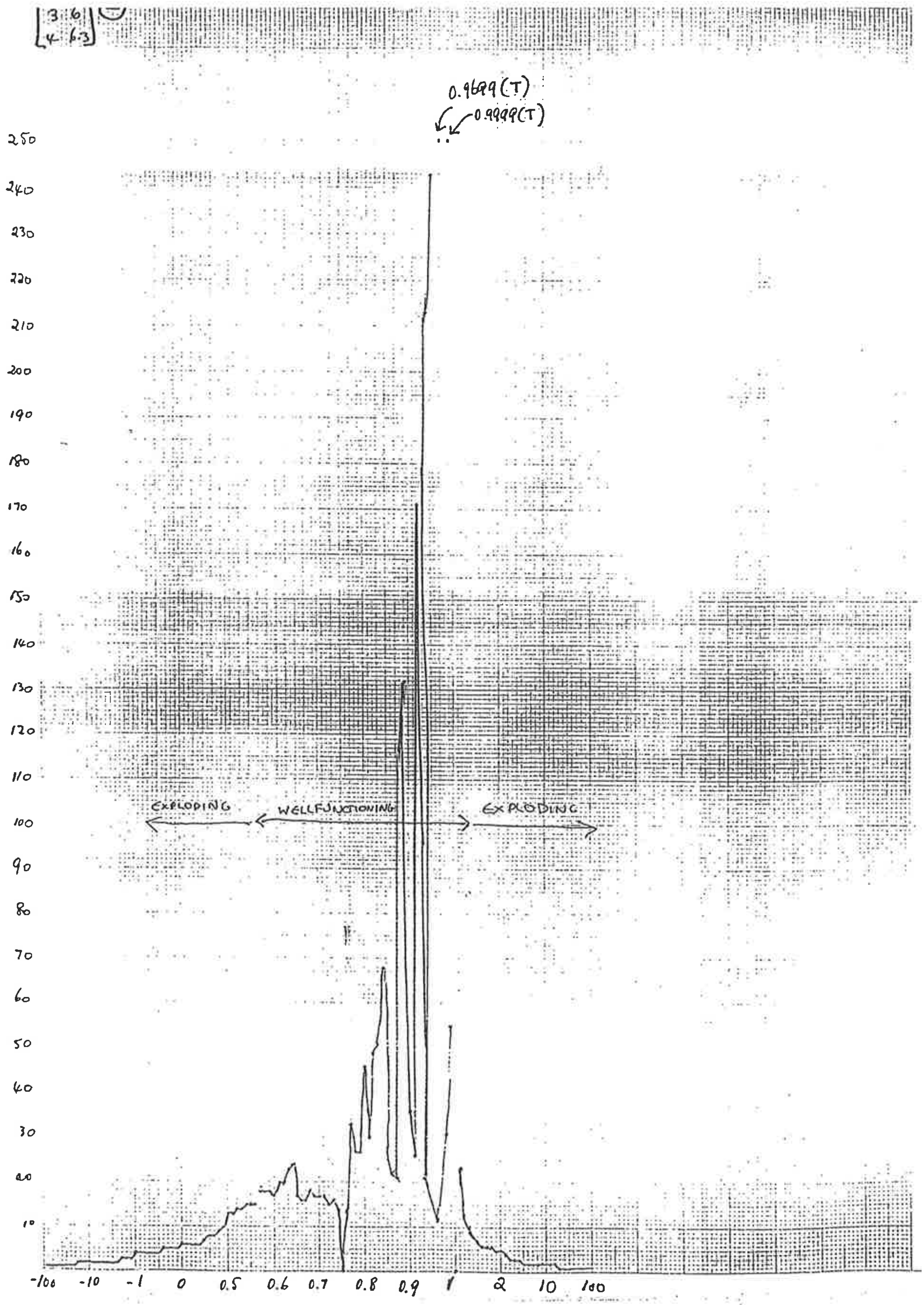
ENGINE 3

P_PLANT. u $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ m $\begin{bmatrix} 3 & 6 \\ 4 & 7 \end{bmatrix}$





