Some varieties of superparadox

Christopher Ormell

MAG-EDU: U E A

MAG Studies in the Meaning of Mathematics
ORMELL: Some Varieties of Superparadox ISBN 0 907669 22 0

Some Varieties of Superparadox

The implications of dynamic contradiction, the characteristic form of breakdown of breakdown of sense to which self-reference is prone

Christopher Ormell

MAG-EDU University of East Anglia

This book is the second volume in the MAG series *Studies in the Meaning of Mathematics*. Series Editor: Chris Ormell
Published by the Mathematics Applicable Group, School of Education, University of East Anglia, Norwich NR4 7TJ England in association with Ashby Anthologies

First published: 1993
ISBN 0 907669 22 0
Copyright c Maths Applicable Group 1993

C o n t e n t s

Lasermasters by Jarrold & Son. Printed in the Print Unit, U E A
COVER DESIGN: An impossible solid triangle ---see Ernst (1986)

P r e f a c e :   a summary of the argument

THE PROBLEM OF THE PARADOXES came to the fore in philosophy and mathematics with the discovery of Russell's Paradox in 1901.

It is the "forgotten" intellectual-scientific problem of the Twentieth Century, because for more than sixty years a pretence was maintained, by a consensus of logicians, that the problem had been "solved".

Culturally, it knocked much of the the stuffing out of rationality. Those who knew the true situation, knew that a concept which had been introduced as the very *embodiment* of mathematical clarity ---the concept of a set--- was flawed in some fundamental way which seemed utterly to defy analysis. You could not, most logicians thought in 1900, get a simpler, plainer, more transparent concept than that of a 'set'. Suddenly, the following year, it was revealed that this apparent epitome of logical perfection contained a baffling flaw!

The paradox "threw" the orthodox, platonic, logical thinking of the 1900s completely. A commonsense explanation of the paradox was required ---one which would show clearly and unmistakably that the paradox was the result of disregarding some *extremely obvious*, self-evident principle of logic. No such explanation was ever found.

But an "answer" of some sort was urgently needed, because set theory was in constant use throughout higher mathematics, particularly in Analysis, and in the theories of the continuum, higher geometry and topology.

The result was that pseudo-solutions came to be offered to the problem: first Russell's 'Theory of Types', and second 'Zermelo-Fraenkel Set Theory'. The former made a feeble attempt to offer itself as an "explanation" of Russell's paradox, while the latter simply accepted cynically that no solution could be found. So why not adopt *ad hoc* rules that would have the effect of *preventing the paradox from occuring* within formal systems?

The consensus view of the logicians of the 1900s could hardly have been less perceptive: it was that *self-reference* was to blame for the paradoxes! This was like blaming Lockerbie, the place, for "Lockerbie" the disaster. "Self-reference" was of course the location of the problem, hardly its "cause". But leading logicians soon convinced themselves that self-reference, in general, contained a vicious circularity. It is hardly surprising, then, that both the initial pseudo-solutions "worked" by finding "reasons" for blocking the self-membership of sets. So, too, most of the solutions which were offered to the problem of the paradoxes by subsequent commentators sought to find the root source of logical mischief in the very notions of self-reference and self-membership.

All such theories were absurd. "Reference" is an intrinsically elusive idea, and one can no more stop a statement ---such as this one--- from referring to itself than one can prevent the photons from a light source from falling back onto the lamp itself.

If you make a list of the things you must take with you to the airport, it is extremely sensible to put *the list itself* on this list! To deny such a truism is not even minimally credible.

But the logicians of the 1900s prefered to put their heads in the sand.

The problem of the paradoxes became for sixty years the forgotten problem of the Twentieth Century. No one wanted to re-open it, least of all the mathematicians, who learnt in the end how to "live with" Zermelo-Fraenkel set theory. Only after the collapse of "foundationalism" around 1972, did it begin to seem feasible to re-open these old wounds. But it has taken another twenty years for the sheer intellectual discomfort of trying to operate opportunistically, with a "foundationless" mathematics, to tell. Now, at last, I believe, there is sufficient awareness of the discomfort to justify the effort of laying-out a new beginning.

A NEW BEGINNING

A new beginning must start with a firm recognition of the principle that a statement is *presumed to be valid until shown to be invalid*. Once we adopt this principle, it becomes evident that most self-referential language is valid. (Appendix A consists of evidence for this thesis.) The problem is not with self-referential language in general, but with a small, peculiar, paradoxical *subset* of it.

The paradoxes, it becomes clear, are not caused by "self-reference" at all, but by "radical self-ascription", i.e. where a statement instructs us to make radical changes to itself. We can easily show that Russell's paradox falls under this principle, because the membership criterion for the ordinary set, O, (the set of sets which are not members of themselves) is that for any putative member, X, the statement "X belongs to X" should be false. So what about O? It now transpires that "O belongs to O" means that "O belongs to O" is false!

This is an example of "radical self-ascription": of a statement which claims that it, itself, is false.

Radical self-ascription, then, is like a circuit which is wired-up to switch itself into a new mode, or even *off*.

Taken literally, a radically self-ascriptive statement S tells us to *change the meaning of itself*, thus producing a new statement, S'. In the case of the paradoxes, S' is also radically self-ascriptive, and it (S') now instructs us to change its meaning back to S. We shall call a statement which does this a '*reflexively* radically self-ascriptive' statement.

S, it appears, implies S'. S', it appears, implies S.

At this point the orthodox logicians of the 1900s threw up their hands in horror, with the cry "A Paradox!". It was a panic reaction. They should have noticed that S was still self-ascriptive, and that it led back to S'... No stable conclusion had been reached. What had happened was that S led to S' which led to S which led to S' which led to S... and so on ad inf. There was no static "contradiction" of any kind, but an *oscillation of inconsistent meanings in time*. This is very like a contradiction, except that it occurs serially, in time. We may conveniently describe it as a "dynamic contradiction".

The states-of-meaning S and S' may be called 'partial meanings'. The oscillation which occurs is an oscillation of partial meaning.

Once we allow a statement to refer to itself ---and we have little choice in the matter, because large numbers of statements (especially statements in logic, mathematics and philosophy like this one) *do* successfully refer to themselves--- the possibility of "dynamic", or "serial", contradiction arises.

SUPERPARADOXES

The superparadoxes, which give this volume its title, are not a stupendously baffling new logical catastrophe, but simply more general, extended, somewhat artificial, examples of dynamic contradiction. They show, beyond any reasonable possibility of doubt, that reflexively radical self-ascription can be tailor-made to produce nominal oscillations of partial meaning round cycles of 3, 4, 5 ... *n* states. We can "wire-up" radically self-ascriptive statements to do such exotic oscillations, as easily as the more familiar on-off-on-off-on... oscillations of the natural paradoxes.

Nominally such "superparadoxes" are vastly more "paradoxical" than the natural paradoxes, because the number of formally inconsistent propositions we can derive from them increases ex-ponentially. Actually they "blow up" the very idea of the horror of paradoxicality, by showing that it is a wholly predictable consequence of reflexively radical self-ascription. The concept of dynamic contradiction, in a word, acts like some hormone weedkillers, which kill the weed by causing its growth to explode. But these "self-exploded weeds" are not more fearsome, more rank, more deadly, weeds than they were before the treatment: on the contrary, they soon wilt, because they have outgrown their strength.

The superparadoxes are in a similar case. They are not intrinsically particularly interesting as paradoxes, and indeed we may stiffle yawn when we meet the Bean Paradox, which is nominally more than three hundred times more paradoxical than Russell's Paradox. This yawn says it all. What would you expect? *Of course* you will get 1023 mutually inconsistent conclusions if you try to interpret such an ingenious reflexively radical self-ascriptive statement!

What we have done is to reduce paradoxicality, which in the 1900s generated all the potential anxiety of the computer viruses of the 1980s, to a quasi-mechanical principle. It is evident that the translation of this "mechanical solution", once widely digested, into logico-mathematical software will not be difficult. Indeed, simply to formalise the concept of 'partial meaning' will be half the trick. This will then lead to checks for circularity: looking for cases where successive partial meanings go into a closed, non-convergent, cycle. We shall need to build-in such "dynamic contradiction checks" into all the integrated, non-stratified, logico-mathematical software of the 21st Century.

So what is the "extremely obvious" principle which was contravened by Russell's Paradox? It is the same principle that was contravened by the Liar, namely, that a statement which claims that it, itself, is false, sets up an oscillation of partial meanings, a dynamic contradiction. The remedy is to avoid dynamic contradiction as carefully and as unobtrusively as we already avoid the more familiar, static variety. Everyone is aware that there is no need to lay out elaborate *reasons* why we should avoid ordinary contradictions: except, perhaps, to remark that they destroy, at a stroke, a person's credibility as speaker or author. The same brief, minimalist, commonsense "reason" will clearly suffice to warn-us-off contradictions of the dynamic variety.

PERSONAL NOTE I began the first draft of this monograph at Deakin University in Geelong, Australia in the last four months of 1990, and completed it (the first draft) at the University of British Columbia in the first four months of 1991. I would like to thank Nerida Ellerton, Ken Clements, Gaalen Erickson and David Robitaille for the invitations which made this possible.

CHRISTOPHER ORMELL

May 1993 Ashby St. Mary, Norwich

A note about notation in this monograph

This monograph contains no technical, i.e. formal, logic. It is offered as an unequivocally *philosophical* inquiry into the problem of the paradoxes of self-reference, not as a technical treatise on logic.

Normal quotation marks have been used throughout, both for so-called 'shriek' quotation, and for literal quotation of statements. Where a *word* has been referred-to, or defined, or linked with a concept/idea, single quotation marks have been used. Single quotation marks have also been used to indicate *sentences*, regarded as grammatically/syntactically correct sequences of words, *phrases* and *words*. (See Appendix B for an explanation of the conventions adopted vis-a-vis the words 'statement' and 'sentence'.)

CHAPTER ONE The problem of the paradoxes is a principal obstacle preventing formulation of a self-evident logical basis for mathematics

"EPIMENIDES the Cretan claimed that all Cretans were liars!" ---it was a joke, a verbal quip, a trick remark, which circulated round the dinner tables of antiquity. St Paul, we know, heard it, remembered it, and mentioned it obliquely in his *Epistle to Titus* [1]. It was called a 'paradox', a classification which placed it alongside the weightier conundrums of Zeno.

Of course a person may be unquestionably a "liar" without necessarily lying *all* the time. We can only suppose, therefore, that Epimenides was telling the truth on this occasion, though like other Cretans, he was prone to lie... So there was, after all, an interpretation of Epimenides' claim which made sense. Later commentators strengthened the paradox by referring to a different, anonymous liar, ('The Liar') who remarked confidently one day:

"I am lying".

This was altogether more convincing as a paradox. If the speaker was lying, what he was saying must be false. This meant that he was not lying. If, on the other hand, he was telling the truth, what he was saying must be true. This meant that he was lying. Yes, there was a "problem", but it was still essentially only a verbal quip.

Few people, probably, lost much sleep over the problem.[2] Most were not bothered by the logically ambiguous nature of the Liar's remark. It was patently simply a "trick with words", and one moreover, on which nothing else seemed to depend. Even if we conceded that the Liar's remark was irremediably ambiguous, that would hardly affect anything else, or any matter of substance, so the ambiguity would be localised onto a remark of zero practical significance.

This, broadly, was the state-of-play with regard to the Liar paradox prior to 1900. Logicians were at that time concerned with much more momentous questions, chiefly the problems that had to be tackled when one tried systematically to introduce the idea of a *set* into the language of mathematics. "Sets" had been introduced into mathematics by Boole,[3] Frege, Dedekind and Cantor. They were seen as a device which enabled mathematicians to talk with a new rigour and a new precision about the subject-matter of their science. Sets, in fact, were widely regarded as the very *epitome* of clarity: their charm resulted from the fact that they enabled one neatly and effectively to take the "psychology" out of thinking about mathematics and/or logic. Early theorists had often treated logic as being the study of the "Laws of Thought". But that seemed to imply that logic might depend on subjective and personal factors: after all, how well one *thought* depended on all sorts of things, like one's general health, one's mood, one's workload, even how many glasses of wine one had drunk that evening!

The great advantage of the concept of 'set' was that it seemed to embody an exemplary, clinical, objective precision. It was, surely, *the* supremely *transparent* idea at the root of mathematics!

There were elements which belonged to the set, and elements which did not. There was a "membership test" or "membership criterion". If an element satisfied this, it was *in*: if not, it was *out*. A set could be a member of itself. There was nothing odd about this: it simply meant that it satisfied its own membership criterion. It was no more questionable than a person making a list of the things he/she should take to the airport:

Passport
Money
Tickets
Folding umbrella
This list
Keys
Letter of introduction

The fact that 'this list' was a member of the list (i.e. itself) was no cause for panic: *of course* the list itself was needed, to check that everything was there! [4] Dedekind had been the first to notice that a set could, as it were, fold back onto itself, when he introduced the idea that an "infinite set" was a set with the property that it could be put into 1-to-1 correspondence with *part* (a proper subset) of itself.[5] This was neat. It seemed to enable one to treat infinite sets very coolly, very calmly, and without having to say anything whatever about infinity.

These advances were widely considered to be unmistakable signs of progress. Mathematics, it was confidently believed, would soon benefit immeasurably from the rigour introduced via this concept of a set. A great age of logically refined, logically enlightened, powerful, astonishing mathematics was about to dawn!

Cantor had introduced the principle that "actual infinity" must exist. It would be a contradiction only to recognise "potential infinity", he argued, because what could 'potentially...' mean in the total absence of actuality? [6] Cantor then went on to postulate an ascending ladder of higher and higher infinities (sets of transfinite sets) and after a while he realised that this system, on his own principles, ought to boast a top rung. This would consist of the "set of everything" and it would, of course, be a member of itself.

However, he soon became aware that such a "set of everything" would also have to contain the set composed of its own subsets: which would give it a different cardinality from itself! [7] This was clearly absurd. So Cantor reconciled himself to the fact that the very concept of a 'set of everything' would involve a contradiction.

Bertrand Russell, however, was initially unconvinced. He felt that there ought to be a "set of everything".[8] And he was investigating this "set of everything" in a low-key way in May 1901 when it occurred to him that there must be, at least, a set composed of all those rather *ordinary* sets which were *not* members of themselves. (Most sets, like most lists, are not members of themselves.) Then he asked himself whether this totality of all ordinary sets would be a member of itself. It was a reasonable question to ask. He expected a definite answer: a minor theorem, indeed, which might throw a little light on Cantor's argument about the "set of everything". What he got was unexpected: a paradox, no less.

If the totality of ordinary sets was *not* a member of itself, it would, after all, satisfy its own membership criterion: so it would *be* a member of itself! On the other hand, if it *were* a member of itself, it was certainly not "ordinary", and therefore it unquestionably did *not* belong to itself!

At first Russell thought that this paradox, though somewhat disconcerting, must have a simple explanation. Gradually, however, it began to dawn on him that it was much more serious than this. It was not simply "disconcerting", it was dynamite. And there appeared to be no simple way to explain what was wrong.

NEWS OF THE PARADOX SPREADS

Russell consulted colleagues: first Alfred Whitehead, his former supervisor and friend. But Whitehead could find no plausible explanation either. "Never glad, confident morning again..." was Whitehead's unhelpful, and as it turned-out, prescient, comment. [9] Then Russell wrote to Frege. Frege, too, could find no plausible explanation. It was, for him, a devastating blow. It came out of the blue at a comparatively late stage of Frege's career, when, after many disappointments, his work was at last gaining a little recognition. He wrote later:

> "A scientist can hardly meet with anything more undesirable than to have the foundation give way just as the work is finished. A letter from Mr Bertrand Russell put me in this position at the moment the work was nearly through the press".

Frege: *Fundamental Laws of Mathematics*, Vol.2.

Russell commented on Frege's reply to his letter:

"Frege was so disturbed by this contradiction that he gave up the attempt to deduce arithmetic from logic, to which, until then, his life had been mainly devoted." Russell: *My Philosophical Development* (1959) p. 58

The seriousness of the situation could not be disguised. A devastating contradiction had been unearthed at the very moment when it appeared that mathematics was successfully casting-out its previous reliance on uncertain reasoning. Russell decided that finding the root *explanation* of the paradox must be his prime objective: everything else must take second place. The logical security of the entire edifice of mathematics was now, he thought, at stake. For three years he struggled vainly with the problem.

"Every morning I would sit down before a blank sheet of paper. Throughout the day, with a brief interval for lunch, I would stare at the blank sheet. Often when evening came it was still empty... The two summers of 1903 and 1904 remain in my mind as a period of complete intellectual deadlock. It was clear to me that I could not get on without solving the contradictions, and I was determined that no difficulty should turn me aside from the completion of *Principia Mathematica*, but it seemed quite likely that the whole of the rest of my life might be consumed in looking at that blank sheet of paper." Russell: *Autobiography* (1967) p. 154

Then Russell made what he thought was a breakthrough. In his *Autobiography* he comments:

"In 1906 I discovered the Theory of Types" (p. 155)

Whether this should be credited as a "discovery" or an "invention" is, however, far from clear. Russell had, in effect, decided to retreat from working with the "free" or "open" conception of a set. Instead, sets would be henceforward regimented into "Types" ---roughly collections of logical entities with the same basic *level* of structure. ("Level" was determined by reference, so an expression which referred to an expression of level *m* became an expression of level *m*+1.)

In truth this "discovery" was not remotely like a proper "explanation" of the original paradox ---which was now fast becoming known as 'Russell's Paradox'. It was more like a global decision to abandon the original conception of mathematics as based on sets. Henceforward mathematics, Russell considered, would be based on stratified or restricted sets. The old "free" or "open" sets of 1901 were abandoned. He now felt that they provided an essentially paradox-prone, and therefore unworkable, basis for mathematics. Set theory of the original 1901 vintage, i.e. one unhampered by the new swingeing restrictions, would, from now onwards, be deliberately branded as "*naive* set theory".

It was, in a word, a retreat. The problem had not been solved, simply side-stepped. It should be said at this point that Russell's "side step" into the "doctrine of Types" was never very enthusiastically received, though Whitehead, Ramsey [10] and a few others went along with it for a while. One can see why, if one considers the "explanation" Russell offered via his "doctrine of Types" for the paradox of the Liar.[11] According to Russell the Liar is using a logically improper expression when he says

I am lying".

What he *should* say ---according to the doctrine of Types--- is something like this:

"I am making a statement of order *N* which is false".

He could even say something more specific than this, namely:

"I am making a statement of order 1 which is false".

But this does not make sense, because it contains a fundamental reference mistake: *no "statement of order 1" has been made.* What the Liar is now saying (3 lines above) consists of a statement *about* a statement of order 1, and therefore a statement of order 2. And it clearly does not matter what numeral is substituted for *N* in the first formulation: the statement actually made will count as a statement of order *N*+1. So the Liar has been forced into a position in which he is unable to comment on the untrustworthiness of his own statement without committing an ordinary reference mistake, i.e. referring to something incorrectly (something which does not exist).

Now it may be said at this point that this "explanation" of the Liar paradox was only, at best, an after-thought on Russell's part. He was, of course, primarily and mainly concerned with the kind of contradictions which could occur inside formal logic if one permitted "open", i.e. wholly unstratified sets. But the fact remains that the "explanation" of the Liar paradox by means of the "doctrine of Types" outlined above is so defective that it would be hard to believe ---if we did not know that it had actually happened--- that Russell or anyone else could be blind to its inadequacy.[12]

It is quite obvious that the Liar ought to have objected strongly to the rule imposed upon him. Why should the Liar, in fact, show any respect for such an arbitrary rule? It is an attempt to use "the authority of logic" to legislate an unobvious embargo onto the use of ordinary language. But from whence could such "authority" derive? The whole point of logic is that its rules are totally transparent, or as Wittgenstein later put it, that it is impossible to think illogically. Logic has no "authority" to impose opaque, unobvious rules. Indeed the mere attempt to impose opaque rules onto the use of ordinary language is such a manifest misuse of the word 'logic' that it's claim to *be* "logic" evaporates immediately.

How, then, can we describe such an attempt to "explain" the paradox of the Liar, except as a pathetic manoeuvre to try to place an obstacle in the Liar's way? It has no more validity as an "explanation" than the story that the Liar was actually *trying* to say:

"I am lying down"! [13]

However, the "doctrine of Types" does have one credible characteristic: it offers a "general rule" which can be imposed *throughout* logic and ordinary language, rather than simply consisting of an arbitrary, preventitive device. It is, in a word, a "blinding-with-science" type of pseudo-explanation. It sounds somewhat like "science", rather than simply like an ad hoc obstacle-insertion procedure. But let's consider what this "science" entails. It tries to prevent, at a stroke, an infinite variety of harmless locutions. For example, it tries to forbid an entertainer, on arriving on stage, from saying (speaking into the microphone):

"I am making this statement to test the microphone".

This is clearly impermissible if we accept the "doctrine of Types". According to Russell he should have said:

"I am making this statement of order *N* to test the microphone".

And if we accept this injunction, we effectively turn the entertainer's attempted self-evidently valid statement into a mistake. We *make* it into a mistake. We also create an equivalence between the entertainer's original remark and the paradox of the Liar: both now appear to be mistakes of a similar kind...

