CONFLICT AND CO-ORDINATION IN THE AFTERMATH OF ORACULAR STATEMENTS

BY MARIAM THALOS

I. SURPRISE EXECUTION

The oracle paradox was introduced to philosophical audiences by D.J. O'Connor, as one among several types of pragmatic paradoxes. It is known to some, e.g., P. Weiss and A. Lyon, and R. Sorensen, whose treatment will occupy the next section, as the prediction paradox.\(^1\) However, this choice of terms is liable to cause some confusion with Newcomb's Problem, of game-theory fame, which is also called the predictor paradox. I shall use 'oracle', as it is descriptive, and furthermore is evocative of legends which may deepen appreciation of the category of problem with which we are dealing. One version of the paradox features a judge who issues a proclamation that a certain prisoner, because he is deserving of the most cruel of punishments, shall be hanged at one of the seven noons of next week, but shall be kept in ignorance of the day of execution until the morning of that final day, when the hangman will come to call. The prisoner is present on the occasion of sentencing, and known by the judge to be present.

Later, in his cell, the prisoner reasons as follows that his sentence cannot be carried out:

'The hangman cannot consistently with the sentence come to call on day 7, because on the evening of day 6 I would know with absolute certainty that I shall be hanged on the following day. And this would violate the terms of the sentence. Similarly for day 6, since on the evening of day 5, if still alive, I shall anticipate being executed on the following day, as day 7 has been eliminated as a possibility. Days 1–5 can be eliminated as well, in reverse order. Hence the judge's proclamation is self-refuting: I cannot, given that I am aware of my sentence, be hanged in a state of surprise.'

On day 5, to the prisoner’s surprise, the hangman arrives to announce the appointed day, and hangs the prisoner in fulfilment of the sentence.

It is thus indisputable that the judge’s announcement of the sentence does not make fulfilment of that sentence impossible, since according to our story the sentence is carried out on day 5. And so it follows that the sentence is both self-consistent and consistent with the prisoner’s knowledge of it. Hence we require an account of what goes wrong with the prisoner’s reasoning in support of the conclusion that the sentence cannot be carried out. In fact, the judge’s announcement appears to be instrumental in bringing about its own fulfilment, just as the oracle’s part in Sophocles’ play is instrumental in bringing about Oedipus’ tragic fate, which it foretells. Even so, each course of life – the prisoner’s surprise execution on the one hand, and Oedipus’ fate on the other – might have taken place in the absence of an oracular announcement anticipating it, although each fate appears to have been made more likely by an anticipatory announcement. So, just as we require an account of how the prisoner’s reasoning goes wrong, we require too an explanation of how the oracle of the contemporary puzzle, the judge’s announcement, participates in precipitating the situation it foretells.

II. BLINDSPOTS

The most prominent current solution to the oracle paradox belongs to R. Sorensen. He introduces the conception of an epistemic blindspot. A proposition is an epistemic blindspot, for a specified person, if and only if that proposition is consistent, while the proposition that the specified person knows it is inconsistent. And he calls a proposition a conditional blindspot if it is not a blindspot, but equivalent to a conditional whose consequent is a blindspot. If a certain proposition is a conditional blindspot of mine, then it is impossible for me to know, in combination, the conditional and its antecedent. For example, the proposition If Ralph survived, he is the only one who knows it is a conditional blindspot for all but Ralph. Blindspots are therefore relations in which subjects and propositions, in pairs of one each, may stand, just as knowledge is a relation in which similar pairs may stand.

Sorensen contends that the (single) feature shared by all variants of the oracle paradox is that each involves fallacious reasoning about blindspots.

---

and conditional blindspots. In the surprise execution story, according to Sorensen, the prisoner falls victim to fallacious reasoning when he makes a series of eliminations (it cannot be day 7; it cannot be day 6; ...) from the impossibility of knowing the associated conditional blindspot (if the hangman has not come by the evening of day 6, I shall be executed to my surprise on the following day; if the hangman has not come by the evening of day 5, I shall be executed to my surprise on the following day; ... ) and its antecedent (the hangman has not come by the evening of day 6; the hangman has not come by the evening of day 5; ... ). But these eliminations, Sorensen says, cannot be made because, while it is indeed the case that the prisoner cannot know the conditional blindspot and its antecedent at one and the same time, it is nevertheless possible for both to be true at one and the same time. Thus each elimination is made illegitimately.

In Sorensen’s view, treatment of the oracle paradox requires bringing to attention limitations on the possibilities for knowledge. The very term ‘blindspot’, in fact, calls to mind certain shortcomings of visual perception, suggesting that the proper diagnosis of the oracle paradox will identify limitations on knowledge, which are somehow either unacknowledged or misapprehended by oracle victims. In fact Sorensen believes that the error in the prisoner’s reasoning consists in mixing up assertions about what cannot be known with assertions about what cannot be true. It is indeed correct to assert that knowledge of the statement ‘I shall be unexpectedly executed at one of the 7 noons of next week’ cannot be combined with knowledge of the statement ‘I have not been executed on days 1-6’. The two statements are not co-knowable. But they are not inconsistent. The prisoner, however, eliminates potential noons on the basis of this non-co-knowability, as if non-co-knowability were a species of inconsistency. The proposition that it is impossible to know both a conditional blindspot and its antecedent is without a doubt correct. But, according to Sorensen, such propositions cannot be employed to make the prisoner’s eliminations.