0.9999 (T)
250
240
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
EXPLODING
WELLFUNCTIONING
EXPLODIN G
-100 -10 -1 0 0.5 0.6 0.7 0.8 0.9 1 2 10 100



0.9699 (T)
0.9999 (T)
250
240
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
EXPLODING
WELLFUNCTIONING
EXPLODING
-100
-10
-1
0
0.5
0.6
0.7
0.8
0.9
1
2
10
100

Fig 6 Mixed Persistence and Cycling

Mult	No of Orbits	Result
2.0	1	Persistence
1.9	1	P
1.8	2	P
1.7	3	P
1.6	3	P
1.5	4	P
1.4	2	Cycling (period 2)

"Inconsistent Images"

Chris Mortensen
Peter Quigley
Steve Leishman
John Mercier

2003-5

http://www.arts.adelaide.edu.au/humanities/philosophy/inconsistent_images/index.html

Only the front page is included here: other pages with animations and images by the co-authors can be accessed from the links on the front page.

School of Humanities

THE UNIVERSITY OF ADELAIDE AUSTRALIA

Inconsistent Images

Impossible pictures are so-named not because the two-dimensional picture itself is impossible (for then there would be no picture!), but because what is depicted, an apparently three-dimensional thing, is impossible. This raises a significant puzzle: how can it be that one is able to draw a picture of a thing which cannot exist because to do so would violate the laws of logic or mathematics? One surely cannot draw a picture of a standard contradiction, such as "Snow is white and snow is not white." But impossible pictures are different: there is an experience "I see it but I don't believe it" which cries out for explication.

The history of impossible pictures goes back to Pompeii. Later, there are isolated medieval altarpieces, Piranesi's Carceri contains some strange-looking stairs, and Marcel Duchamp drew a peculiar bed. However, impossible pictures were not drawn in any systematic way until Oscar Reutersvaard began his career in Stockholm in 1934, drawing over 4,000 pictures in the subsequent decades and being honoured by the Swedish government in the 1980s. M.C.Escher and the Penroses followed from 1955 onward, Escher in particular producing masterpieces such as Ascending and Descending, Belvedere, and Waterfall.

Impossible pictures should be distinguished from pictures which permit more than one gestalt, such as the duck-rabbit or the candlestick-faces. In the present study, these are classified as incomplete, not inconsistent. The property of incompleteness is a logical dual to inconsistency in more then one sense. Since this duality is well-known, this means that mathematical treatment of incomplete pictures is readily available once it has been worked out for impossible pictures. However, the latter is a hard problem, not yet solved satisfactorily.

A research project conducted in the Discipline of Philosophy aims to address the issue on several fronts. First, impossible pictures need to be described mathematically. This requires the tools of inconsistent mathematics and paraconsistent logic, that is logic which is tolerant of inconsistencies. The general idea is on viewing an impossible picture, the brain encodes an inconsistent theory. This is somewhat analogous to the way that the brain encodes projective geometry as a projection of a three-dimensional reality, except that the "virtual" three-dimensional reality is inconsistent. Clearly, this has connections with cognitive science: it is hardly being suggested that there is an inconsistent reality "out there", rather it is a matter of the brain's capacity to represent in an inconsistent fashion. A start has been made on the mathematics, but much more needs to be done. Second, in particular there is an issue of classification into various types here, types which seem not to be reducible to one another. These types ought to reflect different mathematical theories. Third, Reutersvaard's own program of drawing different pictures is being extended by Steve Leishman and others. Fourth, there is the prospect of virtual reality itself. There has been a conjecture by Bruno Ernst to the effect that one cannot rotate an impossible picture. This is now known to be false: Mortensen demonstrated this in principle at the 1999 Australasian Association for Logic Annual Conference (Melbourne), and Peter Quigley has now implemented it in detail, discovering more than one way of doing so with impossible Necker cubes. It is apparent that this animation is a preliminary to virtual reality, wherein one has the prospect of being able to wander through a whole impossible environment.

Peter Quigley has provided some further discussion and examples, including animations.
Steve Leishman has created a gallery of impossible pictures (best viewed with Internet Explorer 6, Netscape 7 or Firefox).
John Mercier has created another gallery of impossible pictures (thanks to Peter Quigley).

Chris Mortensen 2005