Let's put it this way. The sheer weight of the superstructure of unobvious legislation imposed upon logic when we try to adopt the "doctrine of Types" is utterly disproportionate to the purpose of the device. It is as if we had a problem with a particular dog chasing its own tail. To solve it we are proposing to introduce a Nation-wide system of Dog Registration and a code of permitted ambulations...

The solution is absurd. The dog will continue to chase its own tail in spite of all this superstructure of attempted external legislation. No one would wish to deny that there is a useful general concept of 'logical Type'. Clearly a *fraction* is a totally different kind of thing from an *inequality*, a *matrix* is totally different from a *set*... But the "doctrine of Types", as enunciated by Russell, goes far beyond this unavoidable recognition of logical categories. It seeks, (a) to create a special system of logical Types stratified specifically by recursive reference, and (b) to impose this system rigidly throughout mathematics and logic, i.e. not to permit intermediates.

The puzzle is not whether such a proposal is unacceptable (it obviously is[14]) but why Russell ever imagined that it was a possible solution to the dilemma.

THE PANIC

We may distinguish five stages in the discovery of Russell's Paradox: (1) the logical problem, (2) the crisis, (3) the panic, (4) the cover-up, (5) the toothcomb.[15] That Russell and other equally eminent, famous logicians were thrown by the paradox is understandable. But the result was much worse than this. Soon the problem became a "crisis" ---a "crisis" vis-a-vis mathematics' general intellectual credibility. It was not so much the fact that there was a problem, as the fact that none of the chief protagonists could see any *hint* of a proper solution, or even the *lines* on which a proper solution might be sought. It soon became public knowledge that all the eminent logicians of the age (Frege, Peano, Hilbert, Weyl, Zermelo, etc. in addition to Russell and Whitehead) were impotent in face of the problem. Poincare, who had been sceptical about the value of the Frege-Russell approach from the beginning, had no constructive suggestion to make either. His comment was the impish remark that the Logistic approach "is no longer sterile, it begets contradictions!". It became evident, in other words, to intellectuals and academics in many other areas that there was a massive failure of understanding *of some sort* in the foundations of mathematics.

The fact that Russell had, by 1906, adopted the "doctrine of Types" says it all. The crisis finally got to him. Mathematics could not be left indefinitely hanging in the wind, its credibility utterly broken. Action was needed: something had to be *done*. But Russell's confidence, like that of mathematics, had been badly shattered. He describes how, after moving to Bagley Wood near Oxford in 1905, he used to pause on the railway bridge at Kennington every day and watch the trains go by.

> "[I] determin[ed] that tomorrow I would place myself under one of them. But when the morrow came I always found myself hoping that perhaps *Principia Mathematica* would be finished one day... So I persisted and in the end the work was finished, but my intellect never quite recovered from the strain." (p.155)

This was the background to Russell's "Theory of Types". It was essentially a fudge. It did not solve the problem it claimed to solve, but consisted instead of a panicky attempt to restore consistency to mathematics by trying to invoke "logical authority".

Russell and Whitehead's joint project *Principia Mathematica* was the logical outcome of the decision to try to restore credibility by blinding the "public" (i.e. in this case ordinary mathematicians) with science. There would be a major show of logico-mathematical muscle. This would subtly, even subliminally, enhance the "authority" needed to impose the doctrine of Types.[16] It was something Russell felt he had to do. After all, he had *created* the Crisis by discovering the paradox. It was up to *him* to do whatever was necessary to restore the credibility his paradox had so rudely shattered. He was the person ultimately responsible for the whole sorry mess.

When *Principia Mathematica* Vol. I appeared in 1914 it was obviously a massive piece of work: a dauntingly impressive monolith of utterly inpenetrable logical gobbledegook for most ordinary, philosophically interested, readers. But was it valid? How could you tell, without working through page after page of utterly mechanical, formalistic, charmless, diagramless mathematics?

It was supposed to "reduce mathematics to logic". If this meant anything at all it meant that it demonstrated *clearly* that the edifice of modern mathematics could be generated from a few *wholly self-evident*, wholly transparent principles. But this general, procedural, "philosophical" clarity was equally obviously missing, and the authors themselves admitted that they had had to introduce four unself-evident principles

The doctrine of Types	The axiom of choice
The axiom of infinity	The axiom of reducibility

to pull the trick! This was unfortunate. Supposing that the authors could have done the job by sheer logical willpower *without* these hostages to fortune, it would *still* have been difficult to convince the public that the derivation was valid. The "public" needed (if it were to believe full-heartedly in the enterprise) page after page on which matters previously thought difficult, obscure and abstruse, were made *simple*, *transparently clear* and *easy*. What it got was a massive tome which, on page after page, rendered things far more difficult, obscure and abstruse than they had been before. And, in addition to this cavalier display of reader-unfriendliness, there were the four highly questionable assumptions to swallow!

There was no leeway for giving the authors the benefit of the doubt, though many did. *Either* their arguments derived mathematics unmistakably from transparent logical principles, *or* they did not. Only in the former case would the damage created by the paradox be restored. Every page of *Principia Mathematica* says otherwise. There is no restoration of clarity and light. Instead, only a grim calculus churned mechanically, to deliver nominal "proofs" which could never even begin to meet Descartes' precept that every step in a demonstration in mathematics should be totally clear and distinct.

One is bound to feel some sympathy at this point for Russell. After labouring for ten years, chiefly out of a sense of duty ---and on a task he clearly detested--- his monumental lifework got a decidedly cool reception. Unlike the earlier *Principia* (to which a none-too subtle reference was being made), there was no immediate rallying round, no ready-made consensus, no instant talk of a "New Philosophy".

Closer inspection did little to quell initial doubts: and in the end very few logicians were sufficiently impressed to throw their weight behind it. The consensus view shifted, rather, to another fudge, that of Zermelo (1908), later developed by Fraenkel (1919), and subsequently called Zermelo-Fraenkel set theory.

But if Russell's fudge looked at least initially "scientific", there were no such scruples in evidence in Zermelo-Fraenkel set theory. It made no attempt to justify the restrictions it placed on the logic of set theory. This was another fudge, indeed one might call it a "brazen fudge", because it incorporated no gesture towards the susceptibilities of philosophy whatever, no attempt to dress-up the new rules as any kind of "solution". It was offered, rather, as a working compromise: a set of axioms for set theory which allowed mathematics to continue, albeit on a newly restricted, and evidently somewhat arbitrary, basis.

There is no point, today, so many years afterwards, in retrospectively attacking the *aphilosophical* quality of Zermelo-Fraenkel set theory. It was a conscious fudge, but its alibi was always the failure of the philosophers to offer anything better. Russell's "solution", the doctrine of Types, served as the crucial signal that no proper solution could be found. As the years passed, and still no hint of a proper solution appeared, the sheer pessimism of the situation became intolerable.

What had been defeated was the implicit claim of logico-mathematical rationality to be master of its own house, to understand its own logic ---to be thoroughly and consciously grounded in transparent, rigorous thinking. After 1901 this claim was broken-backed. It was quite clear that the logicians, the ultimate guardians of logico-mathematic rationality, did *not* understand their own logic. In those five years from 1901 to 1906 the previously enormous ---some might say overweening--- sense of educated-public confidence in Western "rational intellectual progress" ebbed insensibly away. (Classical physics, too, incidentally, was shattered by the discoveries of Planck and Einstein during the same years.)

To deal with this pervasive pessimism, a new stage in the saga of Russell's paradox emerged, that of the Cover-up. After a while it became the established norm that no discussion of the problem would be encouraged. Why re-open old wounds? Why wallow in a painful intellectual pessimism, to which there is no cure? Discussion effectively ceased. "What problem?" became the watchword of logicians [17] for sixty years, as they whistled in the dark, and tried to pretend that the defeat of 1901 had never happened.

THE SECOND CRISIS IN MATHEMATICS

Unfortunately the "cover-up" was too successful. By the 1960s most people had actually *forgotten* that the highly formalistic "modern mathematics" which followed *Principia Mathematica* (and later, in the 1930s, the *Bourbaki*) was founded on a fudge.

The first digital electronic computer ENIAC appeared during the War, but it was not until the late 1950s that computers became reliable. This resulted from the use of semi-conductor logic devices (transistors) in place of the earlier thermionic valves.

Suddenly, around 1960, it began to dawn on large numbers of intelligent people that the computer would transform human life, that a "second industrial revolution", no less, was at hand. They looked, therefore, for the "new thinking" which lay behind the new machine. And they saw, or thought they saw, *modern mathematics*! Here was a machine which functioned on a basis of truth-function theory (introduced in *Principia Mathematica*), which employed binary arithmetic, and all kinds of unfamiliar digital-logical algorithms. The conclusion was irresistible: that the highly abstract, formalistic concepts of modern mathematics were the *conceptual kernel* needed to understand both the new digital technology and the new world it would usher-in.

It was but a short step from this realisation to the conclusion that mathematics education was hopelessly out-of-date, and that a new regime based on the core concepts of modern mathematics ought to replace it in colleges and schools. Thus the movement, which came to be known as "modern mathematics for schools", was born. It was aided by the shock of the first Russian Sputnik in 1957. (Sergei Korolyev, the Russian engineering genius, we now know, was behind this amazing technological feat.) The effect of Sputnik I in the West was quite dramatic: it produced a near tidal-wave of dismay and disbelief. Self doubt swept across the United States. What had happened? What had gone wrong with American education? It was this political shock which unleashed millions of U.S. government dollars, and which effectively locked the world into a disastrous educational mistake which lasted in some places for twenty years

With the arrival of "modern mathematics for schools" an enormous new interface was opened-up between mathematics, viewed-in-a-formalistic-fashion, and the thought processes of the man and woman in the street. Suddenly teachers found themselves trying to "make sense of" the new attitude towards mathematics, an attitude dominated by awe of symbolism, intense (i.e. blind) deference to formalistic pedantry, and the trappings of rigour. To "make sense of it" would be to *relate* it, naturally and organically, to the kind of assumptions about reality, the kind of assumptions about deduction, and the purposes which were regarded as "real", by the ordinary family. It was never "on". There was no way in which the children of the entire human race could be initiated, willingly, into such a tensioned, attenuated, hierarchical, mandarin culture.

The result, then, was not in doubt. After riding high initially on its extravagant prospectus ---which billed it as an irresistible, irreversible, dazzling reform--- "modern mathematics for schools" promptly crashed. Once the reality of what it proposed became perfectly *clear* to ordinary children, ordinary parents, and teachers, its credibility simply disappeared.

Teachers found themselves trying to "sell" a stilted, artificial, triumphalist, apparently purposeless, "language" in noisy classrooms. One scheme actually wrote-in transfinite set theory, no less, as part of the main mathematics course for ordinary pupils up to 16! Some schemes laid emphasis on the fact that there was not a single concept of zero, but separate concepts for natural numbers, integers, rationals, reals, etc! It is hard to get children interested in zero, as an idea, in the first place. To try to get them to take on board such fine distinctions was pedagogic suicide.

And, let's face it, there was a direct connection between this stilted, authoritarian, Marie Antoinette-like cognitive style, and those fudged solutions of the 1900s. It was no accident that modern mathematics had drifted so far from the thought-patterns of the masses, so far from commonsense. If one bases a whole enterprise (modern mathematics) on a brazen denial of commonsense, one should hardly be surprised that the defiance implicit in that act shows through at every point.

The twentieth century, we know, has been ---in Wendell Wilkie's famous phrase--- "the century of the common man". There was no way in which the children of the "common man" would swallow a conception of mathematics inbued subliminally at every point with the hauteur of an aristocrat (Russell) who happened to have been born a hundred years out of his time.

Now it would have been somewhat surprising if a cultural setback of this magnitude had had no further repercussions. It did, in fact, have repercussions: the chief of which was that it effectively dragged a whole culture ---which might be described as that of "aggressive modern mathematics"--- into oblivion. It destroyed the credibility of that prestigious pressure group (consisting mainly of certain modernist university mathematicians) which had promoted it with an ill-advised and sometimes near-fanatic zeal. Mathematics was, once again, in disarray. The all-embracing cognitive artificiality which sprang ultimately from those fudged "solutions" to the problem of the paradoxes, had at last been exposed for all to see. This was the Second Crisis in Mathematics, or if you prefer, a resumption of the First Crisis after a lull of sixty years. Foundationalism had collapsed, like the cathedral at Beauvais in 1284, to leave a heap of rubble . ("No [Gothic] structure", Bronowski comments, "as tall as this was ever attempted again".[18])

A RETROSPECTIVE EVALUATION

We are now (1992/3) suffering from the kind of long-term, on-going, fundamental uncertainty which the Second Crisis in mathematics has left in its train. Foundationalism has collapsed, but nothing substantial has been put in its place. We appear to be careering along in an essentially opportunist way without any "fundamental logical basis" on which mathematics can rest.

Can't we "get by" without a "fundamental basis"? The answer is that we *can* "get by", to a degree, but that the education of our students is suffering insensibly from the pervasive ambiguity and the uncertainty which stems from ---in ordinary language--- "not knowing where we are, or what we are at". Mathematics is, or should be, the primary source of intellectual confidence in society, because in mathematics we can *fully* and*wholly* explain certain puzzling phenomena.[19] Mathematics thus provides a kind of final yardstick for what counts as a "full and thorough explanation" of a puzzling fact. But mathematics will only serve as a fount of clarity and intellectual confidence when it, itself, is free from the miasma implied by a thoroughly confused and uncertain basis.

Should we, then, go back to the "problem of the foundations", i.e. the problem of the paradoxes, and adopt another fudged solution? The answer is obvious. Fudges will not do the trick. We need *transparency* in the foundations of mathematics: there is no substitute for a proper solution to the problem of the foundations. We need, rather, to go back to the drawing board and actually to *solve* the problem of the foundations which baffled those eminent logicians of long ago.

It does not follow from this argument that we need new impressive formalistic edifices like those of *Principia Mathematica* or the system expounded by the *Bourbaki* (i.e. "formal foundations") to serve as a "fundamental basis" for mathematics. Indeed the *opp*ressive effect of such "*imp*ressive formalistic edifices" (sic) has probably done much, during the present century, to harm the spirit of mathematics, and to blunt the motivation of promising students. With the advantages of hindsight we can now see that the strategy of restoring confidence to mathematics by "blinding ordinary mathematicians with science" was always going to be riskily counter-productive. (Because the final role of mathematics is to increase and enhance the clarity and intellectual confidence, with which we *see* the world, not to "blind" the eyes of those who peer trustingly at it.)

But mathematics, I submit, *does* need some sort of "fundamental basis", or as I have elsewhere expressed it[20], it needs a "version of maximum rigour" (VMR) against which to compare new proofs and new conjectures. Given the disrepute into which the notion of "foundations" has fallen, we may however prefer to call the required basis a "clarification of the roots of mathematics", rather than any kind of reformed "foundation".

Returning to look at the fundamental problem of the paradoxes is of course only a part of such an inquiry. (I distinguish a renewed look at the "fundamental problem" ---i.e. one at the level of generality of the original inquiry--- from the recent "toothcomb" phase of minute research on the problem which has followed the collapse of foundationalism. See Note [15].)

The greatest obstacle to a renewed surge of interest in the fundamental problem of the paradoxes is, in my opinion, the legacy of the cover-up. Nothing is more demotivating to such a study than the word which was put around for more than sixty years that "the paradoxes don't really matter", that they constitute a "trivial problem anyway" and one "which can be solved in all sorts of ways, if that is what you want". We need constantly to remind ourselves that this sort of defensive, defeatist rhetoric originated when the hope of finding any kind of genuine solution of the problem appeared to be non-existent.

It is necessary to insist, therefore, that the paradoxes *do* matter, that the problem is *not* "trivial" and that the multiplicity of solutions implied by the final remark above is a multiplicity of fudged, obstructive *so-called* "*solutions*" not a multiplicity of genuine solutions.

PARADOXES MATTER

Let's consider the human effect of a contradiction in any field of human discourse. It is not simply that two different statements have been made about the same matter (usually in different contexts and within different bodies of speech) which have the unfortunate effect of cancelling each other out. A man says that he bought a bottle of champagne on his way home from the country. Later he changes his story and says that he bought three bottles of wine. Does it matter? Can we not simply leave the issue in decent obscurity, and regard his two contradictory accounts as stultifying each other, the second being treated rather like hitting the "delete" key on a computer?[21] Some indulgent commentators may say this, but anyone with an acute ear for consistency will react very differently. No, a trick has been perpetrated on us, his listeners. What does he expect, that we will *not notice* that on one of the two occasions he was *lying* to us? Our confidence in the speaker is shaken to the core: his credibility is now in question, not simply on this trivial matter, but on many others too.

What has been shattered is the speaker's linguistic honour: he has been caught-out in a clear-cut, unmistakable, serious act of linguistic deception.

If a contradiction, then, is by custom and precedent, viewed in such a concerned light in trivial, everyday contexts, how much more so will it be viewed in a theoretical or policy context? Here the troublesome, discordant consequences of the act are geared-up by the fact that theories and policies are generative of endless derivations... We are now faced, not simply with a pair of mutually stultifying utterances, but with two growing, open-ended, *families* of mutually conflicting implications...

Now let's look at the case of paradox. A "paradox" of course is more than just a stray, passing contradiction. It is a linguistic configuration which seems to *compel* contradiction. It represents a situation where, whether we assent to, or reject, the leading proposition, we are nevertheless *trapped* into a contradiction. There is no escape: we are faced with a contradiction whatever assumption we make about the situation. Now a paradox in any area of *science* or *theory* is worse, by the same token, because it seems to compel open-ended families of contradictions...

So if, as we have seen, a paradox within an area of theory must be treated as a matter of extreme seriousness, then a paradox *in logic itself* ---the very science which sets itself the deliberate, conscious task of thinking clearly--- this must be, of course, considerably worse.

But now let's assume that a new, light, precise, clarifying concept has recently been introduced into logic, with the object of increasing its overall illumination and improving its deductive powers. To discover a paradox lurking there at the heart of such a concept (i.e. that of a set) is a veritable catastrophe!

This sequence of reflections may perhaps convey something of the initially catastrophic impact of Russell's Paradox: seen, that is, as it *was* undoubtedly initially seen, from an unsullied, innocent, pre-1901 viewpoint, totally unaccustomed to the presence of paradox.

From our viewpoint of today, of course, it requires an effort of considerable imagination to recapture that initial moment of intellectual disintegration. We have all become thoroughly familiar with, perhaps even acclimatised to, the presence of paradox. We have come, I think, to take the presence of a kind of *scattiness* for granted. And we have also been assured endlessly for nearly a century that "there is no problem" over this. If, however, we are ever to "tune in" to the problem of the paradoxes sufficiently sharply to be able to recognise the difference between a proper solution and a fudge, our first step must be to throw off this blase attitude. We must try, in a word, to recapture the full enormity of that logical shock which broke over Russell and his friends one day in May 1901.

A MULTIPLICITY OF OBSTRUCTIVE "SOLUTIONS"

Yes, there have been many suggested so-called "solutions" to the problem of the paradoxes based on the principle of obstruction. A huge number of excuses has been found for putting a general embargo on self-referential language. By introducing 'Types' (Russell) 'metalanguages' (Tarski) 'Namely-riders' (Ryle) and many similar tricks... it is possible to *prevent* the occurrence of self-referential language, and hence to prevent the occurence of the paradoxes.

But to "achieve" this is not to *solve* the problem of the paradoxes.

Obstructive tactics of this kind no more "solve" the problem of the paradoxes, than locking up all the suspected criminals in a town "solves" the problem of a particular crime.

The "problem", in a word, is not to *eliminate* the paradoxes, but to *explain* them. This point is absolutely pivotal to what follows. The solution to which our inquiry points has the perverse result of revealing a new, hitherto unsuspected, genre of superparadoxes: statements with considerably more implicit, built-in compulsion-to-contradictoriness than the original paradoxes. That it reveals this new stratum of superparadoxes is obviously a powerful argument in its favour. If, however, "the problem" were simply to eliminate paradoxes from our language, that telling sign of genuine explanatory power might look more like an "exacerbation of the problem" than a "solution".

Of course the last thing we want is to find paradoxical statements popping up and circulating undetected among the propositions of philosophy, logic or ordinary language. The crucial word here is 'undetected'. A genuine explanation of the paradoxes will enable us to *pre-label* them clearly and unmistakably: thus completely defusing the problem of their presence. And an explanation which reveals a new, unexpected genre of paradoxes obviously has more "pre-labelling power" than one which does not.

Let's get to the heart of the central issue. The paradoxes represent an acute embarassment for logic, because it appears that, by the use of unexceptionable reasoning ---and at a level of logic where it was imagined that clarity was guaranteed--- we can find ourselves in a situation in which we are trapped into contradiction. This is what a "paradox" consists of. It follows that some of that apparently "unexceptionable reasoning" must have been invalid reasoning. The challenge is to say what: and to explain *why*, contrary to appearances, it is invalid.

RUSSELL's RE-EVALUATION OF THE PROBLEM

It may be noted that Russell returned briefly to reflect on the "problem of the paradoxes" in his late work *My Philosophical Development* (1959). The personal controversies, the conflicts, the painful dilemmas of that First Crisis in mathematics were now, presumably, a distant memory. Russell was able at last, from the vantage point created by old age, and more than half a century of thought and reflection, to take a calm, settled, dispassionate view of the seemingly insoluble problem with which he had grappled so painfully and conscientiously during the mid 1900s.