But why not? The prisoner’s reasoning, as it pertains to day 7, for example, might be rephrased, using Sorensen’s own terms of art, as follows: ‘If alive on the evening of day 6, I shall know at the same time both a conditional blindspot — namely, “If the hangman has not come to announce the day of execution by the evening of day 6, I shall be unexpectedly executed on the following day”, which follows from the judge’s proclamation — and its antecedent. Since, as Sorensen teaches, these two propositions are not co-knowable for me, I shall not be alive on the evening of day 6. Hence day 7 can be eliminated.’ But where is the error in this reasoning? The prisoner agrees with Sorensen’s contention that blindspots exist, and even with the identification of his own blindspots. Since a conditional
blindsport and its antecedent are not co-knowable, the scenario imagined by the prisoner – namely, that he remains unexecuted on the evening of day 6 – must be ruled out, since it would give rise to knowledge of what cannot be known.

Sorensen might reply that since it is impossible to know both a conditional blindsport and its antecedent, the prisoner should not accept the proposition that if alive on the evening of day 6, he will know at the same time both a conditional blindsport – namely, ‘If the hangman has not come to announce the day of execution by the evening of day 6, I shall be unexpectedly executed on the following day’ – and its antecedent. But again, why not? Acceptance of conditionals whose consequents are impossibilities is not impermissible. In fact, such conditionals are the basis for *reductio ad absurdum* arguments, of which class of arguments the prisoner’s aspires to be a specimen.

So, while the hypothesis according to which blindsports exist is surely correct, it is also entirely unilluminating, since the prisoner makes no obvious errors in applying it, and since the prisoner’s reasoning is only all too commendable by its lights. What is more, it says nothing concerning whether the prisoner, on learning of the philosophical labours undertaken on his behalf by Sorensen, could extract himself from his current difficulties.

Sorensen’s thesis that knowledge admits of blindsports, while perfectly correct, does not concern itself with treatment of the salient questions: (a) why does the prisoner undertake to make eliminations? Why, in other words, does he undertake to formulate an anticipation of which day he will be executed on? And (b): is the judge in any better position to anticipate the day of execution, should she later find herself in the embarrassing position of having forgotten what she inscribed on the execution order? And these, as I shall argue, are questions which must be addressed by any theory that purports to explain the announcement’s contributions to the prisoner’s difficulties. No doubt the prisoner utilizes Sorensen’s principle concerning blindsports, according to which it is impossible to know both a conditional blindsport and its antecedent at the same time. But is this principle, in combination with knowledge of the proposition expressed in the judge’s sentence, the only grounds for the prisoner’s eliminations of eligible days? No. For, as I shall argue, the prisoner draws also on the proposition that both prisoner and judge know of the prisoner’s learning of the sentence, via being present at sentencing. And I shall further argue that knowledge of (a) the blindsport principle, and of (b) the fact that this principle applies only to what one can know; not to what can be true, does not eliminate the difficulties of reasoning in the prisoner’s situation. In fact, nothing can. And this is a very important accomplishment on the judge’s part.
III. CO-ORDINATIONS AND ANTI-CO-ORDINATIONS

Traditional formulations of the oracle paradox involve a single, rigid order of eliminations among options whose number is known to the victim of the paradox. There are versions, however, which do not involve a rigid order of eliminations, and (yet other) versions in which the number of options is not known to the paradox victims. Thus neither the rigid order of eliminations nor the knowledge that a certain number of options exist is essential to the puzzle.

Here, then, is a two-day version of the surprise execution. Suppose the judge, immediately after passing sentence, makes an elaborate show in the prisoner’s presence of inscribing on official documents an execution order, ceremoniously announcing, as she does so, that she is now inscribing on it the precise day of this weekend (Saturday or Sunday) on which the hangman is to call on the prisoner. On the basis of attendance at this inscription ceremony, the prisoner comes to have further corroboration for his belief, first formed as the judge passed sentence, that there is a day this weekend on which he is to be executed. If, on the other hand, the prisoner had been kept entirely ignorant of the proclamation that he is to be executed one day this weekend, and furthermore to be executed unexpectedly, he would not have had a basis for formulating beliefs concerning which day he will be executed; he might not even have had an incentive to ruminant on the subject. So he would have been in that instance very susceptible to surprise execution. And so the judge could have achieved her aim to have the prisoner executed unexpectedly by keeping him in the dark about the sentence. But she does not, and instead declares the sentence openly, in the prisoner’s very presence, and achieves the desired effect thereby. And this is what cries out for explanation.