It is worth quoting his final thoughts, therefore, on the fundamental criteria which, he considered, a solution to the problem ought to satisfy:

"While I was looking for a solution, it seemed to me that there were three requisites if the solution was to be wholly satisfying. The first of these, which was absolutely imperative, was that the contradictions should disappear. The second, which was highly desirable, though not logically compulsive, was that the solution should leave as much of mathematics intact as possible. The third, which is difficult to state precisely, was that the solution should, on reflection, appeal to what may be called 'logical common sense' ---i.e. that it should seem, in the end, just what one ought to have expected all along.

Russell: *My Philosophical Development* (1959) p.61

This is an interesting passage for several reasons. In the first place the one can still feel the sense of urgency Russell is placing on the need to *remove* the contradictions. (He means the paradox about the set of all ordinary sets, and the family of derivative formal variations of it.) The striking thing about this is that, even after more than fifty years have passed, and a range of workable ad hoc operational devices have long since been in action to prevent the occurrence of the contradictions, Russell is still reacting somewhat like a scalded cat. He does not seem to have taken any comfort from the existence of Zermelo-Fraenkel set theory, the main lesson of which is surely that "removing the contradictions" ---as a purely technical matter--- can be quite easily accomplished.

In the second place, Russell is still placing most of his emphasis on finding a solution satisfactory to *mathematics*. When he comes to the question of the *logic* of the exercise his "requisite" is quite vague. He does not even entertain the possibility that so-called solutions to the problem might be harmful to logic.

Russell's requisites one and two, in a word, still embody a tincture of the sense of panic which did so much to mar the quality of all the early so-called 'solutions'.

Russell's requisite three, on the other hand, is a step in the right direction. (One suspects that requisite three is the fruit of later reflection, rather than a literal account of how he approached the problem in 1903/4.) The trouble is, however, that it does not go sufficiently far in the right direction. Of course a solution, to be satisfactory, must "embody logical commonsense", and of course it should seem in the end to have just the quality "one ought to have expected all along". It should, in a word, seem to be "the natural solution" and should seem to be "self-evidently valid". But what Russell conspicuously does *not* say at this point is that it should place a ban only on forms of speech/argument which are actually and visibly "contradictory". To try to ban more than this would be to harm the integrity of logic, because it would be to treat the authority of logic as if it were controllable by whim. Russell seems to have had no stomach to consider ---still less to accept--- even the *possibility* that he might have to work with such a modestly transparent and unauthoritarian logic.

There is no indication indeed that Russell, even momentarily, considered this point of view. He was so wrapped-up in defending the consistency and the estate of mathematics (including its Cantorian provinces), that he seemed entirely to overlook the fact that he might easily end-up trying to stretch logic beyond its proper role, and hence damage it.

Later on the same page Russell comments on the applications of his "third requisite" to some of the more ingenious restrictive so-called 'solutions':

> "Professor Quine, for example, has produced systems which I admire greatly on account of their skill, but which I cannot feel to be satisfactory because they seem to be created *ad hoc* and not to be such as even the cleverest logician would have thought of if he had not known of the contradictions."

Since Quine produced his minimally restrictive version of axiomatic set theory long after the original "Crisis" years, we may deduce at once that Russell has unconsciously foreshortened the timescale. (It is evident that exactly the same comment could have been made about Russell's own Axiom of Reducibility, which *turned out* to be needed to take the sting out of his "doctrine of Types".)

But in spite of all these qualifications, Russell's "third requisite" marks, I believe, a step forward.

Yes, it was the flouting of Russell's third requisite which was what was fundamentally wrong with all those so-called "solutions" of the 1900s and later years. Russell's third requisite does, once fully unwrapped, say a lot. We *do* need a solution whch satisfies logical commonsense and "which seems, in the end, just what one ought to have expected all along".

NOTES

[1]"It was a Cretan prophet, one of their own countrymen, who said 'Cretans were ever liars, vicious brutes, lazy gluttons' --- and how truly he spoke!" Epistle to Titus, I, xii-xiii. *The Revised English Bible*, Oxford, (1989) New Testament, p. 194.

[2] Though Tarski (1969) claimed that Philetas of Cos was tormented by it and that it caused his premature death. See Sainsbury (1988) p. 1.

[3] George Boole (1854). E. T. Bell (1937) comments "As Bertrand Russell remarked some years ago, pure mathematics was discovered by George Boole in his work *The Laws of Thought*..." (p. 479).

[4] This is a crucial example, which is absolutely pivotal in understanding the problem of the paradoxes. No "solution" can be worth the name unless it beards the lion (the paradox) in its most powerful form. (It is no use trying out a new drug for attacking bacterium X on weakened or attenuated strains of X. Only if the new drug can be seen to destroy the bacterium in its most virulent form will it count as a success.) So considerable care is needed in first formulating the most formidable variants and/or interpretations of the paradoxes. This is the principle that one cannot explain a paradox satisfactorily unless one explains the "most difficult interpretation" of the paradox.

The example shows that a set can be a member of itself in the most practical, "ordinary" context imaginable. This immediately wrongfoots most of the celebrated attempts to "solve" the paradox ---which have operated by finding a supposed "category mistake" in the concept of a set being a member of itself.

It might be thought that the concept of a 'list' is one of much less generality than that of a set. However this judgement rests on the assumption that the notion of transfinite sets is valid. There are, it now appears, compelling reasons against this, because the central idea of a 'set' is that one is naming *all* the so-and-sos. The set of Xs is essentially the totality of "all Xs". But we certainly cannot claim to have conceptualised *all* the Xs in transfinite cases like that of the irrational numbers. On the contrary, we know that the totality of irrational numbers is a systematically elastic one, so that no final audit of "all the irrational numbers" is possible, even in principle. This suggests the device which the author introduced in 1984 of calling transfinite totalities SETs ('systematically elastic totalities') and restricting the use of the term 'set' to finite and countable totalities. In the countable case, although we cannot actually write down the names of all the elements, we can write down all the *main* ones (the early ones in the sequence, having first ranked them in order of importance) and we can subsequently name the more minor ones down to as low a level of priority as you like... This is, effectively, the situation which obtains vis-a-vis empirical sets like "the set of all insects in Norfolk" or "the set of all the stars in the sky", which, no one seriously disputes, are bone fide sets.

Given this basis, it appears that all genuine sets are countable, and that their elements are therefore capable of being listed! (In a great many alternative ways obviously.)

[5] Thus the natural numbers can be put into 1-to-1 correspondence with the self-powered primes:

$$\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \ldots \\ 1 & 2^2 & 3^3 & 5^5 & 7^7 & 11^{11} & 13^{13} & 17^{17} \ldots \end{matrix}$$

in the sense that for each member of the top sequence there is a unique member of the lower sequence and vice versa. This correspondence occurs in spite of the fact that the first 823, 543 members of the top set only contain five of the bottom set!

It is evident that this property is not possessed by any finite set. It therefore enabled Dedekind to define infinite sets, in effect, without having to say anything about the logical status of infinity: an achievement much-prized at the time by those who were either conscious of the difficulty of defining the logical status of infinity from a commonsense basis, or like Cantor, were unwilling to do so.

[6] This was like the concept of 'unrequited love'. How could one make sense of such a concept *unless* one had first formed the concept, and found examples of, 'requited love'? The argument is, of course, unsound. Let's face it, 'potential infinity' is a *metaphor* likening the operation of increasing a number to the operation of approaching something huge: it no more guarantees the existence of the "something huge" than the object highlighted in any metaphor is guaranteed by its mention.

[7] The set composed of the totality of all its possible subsets is called the 'power set'. Cantor proved ---what is in fact just a generalisation of his famous "diagonal argument"--- that a set could not be put into 1-to-1 correspondence with its own power set.

[8] And one must agree that Cantor's paradox completely contradicted the principle of "actual infinity" on which the whole theory of the transfinite ultimately rested. (Note that Russell still used the old-fashioned term 'class' rather than 'set' at that time.)

[9] See Ronald Clark (1981) p.27.

[10] In the Preface to the *Principles of Mathematics* 2nd Edn, Russell is loud in his praise of Ramsey's re-worked version of the Theory of Types. The theory, he comments,

"...as it emerges from Ramsey's discussion, ceases wholly to appear unplausible or artificial or a mere *ad hoc* hypothesis designed to avoid the contradictions." (p. xiv)

Russell then gives an example in which a propositional function *Øx* becomes obviously nonsensical when an *x* of the wrong logical character is inserted. Russell and Ramsey were so eager to find "a solution" to the problem of the paradox, that they seemed to be determined to overlook the plain fact that such a category mistake is palpably *not* present in the paradox. They also overlooked obvious examples like the Airport List example (Note [4]). Classes *are* just the kind of things which *can* be members of classes, indeed in mathematics most examples of classes have other classes as their members. (See Moorcroft 1993 p. 101: "Any set theory worth the name requires, almost from the outset, taking sets to be members of other sets.")

[11] See Russell (1962).

[12] After applying his Doctrine of Types in the form given in the text, Russell says airily: "Other ways of evading the paradox have been suggested, but have not been satisfactory." (p. 59) Notice that Russell openly *admits* here that he is "evading" rather than explaining the paradox.

[13] Yes, a particular chap who called himself a 'liar' may have said this on some occasion with the intention of drawing attention to his horizontal posture. Tackling such a case is, however, quite irrelevant to the problem of explaining the paradox, for the reasons given in Note 4 above.

[14] Because the Airport List example shows that the stratification is untenable. When we apply the Theory of Types to the Airport List we get a contradiction. The conclusion is not that the Airport List is contradictory (it obviously exists and serves a useful purpose), but that the Theory of Types is contradictory. Ayer (1982) comments "The theory of types achieves its purpose, but in a somewhat arbitrary fashion and perhaps at too great a cost", p.31.

Of course few later commentators swallowed the principle of Russell's "doctrine of Types". (Though Tarski later used a similar argument to assert the necessity of a hierarchy of meta-languages.) It was treated, from the beginning, as a source of unnecessary rigidity and artificiality within the system laid out in *Principia Mathematica*. It implied, for example, that there was not merely one number zero (0), but five different zeros, depending on whether one was talking about natural numbers, integers, rational numbers, reals or complex numbers. Such a multiplication of logical barriers tends to produce paralysis within a system. To escape from the worst paralysing effects of such rigid typing, Russell and Whitehead introduced their so-called 'Axiom of Reducibility'. This had an even worse press. Weyl and Poincare were scathing about it. Some called it "a sacrifice of the intellect". Ramsey (1931) commented: "Such an axiom has no place in mathematics: and anything that cannot be proved without using it cannot be regarded as proved at all".

[15] The final phase, which I call the 'Toothcomb' Phase, has consisted in going through the paradoxical arguments yet again with a fine toothcomb, i.e. a new level of attention to microscopic logical irregularity. What other response was possible? It appeared that previous analyses had failed miserably to locate the true explanation of the paradoxes. During this "Toothcomb" Phase (i.e. since the collapse of foundationalism in the 1970s) a considerable number of clever, highly specialised, discussions of the logical minutiae of the paradoxes has appeared. (See Note [1] of Chapter 2.) The drawback of such approaches is, however, that it is almost impossible to maintain a sensible overall perspective whilst craning for such minute logical imperfections. This gives many of the "Toothcomb Phase" contributions an unmistakable air of scholasticism. Cargile (1979), for example, subjects Ryle's "Namely-Rider" theory to a minute, case-by-case examination (pp. 262-266) which seems to imply that it deserves such attention, whereas its general implications immediately reveal that it won't do.

[16] And there was a certain amount of gilding the lily by releasing stories such as that of a comment written by Whitehead on the top of one of Russell's manuscript pages "Dear Bertie, The following seems to me rather beautiful." Clark (1981) p. 28.

[17] E.g. Thomson with his "small theorem" (1962) which stated blandly

"Let S be any set and R any relation defined at least on S. Then no element of S has R to all and only those S-elements which do not R to themselves."

This is an example of an approach which tries to solve the problem by generalising it, and thereby casting it in a pretty unfamiliar form, so that it no longer *looks* obviously paradoxical. (One could print Russell's Paradox in Chinese characters, no doubt, but it would still remain acutely paradoxical!) The retort to Thomson is that our logical intuition leads us to *expect* there to be an x in S which consists of those members of S which do not have the relation R to themselves. Some members of S have the relation R to themselves, some do not. Why, then, cannot we simply go ahead and consider the set of those which *do not*? Thomson does not "explain" this puzzle by merely wheeling-out the paradox itself (formulated now in terms of S and R). Others who tried to talk the problem into non-existence, by different methods, included Strawson (1949), Ryle (1951), Skinner(1959), Tucker (1963)

[18] Bronowski (1973) p. 110.

[19] Not only do the explanations fully explain what they set out to explain, but the results can also frequently be seen to be unquestionably true. So mathematics can, in principle, offer a kind of "safe haven" for both explanation and truth, two concepts frequently battered and abused by apologists for relativistic epistemology. Donald Scherer (1974) has even shown in an interesting article that there is a definite sense in which explanations of *empirical* events can "fully explain" what they set out to explain. Lucas gave a similarly robust treatment of 'True' in 1969. (He pointed out that to say that something is "true" is to say that it can be trusted.)

[20] In my draft monograph *Looking for Meaning in Mathematics* MAG-EDU 1985.

[21] C.f. Strawson (1952), p. 3: "Contradicting oneself is like writing something down and then erasing it."

CHAPTER TWO
Getting the orientation right: recognising the overall *validity* of self-referential language

WE TURN NOW to address the fundamental two-part question raised, in effect, at the end of the previous chapter: (1) "How is it possible to deal with the paradoxes of self-reference[1] *except* by banning some categories of language?" (2) How is it possible to introduce a logical ban on some category of language/argument without *harming* the integrity of logic?".[2]

In attempting to answer these questions we need, I believe, an initial Copernican-style conceptual inversion. It is no use first wreaking havoc on the integrity of logic and *then* belatedly asking how we can minimise the damage. What is needed is, rather, a heightening of consciousness from the beginning about the sort of harm that blanket logical embargoes can entail. This implies that we should focus first on what makes sense, even weakly, even trivially. All the early attempts to solve the problem of the paradoxes, I shall argue, got their orientation wrong. Instead of following their mistake and asking ---as became the norm--- what forms of speech/argument were to be *outlawed*, we should, I suggest, first establish what is to be *accepted*.

Most of the so-called 'solutions' to the problem of the contradictions have had the overall effect that they tried to outlaw self-referential speech *in toto*. The doctrine of Types, as we have seen, did this. The alleged necessity of placing judgements of truth and falsity within a metalanguage (Tarski's Theory of Truth) did this.[3] Ryle's proposed namely riders [4] did this. Most of the technical devices introduced into axiomatic set theory to remove the contradictions, pulled the trick by outlawing the possibility of a set being a member of itself, either in general, or, more economically, in the kind of circumstances in which Russell's paradox would otherwise be capable of formulation.[5]

Self-referential language, in other words, was *blamed* for the paradoxes. It was an easy scapegoat. Who would want to defend such an unusual, unnatural and generally trivial form of speech?

It was a sitting target. Epimenides and the Liar tried to comment on what they, themselves, had said. Grelling's paradox was about adjectives which applied, self-referentially, to themselves. Russell's paradox was about sets which were members of themselves. In each case there was a form of self-reference: ergo, self-reference must be the source of the contradictions!

Yes, of course, the kind of paradoxes with which we are concerned are paradoxes of the self-referential type. That is what identifies them: that is their genre. They are, certainly, cases in which self-referential language leads us into trouble. But it does not follow that because *some* self-referential uses of language are logically invalid, *all* self-referential uses of language are logically invalid. Only the most slapdash logician would draw such a faulty conclusion.

It is like saying that the Pan American 747 was flying over the Scottish town of Lockerbie when disaster struck, therefore Lockerbie must be the cause of the crash.

No, if we wish to focus accurately onto the issues which are at stake, it is essential to begin with a recognition *that self-referential language is, in general, a valuable and valid form of speech.* In other words, most self-referential language makes sense. There is much more self-referential speech out there, in the working language of mankind, than is generally recognised. What's more, it makes perfectly good sense. It is precisely because it makes perfectly good sense most of the time that we are frequently unaware of its presence. The paradoxes are the exceptions ---the celebrated exceptions--- to this general pattern. But until we are able plainly to see the paradoxes as contrasted against a substantial background of *valid* self-reference, we are likely hopelessly to misunderstand the essential nature of the problem.

The integrity of logic is harmed if we say that certain speech forms are logically impermissible which are not evidently and visibly so. But this is just what all the so-called 'solutions' to the problem of the paradoxes did. They said, or implied, that all forms of self-referential language were invalid.

Now the statement [6] made by the entertainer who walks up to the microphone and says:

"I am making this statement to test the microphone" has none of the obvious characteristics of "dubious sense": it does not look even *remotely* like *nonsense*. On the contrary, it makes perfectly good sense. Its meaning is open, for all to see. There is no subterfuge, no sting in the tail. There is no credibility, therefore, in the claim that it is *invalid*: and I suspect that no one who assumed on ideological grounds that it "must be invalid", would attempt to claim that it was invalid *in exactly the same way* as the statement made by the Liar. But many logicians have tried, nevertheless, to convince themselves that it is in some subtle, insidious, way "circular",[7] and therefore that it is not worthy to be accepted as a piece of logically authenticated language.

Other logicians have retreated a little at this point, and have tried to draw a line between invalid and valid self-reference: in effect conceding that a statement might validly comment on the *words* which expressed it as physical utterances, or even as forming, in total, a syntactically correct sentence.[8] But a statement could not validly comment, they argued, on its own *meaning* or on the question or conditions of its own *truth or falsity*. Such 'semantic self-reference', they maintained, was evidently invalid.[9]

I'm afraid they overlooked examples like:

"This statement is an example of a statement of extreme triviality and is also true"

which manages successfully to comment both on its own meaning (by classifying itself as 'trivial') and also on its truth. If we consider the statement carefully, we must, I think, conclude that it is indeed "a statement of extreme triviality" and that it is indeed true.

In other words, the notion that one can draw a line within self-reference and say, for example, that all "semantic" self-reference is invalid, will not do. Not only are there infinitely many clear-cut examples of valid semantic self-reference (like that offered above), but there are also all sorts of maddeningly borderline cases between the "semantic" and the "physical" ("non-semantic") kinds of self-reference. The notion that there is an absolutely precise distinction between references to "meaning" on the one hand and references to "words" on the other, seems to have been predicated on the Cartesian assumption that "meaning" was somehow utterly, metaphysically, distinct from the words which carried it. But this assumption is shipwrecked by Wittgenstein's insight that the meaning of words is derivative from their use.

"The uses of this statement are inevitably limited to treatises on the logic of self-reference"

is a statement which appears to comment perfectly successfully on its own uses. Whether it is a "semantic" or a "non-semantic" self-reference, however, is a question over which it is hardly necessary to agonise.[10]

No, we need, I believe, to recognise the principle quite firmly, quite explicitly, and from the beginning, that *a statement should be considered valid until it has been shown to be invalid*. A person, in English Law, is considered innocent until proved guilty. There are huge numbers of statements in common use which are wholly or partly or slightly self-referential. (See Appendix A for a selection of examples.) They have passed an extremely searching examination: namely, survival within the ordinary working language of mankind. A logic which tried to imply, or to insinuate, that they were "really" invalid, would be taking a major step, in my opinion, towards making itself ridiculous.

SELF-REFERENCE WORKS!

Many statements which refer to themselves, such as this one, do so glancingly, and there is in these cases little sense of logical "shock" as one becomes aware of their self-referential character. The notion that a logical tribunal could ever be set up to look systematically at the entire body of the language of mankind and weed out all such glancing moments of self-reference is clearly ludicrous.

It is an established principle that a cat may look at a king.

We should agree in the same spirit, I believe, that a statement may look at itself.

This is what the debate about self-reference is all about. "Reference" is no more substantial than a look. Trying to ban it is as absurd as past legislative attempts to stop people *swearing* (in the

Commonwealth 1649-60), or using *politically incorrect* words (some U.S. campuses) or *franglais* (in Paris, Quebec).

The photons from a candle may, after reflection in a mirror, fall back upon the candle, and even onto the flame itself. It is not a physical impossibility ---more an optical commonplace.

It is especially ludicrous that philosophers should suggest that self-reference is impossible, considering that philosophy is an inquiry into the sense and validity of general forms of thought, and hence a legitimate target for itself. Philosophy is full of self-referential arguments and questions, such as whether what Locke says about human knowledge (e.g. the difference between Primary and Secondary qualities) could be *known* if what he said were true.

It is even more ludicrous that logicians should countenance a systematic distrust of self-reference, considering that logic is the study of valid argument and that all its strictures must, of course, apply to themselves. (For example "The statement-form 'p and not-p' is contradictory" clearly applies when we substitute for 'p' the statement "The statement-form 'p and not-p' is contradictory".)

We have now entered the micro-electronic age, and it is evidently quite important that general logic, the logic of mathematics and the logic of computer programming should be put on a common footing. Reflecting on this desideratum we are driven to consider the existing customary use of self-reference within mathematics. Such customary use should surely be conserved in any concordat which puts the logical assumptions of logic itself, computer programming, mathematics and ordinary language onto a common footing.

Now mathematics began using self-referential forms of argument a long time ago. Indeed we may discern a smidgeon of self-reference in the very idea that an ordinary algebraic equation could make sense as a statement. It is a good thing that punitively anti-self-referential logicians like Russell, Tarski and Ryle were not around when the early Arab scholars started writing down statements like

$$x/2 + x/4 = x - 3$$

because 'x' clearly refers to a number which one will only find-out later, *after* one has first taken it on trust as a valid reference. Algebra, in a word, is a science founded on the perception that a "reference" can be held in suspension for a period of time, and may be involved *inter alia* initially in (defined) relationships to itself. We use arguments embodying a degree of self-reference throughout mathematics, for example when converting a recurring decimal like 0.121212... into a fraction. (This is the fraction which, when it is multiplied by 100 and reduced by 12, equals itself!)

There is an element of self-reference in Dedekind's definition of an infinite set as one which can be placed in 1-to-1 correspondence with a proper subset of itself. We remarked earlier on the considerable intrinsic value of this idea.

As long ago as the 1960s, programmers were writing down lines like this

```
FAC(N) = 1 IF N=1 ELSE FAC(N)=N*FAC(N-1);
```

and the Algol-60 compiler took this happily in its stride. Indeed it was one of the charms of Algol as an early programming language that one could specify mathematical sub-routines so elegantly. It is a clear indication that self-reference can *work* in a computer language.

These, then, are some substantial reasons for positively valuing and celebrating the self-referential aspect of mathematical language. Self-reference in mathematics contributes a great deal of flavour to the whole: rather, one might say, like spices in cooking. From the nature of the case it is of course the *combination* of an aspect of self-reference with *other* conceptual ingredients which produces the total effect. In cooking a dish made entirely of spice might, very likely, have a dreadful taste. Similarly, statements which are wholly self-referential, employing no other concepts, are unlikely to be of much value in mathematics. We should however ---I believe--- treat them as trivialities, like "A thing is a thing and not another thing", rather than attempt to read-into them a logical malevolence which is not there.

COMPLAINTS ABOUT SELF-REFERENCE

It would be quite wrong to try to envalue self-referential language ---as the early part of this chapter might be construed as doing--- without at some stage confronting the central case *against* self-reference as a mode of language. Let's look at the complaints which have been levelled at self-reference as a mode of language. What are the worst charges against self-reference? Can we find, perhaps, an element of truth in them? Could there be, after all, "something in" the recurrent targetting of self-reference as the primary cause of the paradoxes of self-reference?

Russell summed-up the complaint against self-reference as follows:

> It will be found that in all the logical paradoxes there is a kind of relexive self-reference which is to be condemned on the same ground: viz. that it includes, as a member of a totality, something referring to that totality which can only have a meaning if the totality is already fixed. (1959) p.63

Some commentators, e.g. Sainsbury (1988) pp. 138-138, and Russell himself call this the 'Vicious Circle Principle'.

But this argument fails wholly to account for the Airport List. It (the list) unquestionably includes "as a member" of the totality something "referring to that totality" namely itself. It *does* have a meaning, and the totality can in fact be "*fixed*" by its inclusion as the last listed member.

The argument also fails to account for the existence of ordinary classical algebra, e.g. where the recurring decimal mentioned above, *r*, satisfies the equation

$$r = 100r - 12.$$

How, it might be asked, can this definition of *r* possibly be acceptable, since it says that *r* is something (100*r* - 12) which we cannot begin to evaluate until "*r* is already fixed"? The fact is, as Russell knew perfectly well, that mathematics is full of such examples. Indeed we might say that mathematics has been pulling this trick of defining things in terms of themselves, and then *subsequently* establishing stability of meaning, for more than two thousand years.

The general case against self-reference as given, for example, by Ryle, is in the same case. According to Ryle, a statement may not refer to its own meaning, because that reference would only make sense if we already knew the meaning of the statement. But if we say

What this statement says is trivial and is also true

we are not, in practice, going to stumble over the meaning of the first four words ('What this statement says') because, even as we read it for the first time, we have a pretty good *general idea* what it is saying. "Taking the meaning of a reference" is always a very inexact science. Indeed if we had to know *everything* about any object to which we were referring *before* we could complete a logically correct statement about it, speech would never begin! There would be no possibility of permitting such a dangerously nonsensical statement as "The cat is on the mat" without giving a namely rider for the *cat*, and a namely rider for the *mat*. But then the items mentioned in these namely riders would inevitably lack referential precision: so we would be forced to give additional namely-riders. For example my initial namely-rider might be "The china cat your aunt gave me in 1970". But which "aunt" is this? "Namely, your mother's sister Freda" Who is this "Freda"? "Namely Freda Bainbridge who used to live in Osterley Park"... and so on, ad infinitum.

It is, in other words, extremely easy unthinkingly to endorse a standard of official referenceworthiness (i.e. a set of standards to which a reference must allegedly conform) which is completely unrealistic.

The key idea behind all this, however, is that of 'stability of meaning'. Algebraists from Diophantus onwards learnt that you could start out with what looked like a totally unstable meaning such as *r* = 100*r* - 12 and yet you could, contrary to appearances, stabilise it. There *is* such an '*r*' ---a single definitive numerical value for *r*--- and this self-referential definition (equation) tells you all you need to know to find it.

SELF-REFERENCE IS NOT TO BLAME FOR THE PARADOXES

The degree of panic and logical confusion which became attached to this question may be further illustrated when we turn finally to the black sheep of the family of self-referential statements, namely the paradoxes. Here surely, we find clear-cut, totally unacceptable examples of the logically mischievous nature of self-reference? Hardly. When the Liar said "I am lying" he was, we know, referring to the three words given in quotation marks in the early part of this statement.

The moment we recognise this, we tacitly accept that the Liar's reference has *worked*! The Liar tried to get us to focus on his statement "I am lying": and, let's face it, he *succeeded*! His statement may have ended-up being intensely paradoxical, but at least we got as far as working out *to what it referred*, namely itself.

If we had not already "taken" the Liar's reference, we could not see that there was a paradox.

In other words, the Liar's statement may have had all kinds of drawbacks as a linguistic gadget, but at least it succeeds in one respect: it gets us to look at the status of the three words "I am lying". As a referring gadget, it works, whatever else it may fail to do.

The problem of the paradoxes would not arise as a problem in the first place, then, if self-reference were really logically impossible ---as most commentators tried determinedly to persuade us it was, for ninety years.

The paradoxes of self-reference, it is evident, are not caused by self-reference, but by a particularly destructive form of self-ascription.

A man points a gun at himself and pulls the trigger: the result is a tragedy. There is little cogency in the analysis that "pointing anything at yourself is wrong". What did the damage was not the *act of pointing* in itself, but the cardinal fact that the thing being pointed was a loaded gun *which was then discharged*.

NOTES

[1] Sainsbury (Ch. 5, 1988), in a meticulously careful discussion, treats the paradoxes of "self-reference" as paradoxes involving "classes and truth". In effect he discounts self-reference as their defining characteristic, instancing indirect versions of the Liar paradox in support of this decision. It was Popper (1954) however who introduced the term 'indirect self-reference' for references a statement might make to itself via one or more other statements. Sainsbury appears not to recognise this as a kind of self-reference. But not to recognise that indirect self-reference is a species of self-reference begins, in my opinion, to introduce a slightly artificial slant into the problem, because it dulls our awareness of the role of self-reference in these paradoxes. On the credit side, Sainsbury *recognises* the general validity of self-referential language on page 123, commenting that:

"...self-reference is essentially used in universally accepted theorems of logic and mathematics ---for example, in Godel's incompleteness theorem."

Sainsbury identifies three kinds of overall approach to the paradox of the Liar, the "gap response", the "hierarchy response" and the "syntactic response". (The "gap" response consists broadly in the view that the Liar's statement contains some kind of category mistake, so that it can be excused from the normal judgement that a well-formed statement must be either true or false. The "syntactic response" consists broadly of the view that self-reference of the direct or indirect variety is to blame for the paradox.) The argument of the present essay, it should be pointed out, is more radical than each of these approaches. Sainsbury's discussion is useful because it relates closely to numerous contributions in the large over-specialised recent "Toothcomb Phase" literature on the problem of the paradoxes. (See Martin Ed. (1984) and Cargile (1979)) The complaint about much of this literature, however, is that it tries so *hard* to find fallacies in the *minutiae* of logical detail, that it ends-up getting the problem as a whole completely out of perspective. In my opinion most of Sainsbury's objections to arguments in the over-specialist literature are sound, but he is on less promising ground, I think, when he gives credence to the idea of 'logical gaps'. This is a notion which has a perfectly good raison d'etre in relation to some *formal* languages/systems (which contain formulae which can neither be proved nor disproved), but which is quite contrary to the assumptions and conventions of ordinary language.

In ordinary language a meaningful, non-vague statement is treated as being either true or false ---or else composed of parts which are true or false. Any so-called 'gap' is regarded as being occupied by vaguenesses and statements which embody category mistakes. But the Liar's statement is certainly *not* a vagueness of any kind, nor is it a category mistake in the ordinary sense. The statement made by the Liar appears indeed to be just the sort of thing (= of the right logical category) we expect to be true or false. Of course we eventually conclude that *for some unknown reason* it is "neither true or false", but we can't conveniently simply forget that the problem has now devolved onto that "unknown reason". The discussion of these matters is also hardly aided by couching it in terms of "sentences", which are essentially disembodied sequences of words, rather than in terms of "statements", which have *authors* and which imply some kind of apparent *intention* to communicate.

[2] By "harming" the integrity of logic I mean attempting to ban locutions which are not self-evidently meaningless.

[3] Many leading philosophers followed Tarski's lead, e.g. Carnap, Popper, Black. Sainsbury gives a rather sophisticated

example of an hierarchy explanation of the Liar paradox, that of Burge (1979), and he later modifies this to produce his own "gap plus hierarchy" approach (pp. 129-132).

[4] Ryle (1951) says:

"...for it to be true that I do or do not give my name, there must be a 'namely-rider' of the pattern "...namely, 'John Jones'". Nor could it be true that a person had a profession or a disease without there being a namely-rider of the pattern "...namely, theLaw" or "...namely, asthma"." (p.46).

Ryle's account of the Liar is as follows:

"When we ordinarily say "That statement is false", what we say promises a namely-rider, e.g. "...namely that today is Tuesday". When we say "The current statement is false" we are pretending either that no namely-rider is to be asked-for or that the namely-rider is "...namely that the current statement is false". If no namely-rider is to be asked-for, then "the current statement" does not refer to any statement. It is like saying "He is asthmatic" while disallowing the question "Who?". If, alternatively, it is pretended that there is indeed the namely-rider "...namely, that the current statement is false" the promise is met by an echo of that promise. If unpacked, our pretended assertion would run "The current statement (namely, that the current statement (namely, that the current statement (namely, that the current statement..." The brackets are never closed; no statement of which we can even ask whether it is true or false is ever adduced." (p. 52)

The same argument, of course, forbids self-reference in toto.

[5] E.g. the restrictive system suggested by Quine (1953).

[6] Many commentators nowadays write of "sentences" as if they "said something" (i.e. had *particular meanings*), rather than simply as strings of words capable of expressing meaningful statements. Some commentators object outright to the idea of a "meaningless statement". Such procedural matters are discussed in Appendix B. Generally I shall attempt to stay as close to natural usage as posssible, speaking of *statements* as the fundamental carriers of specific meanings and *sentences* as grammatically acceptable strings of words which someone might want to use to make statements of various kinds. I shall use the term *propositions* only for results in mathematics and other formal systems which are, or which claim to be, timelessly true, potentially true, or false. (In all three cases a type-token distinction is also needed.)

[7]'Circularity' may be taken to imply the emptiness of meaning which undoubtedly occurs when a term A is defined in terms of B, whilst B is defined in terms of A. This is essentially a failure of reference, because nothing has been introduced onto which the mind can focus vis-a-vis the rest of the world. However such clear-cut "circularity" is not present in typical self-referential statements such as those given in Appendix A. The emptiest example of self-reference is probably "This statement is true" but it is hard to detect any greater emptiness here than in (say) "A statement is a statement".

[8]For example, Sainsbury (1988) fn. 22, p.135-136.

[9] Early commentators on the paradoxes tended to take this line, but increasingly this line of defence has broken down. The first breach was probably Popper's paper of 1954. Cargile (1979) gives several examples of the absurdity of not permitting general statements of logic to apply to themselves. E.g.on p. 266 he observes that "No proposition is both true and false, and that includes the proposition that no proposition is both true and false". (Later he spoils the effect by disallowing self-membership of sets (p.302))

[10] I.e. over whether a reference to the "use" is, or is not, a reference to the meaning.

ABOUT THE MAG

The MAG was founded by Wilf Flemming and Chris Ormell in 1978 with the purpose of fostering the use of "projective" mathematical modelling in schools and colleges. The approach is based on a *Peircean* interpretation of mathematics as the "science of hypothesis". (See Peirce 1956) (Nidditch (1960) p.287, finds an early expression of this view of mathematics in Dugald Stewart. He quotes Stewart (1818): "[mathematics traces] the filiation of consequences which follow from our assumed hypotheses") Treating mathematics in this way enables us to develop mathematical rationality in active collaboration with the learner's practical *imagination*. The scope for interesting materials based on this principle appears to be unlimited.

CHAPTER THREE
When self-reference combines with radically adverse self-ascription

GIVEN THAT statements which refer to themselves are normally valid, what is the chief *danger signal* that a paradox might be about to erupt? The answer is obvious: it occurs when the "comment" the statement makes about itself is self-destructive, that is, when it is radically adverse to the statement's initial, presumed sense.

This is the real source of the logical mischief.[1]

The words 'The statement written on this blackboard is false' written on an otherwise clear blackboard certainly succeed initially in drawing attention to themselves. We also assume quite quickly that they constitute a "statement" (because they form what looks like a genuine, syntactically-correct sentence which has an author and an apparent intention to communicate). This "statement", we see at once, is commenting on its own trustworthiness. But what it actually says is that it, itself, is *false*! Since we begin, naturally, with the premiss that any statement is, and claims to be, *true*, this is quite evidently a contradiction ---one, indeed, evidently enabled by the channel of self-reference.

But it is not like an ordinary contradiction, which occurs in the form of a conflict of meaning between free-standing, fully-formed, separate statements or sub-statements. Here we have a kind of internal contradiction *within* the statement itself.

In other words, we have only just got to the point of applying the suggested ascription to the suggested reference, when we realise that the statement is self-contradictory.

Our initial reaction is of course to try the alternative hypothesis: that the statement is false. But this leads us to the obverse self-contradiction: namely that the statement, now presumed false, is claiming to be true!

At this point we have traditionally tended to declare "a paradox". Both initial assumptions ---that the statement is true and that it is false--- have led us to these swift, internal, self-contradictions. The situation is "paradoxical", because we appear to have no way whatsoever to answer the question whether the statement is true or false.

This, I suggest, is a fairly accurate blow-by-blow account of our "interpretive experiences" when we first encounter a paradox like the blackboard paradox. Similar stages occur in the case of the other natural paradoxes of self-reference.

It is not difficult to show, for example, that Russell's, Grelling's and the Barber paradoxes arise from statements making radically adverse comments on themselves. In the case of Russell's paradox the statement "x is a member of the ordinary set"[2] *means* (by the definition of the ordinary set) that "x is a member of x" is false. If we now replace 'x' by 'the ordinary set' we get the result that "the ordinary set is a member of the ordinary set" *means* that it, itself, is false.

Our first conclusion, then, is that the paradoxes are caused, not by self-reference, but by radically adverse self-ascription. This follows simply from attending carefully to the form of the paradoxical arguments.

But the "method of naturalistic observation" can be taken further. Let's consider carefully what the blackboard statement tells us. It tells us that the statement written on the blackboard is false. We concluded that a "contradiction" had occurred, i.e. a conflict between two distinct "dictions", two free-standing statements. But this was not quite what happened. What actually happened was that, the statement we were interpreting *changed*, as we interpreted it ---like a chameleon under our very eyes--- from an apparently true statement into an apparently false statement.

So what does this really mean? We are still in the business of trying to resolve the meaning of the statement. If the statement on the blackboard is false, it follows that what it is claiming should be negated, in other words, that the statement on the blackboard is true.

So what does this, in turn, really mean? Evidently, that the statement on the blackboard is false. Which means that the statement on the blackboard is true... Which means that the statement on the blackboard is false... and so on...

Our second conclusion, therefore, is that no "conclusion" to the paradoxical argument is ever reached. It is not that two settled, stable conclusions are reached, which contradict each other: rather that the sense of the statement keeps oscillating between the provisional judgement that the statement on the blackboard is true, and the provisional judgement that the statement on the blackboard is false.

One could describe this by saying that the "provisional sense" of the statement on the blackboard is *looping* between:

the statement is false

the statement is true

By 'looping' we mean that the first self-reference leads us back to the statement itself and the radical self-ascription then leads us to start again. It is rather like the process of polymerisation in chemistry, which is not a single act of chemical combination, but an on-going chain of chemical combinations. We might describe this process of "going back" to re-interpret the sense of the original statement more than once as *repeated self reference*, or *continuing self-reference*. (We shall also sometimes refer to it as "looping", or as a case where the self-reference "cycles on and on".)

We now look at the various forms which "continuing self-reference" can take: then we return to the theme of the role of continuing self-reference in explaining the paradoxes.

Strange as it may seem, not all examples of "continuing self reference" are paradoxical. For example, consider the statement:

"This statement should have the additional phrase 'Oh Boy!' appended at the end."

It is clearly a self-referential statement. Let's read it carefully, study the instruction it embodies, and carry it out!

It instructs us to add the two additional words 'Oh Boy!' at the end. So, when we do this, we get the result:

"This statement should have the additional phrase 'Oh Boy!' appended at the end, Oh Boy! "

It is now still, clearly, a self-referential statement[3], which is still advising us to add the words 'Oh Boy!' at the end.[4] So what happens is that the statement, interpreted literally, grows ad infinitum...

This is a case of continuing self-reference, but it is not, in the ordinary sense of the word, a 'paradox'. It is not even "nonsensical", interpreting that word in the normal way.[5] It is however, in some sense, logically unstable. We might call it a 'ballooning statement', or an example of 'ballooning' self-reference. It is a sign that once self-reference is accepted in general as a valid mode of language, we are going to have to introduce certain (hopefully self-evident) embargoes on recursively unstable sense, as well as on recursively unstable nonsense. We are going to have to *classify* the different kinds of absurdity which can occur.

Some readers might ask at this point: why not simply place a general embargo on any kind of self-referential loop, on any kind of cycling on and on? Unfortunately this simple rule won't do. There are a few cases in which the reference "cycles on" and yet leads eventually to a stable conclusion.[6]

To give an immediate example:

"This statement should have the phrase 'Oh Boy!' appended at the end if that phrase occurs in the statement a total of less than four times."

If we follow the "literal procedure" carefully we finally end up with this:

"This statement should have the phrase 'Oh Boy!' appended at the end if that phrase occurs in the statement a total of less than four times, Oh Boy! Oh Boy! Oh Boy!"

This is now inert. It is not recommending that we add any more appendages. It now makes sense, therefore, as a statement with a rather silly, but nevertheless settled, stable, meaning.[7]

RADICAL SELF-ASCRIPTION

IF WE SAY THAT "the statement made by the Prime Minister yesterday was far-sighted" we *refer* to the statement the Prime Minister made yesterday and we *ascribe* to it the characteristic of being far-sighted.[8]

If we say

"The word before the penultimate word of the statement printed in 10 pt type on this page should not be there"

we *refer* to a particular word in the current statement printed in quotation marks in 10 pt type above, and we *ascribe* to it the characteristic that it "should not be there". Such an ascription is "radical" in the sense that it claims that a significant *change* should be made, namely, that the word in question should be removed.

The result, of course, is a paradox. For the word in question is 'not', and if we remove it ---as the statement in effect asks us to do[9]--- we get a new statement:

"The word before the penultimate word of the statement printed in 10 pt type on this page should be there."

The new statement refers to the same statement as before.[10] It also makes an effectively "radical" ascription, because it is now a *radical* thing to say that the word which we have just removed "should be there".

Notice that, like the ballooning radical self-ascription, the two implications (that the word should "not be there" (NBT) and 'be there" (BT)) [11] both remain radically self-ascriptive. This is another paradoxical example of continuing self-reference, though it is not a traditional paradox. No stable conclusion is ever reached. NBT implies BT which implies NBT which implies BT which implies NBT... and so on ... ad inf. [12]

[It is evident that the paradoxes are technically defined by paradoxical *questions*, questions to which, it appears, no answer can ever be given: because each possible answer (of two mutually incompatible answers) implies the other. In this case the paradoxical question is:

"Should the word 'not' be there?"]

Sometimes a self-reference "continues" for a while, but not indefinitely.

Consider the following statement:

"The third, redundant, word in the current statement should not be there".

This makes a radical self-ascription which, in effect, asks us to remove the word 'redundant'. So we get

"The third word in the current statement should not be there".

This, too, is radically self-ascriptive and it asks us, in effect, to remove the word 'word'. So we get

"The third in the current statement should not be there".

What has happened now is that after two radical self-ascriptions the statement has damaged itself so much that it has ceased to be syntactically well-formed.[13] It is like a computer program which runs for so long and then fails.

So the mere presence of "radical self-ascription" is not sufficient to produce a paradox. To get a genuine *paradox* the message[14] which results from the radical self-ascription must also, symmetrically, lead us back to the original statement. I shall call a statement which does this a '*reflexively* radical self-ascriptive statement'.

Let's define the word 'omu' to mean *upside down* and the word 'nwo' to mean *right-way up*. [15]

Now consider the statement

"The last word of this statement should be read omu".

This is both radically and reflexively self-ascriptive and it produces a paradox. At first it tells us to read the last word upside down. Then it tells us to read it the right way up.

Of course in all such cases an ambiguity is apt to creep in: is the reference essentially to the original statement, or to the current version which we have obtained by following the implicit instructions contained in its radical self-ascription? One can adopt the convention that either of these is the "presumed reading", the other then having to be signalled by a special device. I shall adopt the convention in the sequel that such a statement always refers back to the *original sentence* unless it is signalled to the contrary by including the words 'the current version'.[16]

Let's define the word 'xox' to mean *upside down*.

Now consider the statement:

"The last word of this statement should be read xox".

This, surprisingly, is not a paradox, even though the self-ascription is reflexive, and is normally "radical". In this particular case it happens *not* to be 'radical', because reading the word 'xox' upside down doesn't change a thing!

TYPES OF OUTCOME: CONTINUING RADICAL SELF-ASCRIPTION

WHEN A STATEMENT makes a radically self-ascriptive reference to itself, various kinds of outcome are possible.

First, the sequence of successive self-ascriptions may terminate. In the commonest case it leads eventually to a breakdown of sense, because the current version of the statement becomes seriously damaged syntactically.

It may terminate after leading to a state of affairs where it exhibits a self-evidently true meaning, a self-evidently false meaning, a meaning whose truth or falsity will be contingent on local factors, a category mistake, an ordinary inconsistency, or a logically necessary statement.

Second, if it fails to terminate, it may "balloon" into a longer and longer form of words, or it may enter a paradoxical or oscillating state where it flips back and forth between two mutually inconsistent partial meanings like NBT and BT.

Terminating examples of radical self-ascription can be classified in just the same kinds of ways as ordinary statements, which may fail to make sense because they contain reference mistakes (the attempted reference does not work because it makes an invalid assumption) ascriptive mistakes (the attempted ascription does not work because it makes an invalid assumption), and category mistakes where there is a total mismatch of logical type.

The non-terminating examples, however, do not quite correspond to anything in ordinary language. Radically self-ascriptive statements can be devised to "balloon" in all sorts of different ways, not merely by the simple appendage of the words 'Oh Boy!' at the end. No one is likely to be puzzled by them. Once we have seen what they do, we naturally treat them in the fashion of the argument in Lewis Carroll's 'What the

Tortoise said to Achilles', i.e. as infinite regresses into which there is little temptation for any sensible person to get stuck.[17]

It is a particular genre of self-referential nonsense, which we may call *ballooning nonsense*. No elaborate argument is needed to see that it is nonsensical.

The paradoxes, on the other hand, have traditionally been classified as 'contradictions', and so their origin in non-terminating self-reference has remained comparatively unremarked.[18]

They may be aptly described as examples of *oscillatory nonsense*, because the partial meaning of the statement keeps flipping from state A to state B, and back to A, and hence to B, ... and so ad infinitum. No stable meaning is ever reached.

NOTES

[1] This is the equivalent of the bullet in the loaded gun in the metaphor at the end of Chapter 2.

[2] The 'ordinary set'= the set of sets which are not members of themselves.

[3] We are here following colloquial practice in treating the 'Oh Boys!' at the end as being "part of the statement", notwithstanding the nominal punctuation.

[4] One may describe this situation by saying that the self ascription is still "radical" or still "live" (= generative of new implications).

[5] Because we can still make sense of it and add another 'Oh Boy!' at the end. It is still, in other words, "radical" or "live" in the sense of the previous Note.

[6] By 'cycling on' we mean implementing the radical self-ascription, getting a new version of the statement and then implementing *its* radical self-ascription... and so on. By a 'loop' we mean an indefinitely continuing sequence of such steps. The pragmatic reason given in the text for not banning it as such is that some radical self-ascription which enters a loop ends-up in a stable condition. To ban such loops in toto would therefore violate the general principle which we stated in Chapter 2, viz. that blanket embargoes which ban locutions which are not self-evidently nonsensical damage the integrity of logic.

[7] It might be said that the use of the word 'if' here is not completely natural, because the statement's author knows very well that no problematic issue arises over the number of 'Oh boy!'s present. This slight imperfection could be corrected by considering an example with a much larger number of 'Oh boy!'s at the end, say 104. There would then be a real question-mark over the exact number of 'Oh boys!' present and the phrasing would be quite natural.

[8] For a good account of the reference-ascription distinction see Austin (1964).

[9] To adopt this attitude is to adopt the attitude we normally apply to the Liar, that is, to learn as much as possible, by any available method, from the statement made. When the Liar announces that he is *lying*, we do not respond simply by losing interest in his remarks, but rather we try to salvage something from the wreck, to "deduce" (in an informal way) what we can from the various things we know.

[10] The word in question ('not') is of course centre-focus of the reference, but it is referred-to as *part* of "the statement in 10 pt type on this page" (p. 26), so there is a partial reference to the statement in 10 pt type, the same statement as before.

[11] We shall refer to this in the sequel as the 'NBT paradox'.

[12] This clearly is the result of a radically self-ascriptive loop, which produces an *oscillation* between the two implications labelled NBT and BT. Recognising the fact of such potential oscillation is the central plank of the main argument of this book. (It is "potential" because few readers, obviously enough, will actually pursue it very far: once the oscillating pattern has become plain, there is no point whatever in pursuing it further.)

[13] One *might* argue that 'third' means the 'third word', though this is not a "presumed" interpretation. In any case the sense of the statement will break down at the next stage, after removal of the word 'in'.

[14] I.e. the current interpretation of its meaning.

[15] It is always acceptable to introduce a paradox by formulating special definitions like 'Cretans' = an island race of liars, 'homological' = an epithet which applies correctly to itself... (Kilmister (1967) calls this 'autological' p. 49.)

[16] In the case of the omu-nwo paradox *either* convention will do.

[17] See *Mind* 1895 pp. 278-280.

[18] Kripke (1975) and Cargile (1979) both remark ---in passing--- examples of meaning going round in circles. In neither case however do they quite see the central significance of this phenomenon.

CHAPTER FOUR
The mechanisms of superparadox: partial meanings and dynamic contradiction

WE HAVE SEEN THAT one of the consequences of recognising the general validity of self-reference is that heightened attention is needed vis-a-vis two new forms of characteristic nonsense: benignly absurd statements of the ballooning kind, and paradoxical statements, whose implications effectively oscillate between two mutually inconsistent states.

I shall describe the latter oscillations as "oscillations of *partial meaning*".

By 'partial meaning' I mean a partial answer to the question "What does this mean?". For example, George may say "What the Prime Minister said yesterday was unwise". I did not hear what the PM said yesterday, but nevertheless I have managed partially to grasp the meaning of George's statement. In a word, I have taken in the message that there is an X (what the PM said yesterday) which, George is claiming, was 'unwise'.[1]

The paradox of the Liar exemplifies oscillation of partial meaning in its clearest form.

Let's consider the statement:

This statement is false. (I)

If we ask ourselves "What is the meaning of (I)?" we can at least achieve a partial answer straightaway, namely that (I) is false.

This is the *first partial meaning* of (I).

As we focus on the implication of this first partial meaning, however, we become clear that it implies that (I) is true.

This is the *second partial meaning* of (I).

Now the second partial meaning of (I) is still only a partial answer to the original question. Focussing on its implications, we realise that it implies that (I) is false.

So we have gone back to the first partial meaning!

This, in turn, leads us back to the second partial meaning...

and so on.

If we denote the two partial meaning by M1 and M2 we may say that M1 'leads to' M2 and M2 'leads to' M1.

The oscillation of partial meaning which occurs is an oscillation passing through the stages

M1 to M2 to M1 to M2 to M1 to M2 to M1 to... and so on ad inf.

It is important to conceptualise the proper relationship between M1 and M2. We arrive at M2 simply by working out what M1 *means*. Thus M2 does not stand against M1 in a state of static contradiction between two free-standing "dictions". M2 is simply a clearer (slightly more developed) version of what M1 means.

M2 in a word *supersedes* M1 as our current partial answer to the question "What does (I) mean?".

Similarly M1 *supersedes* M2 as our current partial answer to the question.

Neither M1 nor M2 are fully stabilised meanings: they are, rather, meanings in-the-process-of-gestation.

We never reach any stable conclusion to the series of mental transactions involved in trying to work out the meaning of (I). Instead we get an *oscillation of partial meaning* which continues, in principle, indefinitely.

HOW LOGICIANS MISREAD THE SITUATION

Most logicians have simply noticed that there are two initial assumptions one can make about (I), namely, that it is true (M1) or that it is false (M2).

They have then observed that M1 implies M2, so that is a "contradiction".

Also M2 implies M1, so that is a "contradiction".

We appear to be forced into a contradiction, whichever assumption we make! This situation is then pronounced 'paradoxical'.

R. C. Skinner in an ingenious paper (1959) claimed that the solution to the "paradox" is that it is *not* a paradox! His argument was that M1 implies M2 which implies M1, which is alright. And M2 implies M1 which implies M2, which is also alright!

Both the traditional and the Skinnerian response stop prematurely. Both assume that a "conclusion" has been reached, though they locate this "conclusion" at different points. Actually no "conclusion" has been reached. The so-called "conclusions" to the traditional and Skinnerian arguments are just as maddeningly "partial" and inconclusive as the initial assumptions with which we began.

DYNAMIC CONTRADICTION

An ordinary contradictory statement, which tries to tell us simultaneously that p and that not-p, clearly fails to make sense. It clearly fails to convey any definite message. It *stultifies itself*. A speaker who comes up and announces "It is both raining and not raining" visibly fails to communicate with us; he shoots himself in the foot, verbally, before our eyes.

This is why no one ever thought that an elaborate theory was needed to establish that contradictions were unacceptable in human discourse. They are obviously unacceptable. End of discussion!

Now a statement which gives rise to an oscillation of inconsistent partial meaning is, I suggest, in an exactly similar case. It, too, visibly fails to convey a definite message. It, too, obviously stultifies itself. How can it hope to communicate with us, when it keeps *vacillating* between two inconsistent partial meanings? The only difference is that now it is stultifying itself serially, dynamically, in time[2]; whereas a classic "contradiction" stultifies itself statically, timelessly, or in technical terminology, 'in parallel'.[3]

We therefore require a new category of *dynamic* or *serial* contradiction, into which we can classify examples of oscillatory self-stultification.

It is to be contrasted with classic contradiction, which we may now more expressively describe as *static* or *parallel* contradiction.

This, I suggest, is a long overdue reform of the fundamental basis of logic, which then becomes the art of avoiding contradiction of both the static and the dynamic varieties.

AN EXPANDED LOGIC

A logic which embraces the object of avoiding *both* contradictions of the parallel *and* of the serial varieties, will be, of course, an 'expanded', and also a less one-sided, logic. It is like adding a new, fundamental logical axiom. At the same time the new axiom is clearly as necessary and unavoidable as the time-honoured axiom which enjoins us to reject classical contradiction.

Such a fundamental change in logic should not, of course, be adopted lightly. But what is the alternative? For nearly a century no simple, straightforward, universally applicable explanation of the paradoxes of self-reference has been forthcoming. First, this state of affairs has cast a dreadful (because heavily repressed) sense of fundamental intellectual gloom across the Twentieth Century. As a result, confidence in reason and rationality has fallen to levels which may begin to endanger the human race. Second, it has driven theorists to adopt a greater and greater logical hyper-sensitivity, so that in the end they have trained themselves to try to seek-out serious logical mistakes buried in the most innocuous, harmless examples of everyday inference. This is, in many ways, an even more disastrous legacy (of logic's failure to solve the problem of the paradoxes) than the collapse of intellectual confidence. It reduces otherwise sensible theorists to gibbering hyper-sensitive sifters of the minutiae of meaning and implication. And of course it offers academia a model of intellectual inquiry to which hyper-pedantry, scholasticism and political correctness become all-too easily the inevitably garbled expression.

It is not being maintained, of course, that there is a symmetry of application as between static and dynamic contradiction. Dynamic contradiction is, by the nature of the case, only likely to occur within the ambit of the self-referential use of language. This, in itself, reduces its application immediately to a small fraction of that of static contradiction. At the same time it is a *theoretical* necessity, which like the positron in physics, may be comparatively rarely met, but which we ignore at our peril.

THE FIRST TASK OF AN EXPANDED LOGIC

The first task of an expanded logic must surely be to look for general patterns of dynamic contradiction: to form some reasoned view of the likely extent of this newly conceptualised form of logical breakdown. The new paradoxes (NBT and Omu) introduced in Chapter 3 show the way. There are evidently *all kinds of ways* in which we can construct linguistic gadgets to produce oscillations of partial meaning --- once we have firmly taken the new conception on board. Everything looks completely different once we have put-on, and got used-to, the new conceptual spectacles!

The chief theme of this book is that dynamic contradiction need not be limited merely to vacillations of partial meaning between just *two* inconsistent states, M1 and M2. There is no theoretical reason why we should not construct linguistic gadgets[4] which will generate oscillations of partial meaning of the kind

M1 to M2 to M3 to M1 to M2 to M3 to M1 to M2 to M3... etc.

that is, whose partial meaning will *cycle* round sets of three or more nominally inconsistent states.

This is, in effect, an *hypothesis* generated by the account of the paradoxes offered above. An explanation of the paradoxicality of the natural paradoxes has been produced. It locates the source of the fallacy in the paradoxical arguments in the breach of a new principle, viz. that dynamic contradiction should be abjured.

Many will immediately assume that the "explanation" is essentially scholastic: that 'dynamic contradiction' is just a fancy name for 'paradox': that no genuine "explanation" has been produced.[5]

Well, this new "explanation" of the paradoxes generates a falsifiable hypothesis: it tells us that there *ought to be* examples of dynamic contradiction to be found where the partial meaning of the statement in question cycles round closed sets of 3, 4, 5, ... *n* ... mutually inconsistent states.

Such examples, if found, will be called 'superparadoxes'. The question now is: can they be found?

THE ARTHUR SUPERPARADOX

A man called Arthur suffered from the peculiarity of *selective prolixity* in speech. That is to say, his sentences always contained redundant words which, psychologists were agreed, were not intended. After his speech-behaviour had been intensively studied by a team of expert observers over a period of six months, they came to the following conclusions.

> When he was *angry* the 3rd, 4th, 5th, 7th and 8th words were redundant. When he was *drunk* the 3rd, 4th, 5th, 6th and 7th words were redundant. When he was *sober* the 3rd, 5th, 6th, 7th and 8th words were redundant.

In their Report the experts concluded: "These patterns have remained constant, and have applied to everything he has said".

Since Arthur is always either angry, drunk or sober, everything he says has to be, so to speak, de-coded ---if one wishes to decipher its intended meaning, that is. In order to do this reliably it is necessary, of course, to know what condition he is in. Unfortunately this depends on Arthur's cooperation. On one celebrated occasion he would offer only the obscure information

"I am either angry, or drunk or sober".

This clearly raised the question: which state was Arthur in?

If Arthur was angry, the statement, on being deciphered, read "I am drunk". If he was drunk, it read as "I am sober". While if he was sober, it read as "I am angry".

The question was superparadoxical, because it admitted three possible answers, each of which implied one of the others.

Arthur's strange schizophrenic-like condition thus enabled him to articulate a superparadoxical statement which is not available to people without the condition.

This was effectively the first superparadox, discovered by the author in 1954.

THE 123-TRUE SUPERPARADOX

Paradoxes similar to the Liar can, we know, be produced without postulating a liar, e.g. "The statement written on the blackboard is false" written on an otherwise empty blackboard. The Liar paradox has been de-personalised. This suggests that it might be possible to de-personalise the Arthur superparadox in a similar way.

To do it, we require epithets applying to statements which will cause us to re-interpret them, but in a more selective fashion than 'true' and 'false'.

First, however, we need a preliminary definition. A statement will be said to be a 3/1-worder if, after the first three words, the sentence expressing it consists of a set of words only one of which is capable of forming a true statement when taken in conjunctiuon with the first three words.[6] For example, "Maastricht is a disciplinarian, chocolate-bar or town" is a 3/1 worder, for "Maastricht is a town" is true, whilst replacing 'town' by 'disciplinarian', 'or' or 'chocolate-bar' gives a false or meaningless statement. The following types of 3/1 worders may be distinguished:

(1) A 3/1-worder will be said to be 1236-TRUE if it would be true if only the 1st, 2nd, 3rd and 6th words were read.

(2) A 3/1-worder will be said to be 1238-TRUE if it would be true if only the 1st, 2nd, 3rd and 8th words were read.

(3) All 3/1-worders which are *not* of types (1) and (2) will be called 123N68-TRUE. [7]

Thus every 3/1-worder belongs to one of the categories 1236-TRUE, 1238-TRUE and 123N68-TRUE. Now consider the following statement:

"This 3/1-worder is either 1236-TRUE, 1238-TRUE or 123N68-TRUE." (A)

(For the purposes of definitions (1) and (2) above, hyphenated expressions count as single words.) [8] It is clear that (A) *is* a 3/1-worder, for after "This 3/1-worder is..." one and only one of the predicates '1236-TRUE', '1238-TRUE' and '123N68-TRUE' must make a true sentence (and the other possibilities 'either' and 'or' are meaningless.) Now if (A) is 1236-TRUE the 3/1 selection "This statement is 1238-TRUE" is true. But if (A) is 1238-TRUE, the 3/1 selection "This statement is 123N68-TRUE" is true. In this case, since (A) is a 3/1-worder, it means that (A) is either 1234-TRUE or 1235-TRUE or 1237-TRUE. But of these the first and last are impossible, since the selections in question are meaningless; therefore if (A) is 123N68-TRUE the 3/1 selection "This statement is 1235-TRUE" is true. Reading it as such, we get "This statement is 1236-TRUE". . . and so on. It follows that (A) is a superparadoxical statement.[9]

As the 123-TRUE superparadox generalised the de-personalised version of the Liar, so we may similarly generalise Grelling's paradox.

THE PRIMO-, SECONDO-, TERTIO- SUPERPARADOX

Let's call an epithet of the form "first x, then y, then z" (where the variables x, y, z stand for mutually exclusive epithets) a *triplet*. Triplets may be divided exhaustively into three kinds:

(1) If "first x, then y, then z" is x but not y or z, we call the triplet *primo-dominant*.

(2) If "first x, then y, then z" is y but not x or z, we call the triplet *secondo-dominant*.

(3) If "first x, then y, then z" is is z, but not x or y OR any other possibility except (1) or (2) we call the triplet *tertio-dominant.* [10]

EXAMPLES For primo-dominance: "first long, then of moderate or average length, then of negligible or infinitesimal length". For secondo-dominance: "first illegible, then misspelt, then orthografic" is misspelt. For tertio-dominance: 'first adverbal, then prepositional, then adjectival" is adjectival.

From the definitions it follows that no triplet can belong to more than one of the three categories. By virtue of the additional ("miscellaneous") clause in (3), the three categories are comprehensive: thus every triplet is either primo-, secondo- or tertio- dominant. Now consider the question.

"Is "first secondo-dominant, then tertio-dominant, then primo-dominant" primo-, secondo- or tertio-dominant?"

The question must admit of an answer! Now if the answer is primo-dominant, by virtue of its form the triplet must be x, i.e. secondo-dominant. But if the answer is secondo-dominant, by virtue of its form the triplet must be y, i.e. tertio-dominant. Finally, if the answer is tertio-dominant, by virtue of its form the triplet must be z, i.e. primo-dominant OR neither of these. But we know the answer must be one of the three,[11] therefore it must be primo-dominant. It follows that the question is a superparadoxical one.

This superparadox and the 123TRUE superparadox were discovered by the author in the period 1954-56, and announced in 1958.

GENERAL FEATURES OF SUPERPARADOX

A 'SUPERPARADOX', may be defined in general as a question possessing n distinct possible mutually exclusive answers. The answers, however, possess only partial meanings. They may be labelled A_1, A_2, A_3,... A_n

such that
A_2 follows from A_1,
A_3 follows from A_2, ...
A_n follows from A_{n-1}
and A_1 follows from A_n.

We shall call the number of distinct partial meanings, n, the 'amplitude' of the superparadox.

The 'degree of paradoxicality' of a such a superparadox will be defined as the number of distinct, formally inconsistent partial meanings which may be 'deduced' from it. Each of the As may be nominally shown to be necessary, by considering the inconsistency of the remaining members of the set of As taken together. $2^n - 1$ different subsets of these As may be formed, and each one is (taken as a composite expression) nominally necessary by the argument indicated above.[12] They are all, however, formally mutually inconsistent![13] Thus $2^n - 1$ may be taken as a measure of the degree of paradoxicality of a superparadox of amplitude n.

The same conventions may be applied, obviously enough, to the natural paradoxes. In the case of the Liar the value of n is 2 and the degree of paradoxicality is 3. This arises from the set of nominally necessary results which may be derived from the Liar's statement (L):

(L) is true
(L) is false
(L) is true and (L) is false

In the case of a superparadox of amplitude 10, the degree of paradoxicality, on this measure, is 1023, and for a superparadox of amplitude 20 it is 1 048 575. It is this *exponential* increase in the degree of paradoxicality (in this technical sense) as n increases, which, I submit, justifies the use of the term 'superparadox'.

But although the more general superparadoxes we consider in the next Chapter are all nominally generators of large bodies of technically inconsistent 'conclusions', they are not more *spectacularly puzzling* than the natural paradoxes. If anything, they are *less* puzzling than the natural paradoxes, because they visibly depend on the adoption of pretty artificial definitions, and the mechanism of radically relexive self-ascription is there for all to see.[14]

It should be said in defence of the superparadoxes, though, that their artificiality is simply an extrapolation of the artificiality of the natural paradoxes. The Epimenides has been regarded as 'a verbal trick' for more than 2000 years. But this observation does not, of course, remove the logical embarrassment which is created if we are unable to explain exactly *which* self-evident logical principle has been flouted in its formulation. What is at stake is the self-consistency of logic, or, if you prefer, the claim implicitly made by logicians that they understand the basis of their own science.

THE GENERAL FORM OF THE MECHANISM OF SUPERPARADOX

The general form of any superparadox is as follows:

(1) There is the *principal question*, whether a specified object O has one of the named characteristics C_1 or C_2 or C_3 ... or C_n.[15]

(2) There is the *definition* of these characteristics.

(3) There is the prior demonstration that O ought to possess one and only one of the characteristics. We shall call this the *prior demonstration of closure*.[16]

(4) Finally there is the demonstration that possession by O of each named characteristic implies the possession of another, different, named characteristic.

A superparadox consists in the combined effect of these four steps. It appears from (1), (2) and (3) that O *must* have one of the named characteristics. But it turns out from (4) that it *can not* have any of them.

We turn in the next Chapter to look at a variety of specific *types* of this general form.

NOTES

[1] The closest analogy to this is probably the difference between our response to a particular number, and our response to *x* in algebra ---where *x* is defined by a specific, soluble equation. We accept '*x*' as a genuine item, principally because we trust that "all will be revealed later" (about its value). We proceed, in other words, on a basis of *hope* ---that a stable meaning will subsequently emerge. In a similar way we accept partial meanings on the basis of the *hope* that a stable meaning will finally establish itself.

[2] There is no reference, of course, to specific *clock* times. To say here that the contradiction occurs "in time" is simply to say that M2 follows M1 temporally and that M1 follows M2 temporally.

[3] The terms 'in series' and 'in parallel' referred originally to wiring arrangements in electrical circuits, but in recent times they have been extended to apply to the logic design of computer hardware. In the case of classic contradiction the two inconsistent meanings are offered as being timelessly true. The "simultaneity" and hence the "parallelism" results from this coincidence of two timeless meanings.

[4] Yes, the issue here is, in essence, the "constructibility", or otherwise, of linguistic gadgets. Let's face it, the majority of the "natural" paradoxes are also consciously constructed "gadgets", rather than spontaneous expressions arising from, or in, normal conversation.

[5] We know that the paradoxes are paradoxical, and that paradoxical statements are nonsensical. The simplest possible way to "solve" the problem will therefore be to say that the paradoxical statements are "nonsensical *because* they are paradoxical". It is of course an explicitly "scholastic" explanation. Tucker's (1963) solution was to blame the paradoxical statements for being "unverifiable". This is, I think,quite close to being another explicitly "scholastic" explanation:it certainly tells one nothing new.

[6] The "later" words (i.e. those after the first three) consist of mutually exclusive predicates or words (non-predicates) which, when taken in conjunction with words 1-3, fail to make sense.

[7] Thus the predicate 123N68-TRUE covers the case when none of the "later" words makes sense in conjunction with the first three. We shall call this the "miscellaneous" interpretation of the predicate 123N68-TRUE.

[8] The statement is certainly a 3/1 worder because its "later" words are either mutually exclusive predicates or non-predicates.

[9] The paradox arises from the fact that (A) must be either 1236-TRUE or 1238-TRUE or 123N68-TRUE, yet the effect of the radically adverse self-ascriptions is such that each of these alternatives implies one of the others. One might say that (A) could be 123N68-TRUE under the "miscellaneous" interpretation of that predicate. But in this case it would be, by virtue of its form, 1238-TRUE. Which is a contradiction, since the predicate 123N68-TRUE, whatever else it implies, excludes this. It should be pointed out that the rigour of the "paradoxical argument" here is of the same order as that which we find in the natural paradoxes. The paradoxicality of the natural paradoxes is a problem *at this level of rigour*. If one tries to apply the kind of hyper-rigour

customarily used in mathematical logic at this point, one does not "solve" either this superparadox or the natural paradoxes: in both cases the arguments which posed the problem simply disintegrate into a cloud of dust.

[10] This is the "miscellaneous" interpretation of the epithet 'tertio-dominant'.

[11] Because of the "comprehensiveness" of the three categories: hence the "miscellaneous" interpretation is ruled-out from the beginning.

[12] The convention adopted is that only *positive* statements are counted (in effect, 'not-A1' is regarded as equivalent to the disjunction 'either A2 or A3 or A4... or A*n*').

[13] This "formal" inconsistency does not imply that we are reverting to treating them as stable conclusions. What is meant is that they are prima facie formally inconsistent.

[14] In the natural paradoxes the "mechanism" is hard to recognise because it is so slight. In the case of the superparadoxes the "mechanism" stands out much more visibly.

[15] See the square bracketted comment on page 26, which identifies the technical form of paradox as being that of a *question*. [It is essentially a question we cannot answer.]

[16] Closure may sometimes be achieved by means of asymmetric definitions, which field "miscellaneous" clauses, like the last two superparadoxes discussed in this Chapter. However the Arthur superparadox and the permutation superparadoxes of Chapter 5 achieve closure without using asymmetric definitions.

A MAG-EDU PAPERBACK:

New Thinking about the Nature of Mathematics

ISBN 0 907669 20 4 ii+ pp.84 Published November 1992

A collection of essays by Philip Davis, Paul Ernest, Ray Monk, Warwick Sawyer, David Henley and Eric Blaire

Edited by Christopher Ormell

"Before we can improve mathematics teaching, we need a clear idea what mathematics is... I know of no better starting point for anyone seeking to make sense of our subject today"

Prof. Jeremy Kilpatrick, *University of Georgia*, USA

"It engagingly raises some issues in the philosophy of mathematics especially for those, particularly teachers, who may have been unaware of them."

Prof. Robert Thomas, Editor *Philosophia Mathematica*,

University of Manitoba, Canada

"This book does a valuable service in turning our attention to two fundamental questions... what is it to think mathematically and to use mathematics?"

Max Stephens, past President of the

Australian Association of Mathematics Teachers

"This book ...is timely in that it recognises that thinking about the nature of mathematics ultimately determines what mathematics is constructed, and how it is interpreted and implied by learners and practitioners at all levels."

Prof. Nerida Ellerton and Prof. Ken Clements

formerly of *Deakin University*, Australia.

£5* postfree from MAG, School of Education, UEA,

Norwich NR4 7TJ

*Cheques payable to "MAG-Ashby": add £1 for E.C., £2 for non E.C. airmail

CHAPTER FIVE
Superparadoxes of various kinds and amplitudes

THE EXTENT OF SUPERPARADOXICALITY

In this Chapter we look at the question of the *extent* and *shape* of the field of possible examples of superparadox. In considering such "evidence" of the possibility of dynamic contradiction, we are not, however, talking only about superparadoxes which have been fully formulated, defined and checked-out. There is clearly an almost unlimited *potential* for finding new superparadoxes of a considerable number of different types. We need to form a view, therefore, of the possible extent of this additional area of potential, as yet unformulated, superparadox ---if we are to achieve a realistic appraisal of the significance of the new area of fundamental logical misfunction.

Static, parallel, contradiction has been known and studied by logicians for about three millennia: in effect, inquiry into the possibilities of static, parallel contradiction has been coterminous with the study of "Logic" itself. The arrival of a fundamentally new kind of basic logical misfunction ---dynamic, serial contradiction--- deserves therefore to be treated with something approaching the same kind of rigour, mental concentration and comprehensiveness.

We are at present, however, only on the threshold of the earliest steps in such an inquiry. [1]

The point of such inquiry is, at the end of the day, to chart and identify *potential* logical hazards: to create a sense of "foreknowledge" ---of the kind which was so unexpectedly and alarmingly lacking during the Panic Years of the 1900s. Every variety of superparadox which we can envisage and classify *in advance* is one less unexpected pitfall for future logicians, software writers and mathematicians unwittingly to stumble into. We shall need, eventually, a consistent body of general principles delineating in comprehensive, formal terms what is a "well formed, oscillatory free" (WFOF) formula, and what is not. Such a body of logical principles will then become the "Watchdog" agency, responsible for keeping modern logical systems free from oscillatory nonsense.

Such is a remit for 'Dynamic Contradiction Studies', if we may call them that, in the next decade. The present chapter, however, can do little more than introduce the concept, initiate the debate, and offer a few guidelines for further study.[2]

CLASSIFYING SUPERPARADOXES

The general form of superparadox given in Chapter 4 will now be developed by considering the kind of *objects* (O) about which the superparadoxical question ("Does O possess characteristic C_1, C_2, C_3... C_n...?") can be posed. In each case we shall offer an abbreviation, or coding, for future reference.

In the cases of the Liar and the Arthur paradoxes the object O is of course a speaker's *mental state* the presence of which is indissolubly linked with the need to reinterpret the speaker's words. The Bean superparadox is, in effect, simply a generalisation of the Arthur superparadox. [3] [Code: MS]

In the cases of the Blackboard and 123-TRUE paradoxes, the object O is the *truth-status* of a statement. [Code: TS]

In the case of Grelling's and the Primo-Secondo-Tertio paradoxes the object O is the *classification* of a word or expression's self-application. [Code: WC]

In the case of the OMU and NBT paradoxes the object O is a *word* vis-a-vis the question how it should be read. [Code: WR]

In the case of Russell's Paradox the object O is a *set* vis-a-vis the question into which overall set clasification it belongs. [Code: SC]

In the case of the Barber Paradox the object O is a *person*: vis-a-vis the question whether he shaves himself. (Similar paradoxes can be constructed in terms of gadgets (e.g. a light which illuminates the switches of lights which do not illuminate their own switches) or even animals and plants. The general form of the question is into which action classification does the subject fall. We shall describe this category of subject in general as 'person/physical'.)[Code: PAC]

If we use the object under scrutiny (O) in the paradoxical question as our main initial basis for partitioning examples, this gives a scheme of six principal types of superparadox, though we have not yet, in fact, given examples of superparadoxes where the object of the paradox is a word (WR), set (SC), or physical subject (PAC).

A simple working coding system for superparadoxes consists of the code letters for the object-type, followed by the numeral representing the amplitude, followed by a letter signifying the type of mechanism.

Some typical "mechanisms" are: permuting letters denoted p
miscellaneous letter transformations denoted m
selective prolixity of speech denoted sp
selective self-application denoted ssa
selective self-membership denoted ssm

Thus the Arthur superparadox has the code MS3sp, the Bean superparadox (see note [3]) has the code MS10sp, the Barber paradox has the code PAC2ssa.

We shall begin by filling the gap (i.e. giving examples of the missing types) and then proceed to a short survey of the potential for "constructing" or "discovering" [4] new categories of superparadox beyond these six major types.

Probably the easiest superparadoxes to construct are those of the 'WR' type, i.e. those which exploit "peculiar instructions to read" one or more words in the paradoxical statement in non-standard ways. There are large numbers of such superparadoxes, and they do not normally introduce the kind of assymmetry associated with "miscellaneous interpretations", of the kind we saw in the 123-TRUE and Primo-Secondo-Tertio superparadoxes. So they are easy to construct, numerous, capable of endless variation and satisfyingly symmetric. We give two representative examples, the ONU superparadox and the PIN superparadox.

THE ONU SUPERPARADOX

The ONU superparadox begins with definitions of the three new words 'onu', 'nuo' and 'uno':[5]

'onu' means 'upside down'
'nuo' means 'by letters, in reverse order'
'uno' means 'normally'

The superparadoxical statement is:
"The last word of statement (WR3m*) should be read onu" [WR3m*]

The object of this superparadox is the final word of the paradoxical statement, 'onu'. How should it be read? By implication we are operating within a convention that allows us only three alternatives, those defined by the words 'onu', 'nuo' and 'uno'. [6]

Assumption (A) is that it should be read onu.
Assumption (B) is that it should be read nuo.
Assumption (C) is that it should be read uno.

It is evident that assumption (A) leads us to read the word upside down and thereby to end-up with assumption (B).

Similarly, assumption (B) leads us to read the letters in reverse order and thereby to end-up with assumption (C).

Assumption (C) of course tells us to read 'onu' as 'onu', thereby leading us back to assumption (A).

In this particular superparadox, WR3m*, the subject is a single word, and the form its radical reflexive self-ascription takes is that the *word as a whole* should be read upside down, not individual letters. There is not just a single superparadox of the type "WR3m", but a whole family of superparadoxes, based on various kinds of reflexive ascription mechanisms like word removal/replacement, letter removal/replacement, letter permutation and letter replacement.

THE PIN SUPERPARADOX

This superparadox, too, requires preliminary definitions:

'PIN' means 'in the order 1, 3, 2'

'PNI' means 'in the order 3, 2, 1'

'NIP' means 'in the order 3, 1, 2'

'NPI' means 'in the order 2, 1, 3'

'IPN' means 'in the order 2, 3, 1'

'INP' means 'in the order 1, 2, 3'

Now consider the superparadoxical statement:

"The letters of the final word of statement (WR6p*) should be read PIN" [WR6p*]

As in the previous example, the subject of the superparadox is the single word 'PIN' at the end of the statement. The question at stake is the order in which its letters should be re-assembled to be read. There are in this case six disjoint assumptions, A, B, C, D, E, F and we presume at the beginning that one of these assumptions must apply: there are of course no more possible permutations of the three letters of the word 'PIN'.

Each assumption, as before, leads to the next. So here we have a closed six-pole cycle around which the partial meaning of the statement oscillates. The mechanism in play may be simply described as one of word letter-permutation. The logical absurdity score of this superparadox is 63, see note [7].

THE RUSSELL EXTENSION SUPERPARADOX

As the primo-, secondo-, tertio- paradox is a natural extension of Grelling's paradox, so there ought to be a natural extension of Russell's paradox on similar lines. We shall designate such a superparadox SC3ssm*.

We begin by defining an ordered set of three comprehensive, disjoint sets, A, B, C. We shall call such an ordered set of sets a *trio*:

{A, B, C}.

Now we may distinguish trios of different kinds. Sometimes a trio, which is a set, is a member of its own first set: in other words, the trio {A, B, C} satisfies the membership criterion for the set A. We shall call such a trio a first-qualified trio.

In a similar way, the trio {A, B, C} may be a member of its own second set, B. In this case we call the trio a 'second-qualified' trio.

In all the other cases, that is, where the trio {A, B, C} is a member of its own third set, C, or where it not first or second qualified for any other reason, we shall call the trio a 'third-qualified' trio. (Notice that this includes the "miscellaneous interpretation" indicated by the phrase 'for any other reason'.) [8]

We now define three categories of trios: (i) the set of all first-qualified trios, (ii) the set of all second-qualified trios, and (iii) the set of all third-qualified trios. We designate these three categories (sets) SFQ, SSQ and STQ respectively.

Now let's address the question:

"Into which category does the trio {SSQ, STQ, SFQ} fall?" (SC3ssm*)

There are three possible suppositions, that it belongs to SFQ, that it belongs to SSQ or that it belongs to STQ. We shall call "it", i.e. the special trio mentioned in SC3ssm*, the 'Extension' trio.

Assumption (1)

The Extension trio belongs to SFQ. If so, it is a first-qualified trio and therefore belongs to its own first member SSQ.

Assumption (2)

The Extension trio belongs to SSQ. If so, it is a second-qualified trio and therefore belongs to its own second member STQ.

Assumption (3)

The Extension trio belongs to STQ. We know that it must belong to one of the three sets SFQ, SSQ and STQ, so the only possibility is that it belongs to its own third member, SFQ.

This is clearly a necessary extended oscillation and a new superparadox in which the partial meaning of each of the three paradoxical statements

"The Extension trio belongs to SFQ"
"The Extension trio belongs to SSQ"
"The Extension trio belongs to STQ"

cycles steadily round the set of assumptions listed above.

It is immediately evident that the pattern of superparadox exhibitted here may be generalised to have as its subject an ordered *n*-tuplet of category sets of the same kind as those defined above. We shall designate this generalisation of the Russell Extension superparadox SC*n*ssm*. Its absurdity score is $2^n - 1$ as before.

THE NURSE ATKINS SUPERPARADOX

This superparadox is an extension of the idea of the Barber paradox.

On an island with only one doctor (Dr Jones) his authority on medical matters is absolute. One day a serious infection struck the inhabitants and Dr Jones began to give out phials of vaccine to patients with the infection to inject themselves with the vaccine every 8 hours starting at midnight, 8 am or 4 pm. (Needles would be issued for the injections at the Island Dispensary only at these hours.)

The next day Dr Jones had planned a lengthy fishing trip, but before he left he realised that the total supply of the vaccine was limited. So he left strict orders:

<u>FOR NEW PATIENTS WITH THE INFECTION</u>

EVERYONE WITH THE INFECTION MUST HAVE REGULAR
8-HOURLY INJECTIONS
ALL INJECTIONS WILL BE AT MIDNIGHT, 8 am or 4 pm

"If you would have given yourself your first injection at midnight, wait instead until 8 am when Nurse Atkins will give you your first injection.

If you would have given yourself your first injection at 8 am, wait instead until 4 pm when Nurse Atkins will give you your first injection.

If you would have given yourself your first injection at 4 pm, wait instead until midnight when Nurse Atkins will give you your first injection."

Nurse Atkins will personally supervise all injections and see that everyone with the infection has regular 8-hourly injections once they have begun the course.

These instructions seemed to be perfectly plain and Nurse Atkins promised to do everything which the rules said she would do.

Unfortunately some hours later Nurse Atkins discovered that she herself had contracted the infection. If she would have given herself the first injection at midnight the rules enjoined her to wait until 8 am; but if she would have given herself the first injection at 8 am the rules enjoined her to wait until 4 pm; and if she would have given herself the first injection at 4 pm the rules enjoined her to wait until midnight. But injections could only be given at midnight, 8 am or 4 pm, so she could not receive the injection at all. On the other hand the rules said that she must have the injection.

The logical force of this superparadox (PAC3ssa*) is clearly weaker than that of the more cerebral superparadoxes. In this it follows the pattern of the Barber paradox, which is only a "paradox" insofar as the village barber cannot keep strictly to what seem initially like perfectly reasonable rules. (In practice he obviously shaves himself and reasons that the rule that he "shave those who do not shave themselves" applies only to non-barbers.) Similarly, Nurse Atkins will give herself the injection at the earliest opportunity, reasoning that the rules left by Dr Jones apply only to non-nurses.

Some have said that the "solution" to the Barber paradox is that no such barber could exist (i.e. a barber strictly observing the rules). The corresponding "solution" here would be that no such nurse (i.e. one strictly observing the rules) could exist.

The seven specific superparadoxes considered in this and the previous chapter exhibit dynamic contradiction, or oscillation, round a cycle of 3 or more poles, but in each case by means of a different mechanism. Many of them have been given in the case where the amplitude is 3 for reasons of convenience. There is nothing special about cases where the amplitude is 3, except that they can be explained more simply and neatly than cases where it is 4 or 5 or more... The important point is that a pattern has been established which the reader can easily generalise. Thus an amplitude 5 version of the Russell Extension superparadox (SC5ssm) will introduce quintets of sets (ordered sets of five sets) and five categories of qualification ("first qualified", "second qualified"... through "fifth qualified"). Nurse Atkins' instructions might have mentioned four injections per hour, at 15-minute intervals, thus producing a superparadox with the code PAC4ssa.

The self-ascriptive mechanisms involved in the seven stated superparadoxes are only a small *selection* of the mechanisms which could, potentially, be employed. They have ranged across such things as selectively discarding words, selectively applying words, reading upside down, permuting letters... It is evident that there are other self-ascriptive mechanisms which might lead to the emergence of superparadoxes, e.g. selectively adding-and-subtracting words, general selective transformation of words, conservative transformation (moving, permuting) of words or words-and-letters, removal of selected letters... Such mechanisms, among others, clearly have the necessary power to create superparadoxes of the future.

NOTES

[1] Hopefully the inquiry into serial, dynamic logic will not take as long to develop as the 3000 years mentioned in the text for parallel, static logic! (The metallurgy of Titanium was developed, using modern methods, in about ten years ---which compares quite well with the length of time needed for the human race to take the same steps in relation to bronze and iron.)

[2] In most cases only *outlines* of the superparadoxes introduced in this chapter have been given. The details are less important, and can reasonably be left as an exercise for the reader.

[3] Mr Bean suffered from a multiple personality disorder in which he flipped in an apparently random manner between ten different, quite distinctive, sub-personalities. Like Arthur, though, he was prone to include large numbers of redundant words in his sentences, the pattern (of "selective prolixity") changing from sub-personality to sub-personality. On one occasion he commented to his doctor: "I am either in sub-personality two, or sub-personality three, or sub-personality four... or sub-personality nine, or sub-personality one!". It was both a truism and a superparadox (MS10sp), because close inspection showed that this innocent remark led to the apparent "logical necessity" of 1023 different mutually contradictory conclusions about his mental state!

[4] The choice between saying that we "construct" such superparadoxes or that we "discover" them is surprisingly lacking in tension here. It seems natural to say that we "construct" them, but we can't do this, obviously enough, until we have "discovered" a pattern which works.

[5] Here, as we remarked on page 28 Note [15], apparently slightly arbitrary conventions are always a necessary preliminary to any paradox.

[6] The star is simply a distinguishing mark to enable us to link the coding to this particular superparadox of the general type WR3m.

[7] It may be noted that this superparadox enables us nominally to "prove" 63 different combinations of mutually inconsistent propositions, so its 'logical absurdity score' is 63.

The 63 may be coded thus: A , B , C , D , E , F , A & B, A & C, A & D, A & E, A & F, B & C, B & D, B & E, B & F, C & D, C & E, C & F, D & E, D & F, E & F, A & B & C, A & B & D, A & B & E, A & B & F, A & C & D, A & C & E, A & C & F, A & D & E, A & D & F, A & E & F, B & C & D, B & C & E, B & C & F, B & D & E, B & D & F, B & E & F, C & D & E, C & D & F, C & E & F, D & E & F, -(A & B), -(A & C), -(A & D), -(A & E), -(A & F), -(B & C), -(B & D), -(B & E), --(B & F), -(C & D), -(C & E), -(C & F), -(D & E), -(D & F), -(E & F), -A, -B, -C, -D, -E, -FA & B & C & D & E & F. In this list 'A' means 'assumption A', 'B' means 'assumption B', etc. '-A' means 'B & C & D & E & F' etc., -(A & B) means 'C & D & E & F', etc.]

[8] So there is an asymmetry here similar to that of the 123-TRUE and primo-secondo-tertio superparadoxes.

CHAPTER SIX
Conclusion: the possibility of a consistent, typeless, nonstratified mathematics

IN LOGIC, AS IN MEDICINE, accurate diagnosis is more than half of the story. We have seen in exactly what the alarming malady of unannounced paradoxicality consists. We turn now to consider, first, the form of a cure, and, second, the prognosis for mathematics, logic and programming *after* the cure.

Let's begin by recapitulating the malady.

We have seen that the problem (the malady) which bothered the philosopher-logicians of the 1900s was twofold: (a) how to *explain* the paradoxes which had erupted so rudely and unexpectedly into the circle of apparently precise concepts which ---it was generally agreed--- were needed to create a logical basis for mathematics, (b) how to *foresee* in advance new eruptions of similar nonsensicality.

The two aspects of the problem were obviously inter-connected. The sign of a good explanation of the paradoxes would be that it enabled one to foresee future instances of the same sort of deep-seated nonsense.

This was paramount. Mathematics could not be expected to operate indefinitely under the constant threat of new unannounced logical disasters of this kind. The need for effective prescience, for advance warning of future outbreaks of paradoxicality ---this, it was agreed, should be given the highest priority.

A "good explanation of the paradoxes", however, would have *first* to show that they resulted from the breach of some very obvious, self-evident, logical principle. The number of places you could search for such a "breach" was quite limited. Russell searched these places, with great thoroughness and impressive concentration, for three years. He found nothing. This was disconcerting. It was a problem of an intensely *baffling* kind, because there appeared to be no locus whatever of logical misadventure. There was no point onto which you could place your finger, and say with obvious correctness, obvious validity, "The argument goes wrong *here*!".

Yet no one had any sense that the mechanisms of the paradoxes were hidden. It was like a conjuring trick, in which everything is done in the open: yet at the end of the trick a paradoxical rabbit suddenly appears where none had been before!

In retrospect we may see that Russell's initial judgement, which led directly into the main paradox (Russell's Paradox), was right. You had to take Cantor's original concept [1] of a 'set of everything' pretty seriously. To lose the "set of everything" would be a body-blow to transfinite theory, because abandoning it would be a tacit admission that the argument for the "actual infinite" ---on which the whole transfinite edifice ultimately rested--- was not valid. In the event Cantor went mad, and Russell, having opened a pandora's box of trouble, seemed quietly to forget that the existence of Cantor's set of everything was so pivotal to the status of transfinite theory.

Today, at last, we are beginning to recover from what Poincare described so long ago as the "disease" of believing in the transfinite. We are, at last, beginning to work free from the intense mesmerism exerted by what Hilbert called the "the paradise that Cantor has opened to us". So, unlike Russell in 1901, we *can* today abandon the concept of the 'set of everything' with few qualms.[2]

The main conclusion of the inquiry recorded in this volume is that the reason no proper explanation of Russell's Paradox was ever found, is that it breached a principle involving a concept no one had even considered as a possibility: that of *dynamic* contradiction, that of a statement which fundamentally stultified itself serially, rather than in parallel.

To get into the position where one could even *begin* to see the possibility of such a new variety of contradiction, one had to see that the neo-platonic presupposition ---that logic and mathematics dealt only in timeless truths and timeless meanings--- was false. It tries to see the world *sub specie aeternitas*. But we cannot see the world in this way, because our very being is rooted in consciousness, which implies the passing of time. We must always look at the world, including the man-made world of mathematics, from some particular "now".

Of course our logic and mathematics is not tied to any particular "now". Its results were, are, and will always be, "timeless" in the sense that *any "now" will do*. So the results of logic and mathematics are undoubtedly "timeless", but they are not "nowfree". One can no more do logic and methematics outside the framework of some particular "now", than one can do coordinate geometry without an origin.

Logic, in a word, needs to be "now-rooted". Its truths may be timeless, but its concepts need to assume the existence of a now.[3]

Dynamic contradiction occurs when a partial meaning M_1 supercedes another, inconsistent, partial meaning M_2, and is then itself superceded once again by M_2... and so on, Partial meaning, in other words, is "meaning in the active process of gestation", or if you prefer, "what we are making of the meaning of this item *now*".

The logicians of the 1900s were trapped in an iron platonic mindset, from which they simply overlooked the possibility of partial meaning. They simply did not see that "(L) is true" or "the ordinary set is a member of the ordinary set" possesses, at most, only partial meaning.

And it has, we know, taken nearly a century of valiant, committed, impassioned argument (not to mention the heartbreak implicit in attempting repeatedly and hopelessly to assault an impregnably entrenched status quo) to break that iron mindset. Wittgenstein tried to throw semantic bombs at it ---which blew up in his face.[4] Popper berated its historicist implications. Ryle hammered it, producing a crack. [5] Lakatos began to prise it apart. [6] Putnam, Kline, Sawyer, Davis and Hersh, Hofstadter ---and many others--- gradually changed the climate of debate, gradually pushed platonism off centre-stage, insisting that mathematics must first make sense in *human* terms.

All this work has effectively cleared the ground, enabling us to begin once again the task of creating a secure logical basis for mathematics. But the meaning of the word 'logical' is not quite the same as it was in 1901. We are now at last able to see that this logic must embody certain elements never envisaged in 1901, namely, the-presence-of-a-now, the logical possibility of mathematically indefinable sequences ("Showdown sequences")[7], partial meaning and dynamic (serial) contradiction. Such a logic will be a 'supple' logic, [8] as opposed to the paralysed nowfree logic of 1901. [9]

SUPERPARADOXES AS EVIDENCE

The numbers of new, hitherto unsuspected superparadoxes which we can find, once we have taken on board the idea of serial contradiction, do not add-up to a new crisis in logic: they simply confirm the diagnosis. They show that the scientific hypothesis, *that "serial contradiction" is logically possible*, has potentially observable consequences ---namely that superparadoxes ought to exist. Such a prediction of a new logical phenomenon is not a commonplace event. What it claimed "ought" to exist, was, initially ---it is important to remember--- only an *hypothesis*. (A falsifiable hypothesis indeed ---because determined construction efforts might have revealed no such logical monsters.)[10]

Finding what the hypothesis predicted is, I submit, highly significant.

It is an example of the basic procedure by which modern science establishes its case: by first making strange, unobvious, hitherto wholly unsuspected, *novel predictions* springing from (derived from) a new *hypothesis*: and by, second, the subsequent verification of these "strange, unobvious, hitherto wholly unsuspected predictions".

This is the modus operandi of, the root magic of, science. This is what impels otherwise sceptical people to sit up and take notice.

This is what turns an "hypothesis" into an "explanation".

In the present case we have applied this procedure to the problem of the paradoxes. First, an hypothesis about the source of the paradoxicality of the paradoxes was offered: one which claimed that they were to be regarded as examples, not of static contradiction, but of dynamic oscillation of partial meaning.

Then this "hypothesis" was driven to produce the kind of "strange, unobvious, hitherto unsuspected" predictions mentioned above, viz. that superparadoxes ought to be constructible!

The actual discovery of the superparadoxes thus plays the role of the result of a crucial experiment in science, as when the products of combustion were weighed and found to be heavier than the original substance before combustion, as when light from stars in line with the Sun was displaced during the total eclipse of 1919.

In recent times Kuhn (1960) has pourred cold water on the notion that such "crucial experiments" are really "crucial". They are, it is true, never as compelling for people thoroughly trapped behind ideological spectacles as one might, optimistically, initially expect. (Descartes, for example, totally failed to "see" the significance of Pascal's brilliant series of experiments which verified the air pressure explanation of the Torricellian experiment.[11])

The human mind is not so easily won over to a new point of view. Yes, but, as Lakatos pointed out, these crucial experiments remain as *landmarks* in the development of science. They are exceedingly significant moments, which serve to establish the "progressive" quality of the problem-shift derived from the new account. They may not "compel" belief, but they are pretty ominous signals for anyone inclined to cling to the outmoded point of view.

A REFERENTIALLY UNSTRATIFIED MATHEMATICS

The idea that mere *reference* was sufficiently weighty and sufficiently precise to become the basis of a secure stratification of mathematics has, as we have seen, broken down. Of course mathematics, regarded as a humanly constructed set of conventions, embodies many distinctions, and hence potential "stratifications". The basic method which we apply throughout mathematics to simplify and clarify mathematical problems (mathematical modelling) [12] consists, in effect, in finding convenient temporary stratifications from which to attack them. What has gone, then, is not "stratification" as such, but an enormous multiplication of rigid, flimsy, scholastic stratification allegedly needed because of the over-riding requirement to exclude the paradoxes.

It is stratification "by reference" which has been shown to be unnecessary. This was a kind of stratification which erected new logical barriers between natural numbers, integers, rational numbers and reals. Of course there are circumstances in which we *want* to distinguish numbers defined on these different bases. But there are far more circumstances in which we don't! This was where the concept of stratification-by-reference made life difficult. Logicians were so alarmed about what *might* happen if stratification-by-reference were breached, that they effectively decreed that it should be universally observed.[13] It was the chief legacy of the Panic which followed the discovery of Russell's Paradox. The result is all around us: a guilt-ridden mathematical culture, which should have risen to the challenge posed by the microchip revolution, but never did.[14]

Hopefully, supple logic will make a difference. Using supple logic as the basis for mathematics allows us to handle the subject in a neater, more fruitful, more relaxed way. The miasma implied by an infinite hierarchy of meta-languages disappears. There is *one* language of mathematics, it is called 'mathematics'.[15] It is about "things" created by human convention, e.g. 1, -1, e, i, 0... ---*honorific existents* which "exist" because, and only because, they have been historically *recognised* by the consensus of those with the most highly developed mathematical talent.[16] It was this world of convention which was messed-up by the introduction of stratification by reference. And it is this world of convention which can be *cleaned up* once we accept the need for the new fundamentally supple concepts of now-rooted logic.

NOTES

[1] "Original" concept, because of course Cantor had reluctantly to abandon it after it transpired that it led to a paradox (Cantor's paradox).

[2] The crucial fact is that it has become clear that all the real numbers which appear, or have ever appeared, in mathematical research journals and books belong to a few *countable* classes of real numbers. The supposed immensity of real numbers has never been found. This has given credibility to the theory that it is not *numerousness* which creates uncountability among real numbers, but the incompletability of the totality. Many commentators still speak, however, of the transfinite calculus in tones which imply that they think it is here to stay, e.g. Kilmister (1967), Tiles (1989), Penrose (1989). Its fundamental logical basis, however, disappeared once the concept of the set of everything had been abandoned. Borel had already recognised this by 1898.

[3] We need 'before' and 'after' in defining limits, potential infinity, convergence, etc. These concepts are literally meaningless unless we are working in a context in which we recognise that a "now" exists.

[4] I am referring to his "Seminar War" with Alan Turing at Cambridge in 1939. See Ray Monk (1992) for a detailed account of this.

[5] By showing (in his lectures on "Meaning") that platonism rested ultimately on the naive view that all meaning was created by acts of naming. This is not, however, mentioned in his paper on meaning in *The Theory of Meaning* (Ed. Parkinson 1968). pp. 109-116.

[6] Lakatos's original papers 'Proofs and Refutations' (1962-64), one might say, vividly highlighted the error of trying to think about mathematics in a "now-free" way.

[7] See Ormell 1992b.

[8] A term introduced in *New Thinking about the Nature of Mathematics*, 1992, pp. 8.

[9] The required change is not, of course, limited to these four new concepts. They, in their turn, begin to modify our concepts of 'timeless', of 'mathematically true', of 'meaning', of 'logical possibility', etc. All this is encompassed in the new idea of a now-rooted and supple logic.

[10]Much as the hypothesis that there should be a second order of "complex" numbers, turning the Argand Diagram into a three dimensional display, and composed of entities of the kind a + *i*b + *j*c, turns out, upon investigation, to be invalid. There are not, and could not be, such numbers.

[11] But we have no warrant whatever to assume that he could *never* have seen the point. Yes, there are walls of initial incomprehension between conceptual frameworks, but much of the point of mathematics teaching is to train students to hop routinely over them: to be logically gymnastic. 'Incommensurability' would presume impossibility.

[12] A highly important point made by Atiyah (1977) in his talk to the Karlsruhe ICME: in pure mathematics "modelling" is often called "finding a representation".

[13] Of course it never *was* universally, or even often, observed. (It may be noted that neat forms of restrictive set theory are still, to a minimal degree, stratified.) But the fact that mathematicians have had *routinely* to flout the official logic of their science has hardly set a good social example (i.e. for the intelligent subset of society). It has induced feelings, perhaps subliminal but real for all that, of guilt. It has been a self-inflicted wound, damaging both the self-confidence and the sense of moral purpose of mathematics, and during a period in which unprecedented demands were being placed upon the subject.

[14] In the sense that the microchip has vastly increased the power, and hence the relevance of, mathematics in society. One would have expected a commensurate increase in the number and quality of young people coming forward to take-on the exciting challenges thus revealed. This has not, broadly speaking, happened. Rather the reverse: many young people have met the now thoroughly muddled, guilt-ridden culture of mathematics, and have rejected it, as a quagmire into which they are not anxious to step.

[15] I have suggested elsewhere that mathematics needs a 'Version of Maximum Rigour' (VMR) to act as the language of final appeal concerning the validity of proofs and concepts. Such a VMR, will, I suggest, be an object-centred formalism: that is, it will consist of meaningful statements about ultimately uninterpreted symbolic expressions. Viewed in this way the VMR of mathematics is not a vast collection of meaningless formulae, but a vast collection of truths (in the ordinary, i.e. Lucasian, sense) about what you can legitimately do with uninterpreted symbolic expressions. So the need for even a single 'meta-language' disappears ---though the idea of adopting temporary meta-languages is of course still available to us as an heuristic, modelling, device.

[16] The present author introduced the concept of 'honorific existent' in his draft monograph *Looking for Meaning in Mathematics* (1985): a definitive MAG edition is planned for 1994 .

Some examples of valid self-reference

That there are thousands of examples of self-referential statements which make perfectly good sense, such as this one, in ordinary *use* in English and other languages may be wholly self-evident. They happen, unnoticed and unrecorded, all the time. But in setting-out the case for the validity of self-reference in a fully measured and organised way, it seems appropriate to compile a short list of such examples. The list will not, in itself, convince anyone who is *determined* to cling uncritically to the once fashionable Official Story that self-reference is logically flawed, but it may occasion the earliest beginnings of doubt, and it may afford some support to those who have crossed the threshold of rejection of the Story, but are still slightly unsure whether they are walking on solid ground. [NOTE: the qualifications in *italic* in some of these examples are often omitted.]

Singular Statements

1. I realise that everything I say is being recorded. (A comment made by a person being interviewed with a tape recorder.)
2. Just testing! (A person speaking into a microphone for the first time after a temporary electrical breakdown.)
3. I declare this meeting closed!
4. I nominate John Smith for the job!
5. This is the first thing I have said for five years. (Said by a man on completing a "mute protest" which lasted for five years.)
6. I hope I am making sense. (A person who is very unsure whether he is being understood.)
7. I realise that you are not listening to a word I say.
8. I am very pleased to be addressing you here tonight.
9. I say, you look as if you are lost.
10. I thank you, doctor, so much for restoring my power of speech! (A patient thanking her surgeon for completing a successful operation on her vocal chords.)
11. I hope you will be able to make sense of my imperfect English. (A person speaking uncertainly in English.)
12. We who publish newspapers are proud of our record during the Gulf War.
13. I am no orator, but I am trying to tell you the facts.
14. You know that you can trust what I say to you.
15. I make exactly the same remarks to each year's intake. (Part of an address to this year's intake, which has also been used for the last five years.)
16. I must fully admit that I *do* tend to over-emphasise what I say more than 200% of the time.
17. We in the BBC, do our best to try to read the news clearly. (A BBC newscaster reading the news clearly.)
18. We hesitate to speak after the end of the programme until we are off the air. (Whispered in the studio shortly after a radio programme had finished.)
19. This is the Voice of America bringing you news and comment from the United States. (A radio broadcast from the United States.)
20. I do, frankly, admit to a more than slight tendency to be longwinded, so that you may find that, on occasion, I use comparatively lengthy statements and roundabout locutions to convey (in other words, communicate) meanings which a more laconic speaker would convey in a fraction of the time.
21. I, who have been as deeply involved as anyone in these disputes, say that we should forget our differences.

Imperatives

22. Disobey orders on principle ---*with the exception of this one*! (Given as an order.)

23. Write these words 200 times before next Tuesday! (Set as "Two Hundred Lines" by an old-fashioned schoolteacher.)

24. Listen carefully to this!

25. Always do what you are told!

26. Do what I do, not what I say ---*apart from on this occasion*!

27. Everybody stop speaking!

28. If you must speak,*whisper* while you are in the Library! (Whispered in the Library.)

29. Write down everything I say to you!

30. Ignore advice given in books, *but not on this occasion*! (Advice given in a book.)

31. Always begin an order with the word 'Always'!

32. Always give your instructions loud and clear! (Instructions given loud and clear.)

33. Always set up a line of type in the same font unless given specific instructions to the contrary! (A line of type set in the same font: part of a *Printer's Manual*.)

34. Always complete a printed imperative with an exclamation mark!

35. Make your orders short!

36. Write any order you are given in this notebook!

Questions

37. Can you hear me on the back row? (A speaker addressing the back row.)

38. What is the point of Question Time? (Asked on the BBC Programme Question Time.)

39. What is the most important question we must ask ourselves?

40. Do we ask too many questions?

41. Is there any point in asking rhetorical questions? (Rhetorical question asked by a speaker.)

42. Can we ask about anything? (Question from an intake who have just been told to ask questions.)

43. Can I ask you your opinion, Sir? (Reporter approaching a celebrity.)

44. Can you read this? (Sign displayed in small print in an Optician's window.)

45. Can you say these words after me? (Language tutor speaking to a group of students.)

46. Can you make sense of my questions? (Asked by a person who just has asked several questions but has yet to received a coherent reply.)

47. Have you got time to listen to me?

48. Have you got time to write this down?

49. Have you written down all the questions you were asked in your notebook?

50. Do you know where you are? (Question asked on a signpost.)

51. Do you often watch this programme? (Part of an appeal for funds on a Public Service TV programme.)

Generalisations

52. The first word of a sentence in written English should begin with a capital letter.

53. The final word of an ordinary, indicative sentence in written English should be followed by a full stop.

54. Altering the order of words in a statement invariably produces a change of emphasis if not a complete change of meaning.

55. A sentence with a subject, a copula and and adjectival phrase is said to be in 'subject-predicate' form.

56. Every sentence in correct English contains a verb.

57. Any sentence containing more than eight words will be described, for the purposes of this manual, as 'long'. (Sentence in a *Printer's Manual*.)

58. Every stentence which does *not* contain a spelling mistake in a printed text generates a minute additional increment of confidence in the Editor's orthography.

59. And remember, my boy, sentences beginning with 'And' are to be avoided ---*apart from this one*. (Advice given in a letter from father to son.)

60. The meaning of a statement is generated partly by the *words* spoken and partly by the *reputation* of the person who speaks. (A speaker explaining methods of oratory in Central Park, N.Y.)

61. All language presupposes the linguistic conventions of its time.

62. Everything we say is a product of our culture.

63. A statement only makes sense to a listener who brings the appropriate linguistic conventions to bear on it.

64. English teachers prefer sentences which do not contain spelling mistakes.

65. English is an expressive language.

66. All the statements made in this programme have been vetted by the Iraqi Government. (Statement made in a TV programme about Saddam Hussein.)

67. Having got to the half-way point in a fillibuster without fully broaching the argument you will bring to bear on the subject, start using self-referential statements like this one, which are of above-average length, and which introduce a confusing mixture of ideas, long words, subordinate clauses and idioms.

68. Words speak less loudly than deeds.

70. The first sentence of a book always tells you much more than you might imagine about the book and its author. (First sentence in a book.)

71. Linguistic insights should always be expressed in as general a form as possible, rather than by means of particular examples.

72. Generalisations should always be avoided ---*apart from this one*.

Philosophical and Logical Examples

73. Whereof one cannot speak, thereof one must be silent. (From Wittgenstein's *Tractatus*.)

74. Philosophical arguments often make claims about the nature of philosophical arguments.

75. The meaning of the word 'use' is generated by its use.

76. All philosophical generalisations are problematic.

77. It is essential in making philosophical statements to use words extremely accurately.

78. The chief method of Philosophy is a technique for clarifying meaning.

79. To say that a statement is true is to say that it corresponds with the facts.

80. To say that a statement is true is not to add anything to the assertion of that statement.

81. To say that a statement is true is to say that it can be trusted. (J. R. Lucas 1969)

82. To say that a statement is true is to say that it coheres to the greatest possible extent with all the other statements we believe to be true.

83. A statement of the form "p and not-p" is self-contradictory.

84. A statement of the form "p or not-p" is logically necessary.

85. Genuine statements of logic say nothing about the nature of the world.

86. Tautologies tell us nothing about the world.

87. To say that a statement is logically necessary is to say that we cannot conceive of any circumstances in which it would be false.

Miscellaneous Examples

88. There is nothing baffling about this statement.

89. Don't let this statement worry you, it is not supposed to say very much.
[Addressed to a person who develops anxiety on being spoken-to.]

90. Don't tell your friends about this statement, it is not worth mentioning.

91. Don't memorise this statement, it is not worth the effort.

92. This statement is not to be read in a highly critical frame of mind.

93. Don't read the last sentence of this review, it does not give you any extra information.
[Last sentence of a review of Douglas Hofstadter's *Metamagical Themas* (1985).]

94. Don't read this statement if you have to strain your eyes to see it.

95. This statement should be read in a relaxed way.

96. Reading this statement will help you to relax.

97. It is not necessary to sit down to read this statement.

98. This statement is not really suitable as the opening sentence of a novel.

99. This statement is not really suitable to be declaimed on stage.

100. This statement is quite suitable as the final statement in a set of self-referential statements of ordinary language which palpably make sense.

Some Varieties of Superparadox: Appendix B

Sentence versus Statement

WHEN DISCUSSING sentences as "gadgets" it is obviously important to make a clear distinction between sentences on the one hand and statements on the other. The track record of philosophical discussion on this issue, however, is hardly reassuring. Some theorists write as if "sentences" could be "true" or "false", which implies that a "sentence" can have "a specific meaning". Others, e.g. Russell and the Logistic school, have introduced the term 'proposition' as something rather more abstract and "timeless" than a "statement".

The aim of the present essay is to use words as closely as possible to what may be described as the "careful use of ordinary language".

For the purposes of the present essay a sentence will be described as a sequence of words beginning with a capital letter and ending with a full stop or exclamation mark which observes the normal rules of grammar and syntax.

EXAMPLES

The following are sentences:

The cat is on the mat.

Now is the time for all good men to come to the aid of the party.

The clock stood at ten to twelve.

Mary had a little lamb.

For the purposes of the present essay "sentences" do not convey meaning, as such, though they "have" meaning in the potential sense that they might used to make statements which had meaning. Thus, if I go into a classroom and see the words

'The cat is on the mat'

written on the blackboard, I do not (usually) expect subsequently to learn that the words 'The cat' referred to a specific animal or that the words 'the mat' referred to a specific floor-artefact in a specific building or that the words 'is on' referred to a specific time. I "see" the words as constituting a sentence. It is not telling me, or anyone else, something. It is not conveying a message. It is neither true nor false. It is simply an ordered set of six words

begining with a capital letter and ending with a full-stop, which observes the rules of grammar and syntax, and which *could* be (though is not being) used to make a statement.

On the other hand, I might have made a prior arrangement with Mary Brown ---whose classroom this is--- that she would leave a *message* for me on her blackboard about the whereabouts of her china "cat" which I am going to take to the auctioneer. I shall call in at the classroom on my way to pick up the cat. By the phrase 'the mat' I understand a particular place near the frontdoor in the outer-hall of Mary Brown's house. In these circumstances the words on the blackboard

'The cat is on the mat.'

are telling me something; they are ---I hope--- true; they constitute a *statement,* made by a specific person (Mary Brown) with the intention of informing me of the whereabouts of a specific object (Mary's Brown's china cat).

A "statement", then, is a sentence-in-use, a sentence with (a) a determinate author (which may be an individual or a group of people) "made" with (b) the intention or *apparent* intention of conveying (c) a message of some sort to (d) the person(s) to whom it is addressed. It was "made" (e) at a specific time, at (f) a specific place. The use of demonstratives presupposes (g) that the person(s) addressed will "take" the presumed, familiar references unambiguously. [(b) enables a statement to be "meaningless".]

In my opinion the *paradigm* use of language is the spoken word. A statement made orally, I suggest, makes a bigger impact on the mind of the listener than anything he or she has first to *read.* In other words, statements made orally gain visibly from the unmistakable physical presence of the factors (a), (d), (e), (f). Additional signs of the author's *intention* (b) and the *point* of the message (c) are also often at hand. In many cases the intended references of the demonstratives used (g) will be in the addressee's field of vision. The only serious deficit which an oral statement suffers is that those to whom it is addressed may *forget* the exact words shortly after they are spoken. This is often remedied in teaching situations by the teacher writing major statements on the board at the same time as saying them.

Statements exist

Once a statement has been "made", it *exists* as a statement. (There are, of course, verbal formulae for conventionally "retracting" a statement which a person may have made inadvisedly. The effect of such conventional "retraction" is never , however, quite the same as if the statement had not been made in the first place.) This "existence" of a statement which has been made, continues virtually indefinitely, at least while there are people who remember it, to whom it conveys a definite meaning --- not necessarily the addressee(s)--- and for whom it has some definite significance.

Such a view of the nature of *statement* ---as a human institution which predates civilisation--- then has to be extended to take into account the effect of writing and printing.

Once written down, a statement's continuing existence is greatly reinforced: the printed or written (or electronically stored) words ---as it were--- "mark" its continuing existence physically. They also permit "returnability", as when an addressee or listener "returns to the words to re-read them". The effect is further amplified when the words are duplicated by copying, printing or electronic transference: now there are multiple instances ("tokens") of the sentence-which-expresses-the-statement in existence.

However, it is a consequence of this process of duplication that the reader's experience of being *vividly aware* of factors (a) through (g) above is somewhat diminished. So the statement is now identified closely with the sentence which expresses it, and the essential factors (a) through (g) which were (all) originally responsible for its being a "statement", rather than a mere "sentence", can be easily forgotten and/or underestimated. The longevity of a statement's existence is enhanced by writing, printing and copying, but at the expense of the unique, fresh, personal, flavour which that statement originally possessed. Oral statements typically possess the flavour without much longevity; written statements typically possess the longevity without so much of the flavour.

Propositions

The point of the term 'proposition' is that when different statements "say the same thing" it is convenient to be able to say that they express the same proposition. It would be a mistake, though, to treat 'propositions' as a more satisfactory logical unit than statements, because it is always a matter of debate whether two genuinely different statements (i.e. statements made by uttering genuinely different sentences) do "say the same thing". Essentially what counts as "the same thing" differs according to the context within which it occurs ---it is *relative* to what is at issue when the statement is made. Nidditch (1960) defines a 'proposition' in *mathematics* as "a statement that is either true-or-false" p.284.

The Meaning of a Statement

I have argued elsewhere that the meaning of a statement is what it legitimately and bi-conditionally leads the reader/listener to expect in his/her experience. This expectation is of course created entirely by the *use* of the words which make up the statement. (The term 'bi-conditionally' covers the two conditions that the reader/listener must (i) *believe* it (ii) *consider* the area of experience to which it refers. No expectation will be created unless these conditions have been met.)

More accurately this "use-expectation" account of meaning applies to the *hard* or *accountable* element in the meaning of a statement. There is also a "soft" (or poetic) element, a kind of "splash" of associations which are too loose to constitute definite "expectations" of experience, but which can set-off trains of aesthetic feeling and reverie.

Type versus Token

We shall use the familiar distinction between 'type' and 'token'. The sentence 'The cat is on the mat' appears twice above. In once sense, then, thees words on p. 48 and p. 49 constitute the same sentence (type), but in another sense they represent different sentences (tokens). They are different tokens of the same sentence type. (In the second instance the sentence was being used to "make a statement", but this does not mean that it ceases to be a "sentence" as well.)

When two newspapers set the same headline 'Tax Cuts on the Way --official' they may be presumed to be making the *same* statement ---"Tax Cuts on the Way --official" --- which is probably a somewhat garbled rendering of a statement made by the Chancellor in the House of Commons. On the other hand, if it subsequently turns out that these headlines are referring to statements made by the U.S. Treasury and the British Treasury respectively, they are *not* making the "same" statement, though they *are* using the "same" sentence (type).

In the first case the two headlines printed on the front page of different British newspapers represent different tokens of the same statement (type). In the second case the two headlines represent similar tokens of different statements (types).

Do sentences "have specific meanings"?

A sentence like

'The brown cow jumped over the lazy Moon'

is typically used as an example, to illustrate points of grammar, spelling and typography. No one is actually *asserting* that some *specific* cow *jumped* anywhere, let alone over *the Moon* at a particular time when the Moon could be described, in some metaphorical sense no doubt, as "lazy". Someone wrote down the sentence for the first time, but their intention was to compile a sequence of words, not to make a statement. The question is, whether we should say that such a sentence "has a meaning" or not. We do, by customary usage, say that words "have meanings" when they are mere free-floating words, i.e. not pressed into place in a syntactically correct sequence of words. So perhaps, by analogy, we should say that sentences "have meaning" of this free-floating kind.

We could say that the sentence above "has meaning" but does not have *a* meaning. It certainly does not have "a meaning" in the serious, purposive, rooted sense in which a typical everyday statement "has a meaning". On the other hand, being already formed, a sentence can very easily be asserted by someone. In the blinking of an eye it can become a statement. This, too, would be a justification for saying that it "had meaning" but not "a" meaning.

Some theorists speak of sentences as being "true" or "false". This kind of loose terminology is commonly found in discussions of the paradoxes. But what could it mean to say that a "sentence" is "true" other than that the statement which would be made by a person who asserted it would be true? In which case it would surely be more direct and to-the-point to discuss the issue in terms of statements in the first place.

I suspect that the tendency to speak of sentences as being "true" or "false" crept in as a result of too much concentration on formal languages. But even in formal languages we need the statement-sentence distinction.

We need "sentences" qua sentences in logic-mathematics as when we speak of the set of points which satisfy the sentence '$X + Y = 1$'. (Russell and Whitehead got over this problem by introducing the "I assert that..." sign, |-. When a sentence was preceded by |- it was being asserted (i.e. a statement/proposition was being *made*), but when it was not so preceded it was interpreted as a sentence merely.)

To say that a statement or proposition is "true" is ---as Lucas observed (1969)--- to say that it can be trusted. (It may say more than this, but it certainly says at least this.) Now the notion of 'trust' implies a *source* of that trust. Can one "trust" a sequence of words as such, which no one has suggested we might trust? One cannot categorically say that this is nonsense, because language has immense flexibility, and someone might want to say that they "trusted" a sequence of words, like the cypher for a combination lock. But it makes more sense to distinguish *sentences* which are neither true nor false (but "have meaning" in the weak fashion) from *statements* and *propositions* which are presumed to be either true or false.

References

Atiyah, Sir M. (1977) Trends in Pure Mathematics. *Proceedings of the 3rd ICME*: pp. 61-74. ZDM: Karlsruhe.

Austin, J.L. (1964) Truth in *Truth* (Ed. G. Pitcher) Prentice Hall: New Jersey.

Ayer, Sir. A.J. (1936) *Language, Truth and Logic*. Gollancz: London.

Ayer, Sir. A. J. (1982) *Philosophy in the Twentieth Century*. Weidenfeld & Nicholson:

London.
Bell, E. T. (1937) *Men of Mathematics*, Vol. I Penguin: London.
Black, M. (1935) *The Philosophy of Mathematics*. Routledge & Kegan Paul: London.
Boole, G. (1851/1951) *The Laws of Thought*. Blackwell: Oxford.
Borel, E. (1898/1950) *Lecons sur la Theorie des Fonctions*. (4th Ed.) Gauthier-Villars: Paris.
Bronowski, J. (1973) *The Ascent of Man*. BBC Books: London.
Burge, T. (1979) Semantical Paradox. *Journal of Philosophy*. 76, 169-198.
Butler, R. J. (1962) *Analytical Philosophy*. Blackwell: Oxford.
Carnap, R. (1928) *Der Logische Aubau der Welt*. Berlin.
Carroll, L. (1895) What the Tortoise said to Achilles. *Mind* 278- 280.
Cargile, J. (1979) *Paradoxes: A study in form and predication*. The University Press: Cambridge.
Clark, R. (1981) *Bertrand Russell and his World*. Thames and Hudson: London.
Davis, P and Hersh, R. (1981) *The Mathematical Experience*. Harvester: Brighton.
Dedekind, R. (1956) Irrational Numbers in *The World of Mathematics*. (Ed. J. R. Newman) I, III, 528-536. Simon and Schuster: New York.
Descartes, R. (1637/1970) *The Essential Descartes*. (Eds. E. Anscome and P. T. Geach) Nelson-Open University: London
Ernst, B. (1986) *Adventures with Impossible Figures.* Tarquin Publications: Diss, Norfolk.
Fraenkel, A. (1919) *Einleitung in die Mengenlehre*.Berlin.
Frege, G. (1950) (tr. Austin. J. L.) *The Foundations of Arithmetic*. Blackwell: Oxford.
Frege, G. (1952) (Eds.) Geach, P. and Black, M. *Translations from the Philosophical Writings of Gottlob Frege*. Blackwell: Oxford.
Godel, K. (1930) 'Uber formal unentscheidbare Satze der Principia Mathematica und verwand' in *Monatshefte fur Mathematics und Physic*.
Hersh, R. and Davis, P. (1981, 86) see Davis, P.
Hilbert, D. (1930) *Grundlagen der Geometrie*. (7th Ed.) Berlin.
Hofstadter, D. (1979) *Godel, Escher, Bach: an Eternal Braid*. Harvester: Brighton.
Hofstadter, D. (1985) *Metamagical Themas*.Viking: London.
Kilmister, C. M. (1967) *Language, Logic and Mathematics*. English Universities Press: London
Kline, M. (1980) *Mathematics: the Loss of Certainty*. The University Press: Oxford.
Kripke, S. (1975) Outline of a theory of truth. *Journal of Philosophy*. 72, 690-716.
Kuhn, T. (1960) *The Structure of Scientific Revolutions*. The University Press: Chicago.
Lakatos, I. (1976) *Proofs and Refutations. The Logic of Mathematical Discovery.* The University Press: Cambridge.
Lakatos, I. (1976) *The Methodology of Scientific Research Programmes*. The University Press: Cambridge
Lucas, J. R. (1969) 'True' *Philosophy* XLIV, 169, 175-186.
Martin, R. L. (Ed.) (1984) *Recent Essays on Truth and the Liar Paradox*. The University Press: Oxford.
Monk, R. (1990) *Ludwig Wittgenstein: The Duty of Genius.* Vintage: London.
Monk, R. (1992) 'Wittgenstein and the Bolshevik Threat' in *New Thinking about the Nature of Mathematics*. MAG-EDU UEA: Norwich
Moorcroft, F.(1993) 'Why Russell's Paradox Won't Go Away'*Philosophy*. 68, 263, 99-104.
Nidditch, P. H. (1960) *Elementary Logic of Science and Mathematics*. University Tutorial Press: London.
Ormell, C.P. (1958/1992a) 'On the Paradoxes of Self-Reference' *Mind*. LXVII, 266, 267-

271, reprinted in *New Thinking About the Nature of Mathematics*. MAG-EDU UEA: Norwich.
Ormell, C. P. (1992b) A Modern Cogito: Absolutely Unpredictable Sequences.*Cogito*. 6, 3, 155-163. (1993) A Modern Cogito 2. *Cogito* 7, 1, 50-58.
Parkinson, G. H. R. (Ed.) (1968) *The Theory of Meaning*. The University Press:Oxford.
Peano, G.et. al.(1893/1908) *Formulaire de Mathematiques*. Turin
Peirce, C. S. (1931) *The Collected Papers of Charles Sanders Peirce.* Vol. I The University Press: Harvard.
Peirce, C. S. (1956) The Essence of Mathematics in (Ed. Newman, J. R.) *The World of Mathematics.* Vol. 3, 1773-1783. Simon & Schuster: New York.
Penrose, R. (1989) *The Emperor's New Mind*. The University Press: Oxford.
Pitcher, G. (1964) (Ed.) *Truth*. Prentice Hall: New Jersey.
Poincare, H. (1908/1946) (tr. Halstead, G.B.) *Science and Method*. UPA History of Science: New York.
Popper, K. R. (1954) Self-reference and meaning in ordinary language.*Mind*.LXIII,161-69.
Putnam, H. (1967) 'Mathematics without Foundations'.*The Journal of Philosophy*. 64, 1, 4-22.
Quine, W.V.O. (1952) *Methods of Logic*. Routledge and Kegan Paul: London.
Quine, W.V.O.(1953) New Foundations for Mathematical Logic in *From a Logical Point of View*. Harvard University Press: Cambridge, Mass.
Ramsey, F. P. (1931) *The Foundations of Mathematics*. Routledge & Kegan Paul: London.
Russell, B. and Whitehead, A.N. (1914) see Whitehead.
Russell, B. (1903/1937) *The Principles of Mathematics.* Allen & Unwin: London.
Russell, B.(1940/62) *An Inquiry into Meaning and Truth.* Allen & Unwin/Penguin Books, London.
Russell, B. (1959) *My Philosophical Development.* Allen & Unwin: London.
Russell, B. (1967) *Autobiography*. Allen & Unwin: London
Ryle, G. (1951/54) Heterologicality *Analysis*. 11, 3. reprinted in Margaret Macdonald (Ed.) *Philosophy and Analysis* Blackwell: Oxford, pp. 45-53.
Ryle, G. (1961/68) Use, Usage and Meaning. *The Theory of Meaning*. The University Press: Oxford.
Sainsbury, R. M. (1988) *Paradoxes*. The University Press: Cambridge
Sawyer, W.W. (1992) 'Does Mathematics rest on Facts?' in *New Thinking about the Nature of Mathematics*. MAG-EDU, UEA: Norwich.
Scherer, D. (1974) 'Explanatory Completeness' *Philosopy* . 49, 188, 198-204.
Skinner, R.C. (1959) The Paradox of the Liar. *Mind*. LXVIII 322-335.
Stewart, D. (1818) *Elements of the Philosophy of the Human Mind II.* Edinburgh
Strawson, P. F. (1949) Truth *Analysis*. 9, 6. reprinted in Margaret Macdonald (Ed.) *Philosophy and Analysis.* Blackwell: Oxford, pp. 260-277.
Strawson, P. F. (1952) *Introduction to Logical Theory*. Methuen: London
Tarski, A.(1937) *Logic, Semantics, Metamathematics*. Clarendon Press: New York.
Tarski, A. (1944) The Semantic Conception of Truth. in *Philosophy and Phenomenological Research*, Vol. 4.
Thomson, J.F. (1962) On Some Paradoxes. in *Analytical Philosophy* (Ed. R. J. Butler) Blackwell: Oxford.
Tiles, M. (1989) *The Philosophy of Set Theory*. Blackwell: Oxford.
Tucker, J. (1963) The Formalisation of Set Theory. *Mind*. LXXII. 500-518.
Tucker, J. (1965) Philosophical Argument II, *Proceedings of the Aristotelian Society* (Supplementary Volume), 47-64.
Weyl, H. (1918) *Das Kontinuum*. Leipzig.

Whitehead, A. N. and Russell, B. (1914) *Principia Mathematica*. Vol.I The University Press: Cambridge.

Wittgenstein, L. (1922) *Tractatus Logico-Philosophicus*. Routledge and Kegan Paul: London

Wittgenstein, L. (1953) *Philosophical Investigations.* Blackwell: Oxford.

Wittgenstein, L. (1978) *Remarks on the Foundations of Mathematics*. (3rd Ed.) Blackwell: Oxford.

Zermelo, E. (1908) Untersuchungen uber die Grundlagen der Mengenlehre. *Mathematische Annalen*, 65, 261-281.

Index

Postscript for initially sceptical readers, 1993

The problem of the paradoxes is the problem of *explaining* the paradoxes, in completely self-evident tern language or from formal logic.

The paradoxical arguments always need to be addressed in their *strongest* form (see Note [4] page 16). Only "

The author *hopes* that the arguments which lead to the new superparadoxes have been given in a clear, und necessarily be their strongest form. (After all, we have corporately had the advantage of centuries/decades Grelling's paradoxes, etc. The present author has neither had the same length of time, nor the corporate support, appear to you to be ambiguities in the new arguments, remember that one ambiguity, based on one personal inte is always a defence argument. The author has tried to anticipate some likely counter-arguments, but it is ir interpretation.

The main thesis of this monograph (that the paradoxes are examples of dynamic contradiction) is not tech offered, or even on the collection as a whole: the superparadoxes are simply offered as *evidence* that it (the the phenomena, and that it is not, therefore, merely waffle.