What can go wrong en route to opinion in the circumstances in which the prisoner finds himself? The reasoning usually attributed to victims of the oracle paradox involves a series of eliminations, based on present beliefs concerning potential future beliefs, which are formulated against a backdrop of knowledge of a certain anticipatory announcement and the context in which it is produced. My contention will be that such beliefs, if they are to be well founded, must be made in the light of the proposition that the announcement-maker deliberates rationally prior to making the announcement. I shall call ‘projections’ those beliefs which are formulated, at least in

---

part, on the basis of anticipations of other individuals’ reasoning processes, and on the assumption that those other individuals are rational beings too. The action of making a projection will be called ‘projecting’. Using this notion I shall argue that the projection undertaken by the prisoner cannot be concluded. The prisoner’s mistake, therefore, is not that he makes eliminations, but that he supposes that this process of eliminations, once entered into, can be halted, and thus that it can legitimately be concluded that the sentence is inconsistent with his knowledge of it.

The prisoner must be prepared, when reviewing his prospects, to take into account also the matter from the judge’s point of view, since it is the judge who is responsible for the means and circumstances of the prisoner’s execution, the judge who has performed the inscription ceremony, and the judge who can be assumed to have performed rational deliberations in the service of all these things. And the prisoner must, if his own deliberations are to be well founded, come inevitably to realize that the judge’s part must be supported by reasons that take also into account his (that is, the prisoner’s) reasonings. For the judge has had to decide upon a day this weekend, in such a way that she can project that when the hangman calls on the named day, the prisoner will not be expecting the hangman. And she has had to do this with awareness of her own intention to proclaim the sentence, and with awareness too that the prisoner will be – precisely as she intends him to be – present at the proclamation. And she knows that the prisoner also knows this.

The prisoner looks at things from the judge’s point of view. Will she write ‘Sunday’? How can she, knowing that on Saturday evening the prisoner will reason that, since he is then still alive, the fated day is Sunday, and so prevent fulfilment of the sentence? She apparently has no choice but to write ‘Saturday’. But knowing this, and knowing that this will occur to the prisoner too, her attention is drawn away from Saturday, and to Sunday once more. Only to be turned away again. The process of deliberations which the prisoner must suppose that the judge has undertaken is one that requires projections: in fact, it requires reciprocal projections. Reciprocal projections are normally self-reflective: the reflections themselves are among those matters on which projections are made. And this is where things are liable to go wrong, when conditions are unfavourable.

The conditions in which judge and prisoner find themselves, in the aftermath of the judge’s oracular statement, are unfavourable. For they are such that the normal procedure for consummating reciprocal projections will not halt. The swinging phenomenon we have witnessed between Saturday and Sunday, and which is present on both sides of the announcement (judge’s and prisoner’s), is doomed to continue without termination.
This fact comes out most transparently when the situation is looked at from
the judge's angle. For the halting of deliberations is a precondition for
surprise execution (since the judge must come to a decision first, and only
subsequently order execution), yet the proposition that deliberations halt
entails that no unexpected execution can be carried out, since halting
indicates that the day of execution can be anticipated. And yet the judge
must – and *has*, according to the story – resolved this small difficulty, so as
subsequently to perform the inscription ceremony. She has achieved what
appears to be an impossible task, simply by making a statement whose
announcement presupposes the accomplishment of such a task.

The flowchart depicted in Figure 1 is one way of formalizing this process.
The decision depicted within the dotted portion must be made at every
cycle, and at every cycle the answer must be no; for one cannot be making a
decision if one is halted. This diagrammatic representation of the procedure
undertaken by the prisoner, and attributed by him to the judge, illuminates why
a surprise execution can be held even on the last day, as many friends of this puzzle
agree, although (as G.C. Nerlich is at pains to point out<sup>4</sup>) this is not to deny that
the last day is a queer case. For this procedure will not terminate even if \( n = 1 \).

The prisoner, now in his cell, considers these matters. The unhalting swinging
phenomenon is present in his own expectations, since he believes it must exist in the judge's as well. So, even if he is careful not to conclude that he cannot be executed in compliance with his sentence, he cannot on Friday project that he will be executed the following day, because his projections concerning the judge's 'move' will not converge; each new consideration swings him to the opposite point. And it is precisely because he is reflecting carefully on the matter that he knows that careful reflection on reciprocal projections cannot bring the swinging phenomenon to a halt. True, he also knows that the judge has in fact broken the symmetry between Saturday and Sunday, but he himself cannot do so by making and reflecting on projections.

The prisoner can know (by going through the sort of reasoning we have here reviewed) that the procedure he has undertaken does not terminate. Does this help? It might lead him to believe that the judge must have used some procedure different from the one he is currently using to anticipate her decision – possibly some randomizing mechanism such as a coin toss. So the prisoner prepares to resign himself to his fate. On further reflection, however, the prisoner realizes that if the judge has used a randomizing device, then surely Sunday must be eliminated as a potential day to be selected by lottery, since (as the judge can anticipate the prisoner will come to realize that randomization has been used) on Saturday evening the prisoner will be in a position to anticipate that Saturday was not selected by lottery. And if not Sunday, then Saturday; and the non-halting phenomenon re-emerges. Hence it does not help the prisoner to realize that the original procedure does not terminate. For any alternative procedure, even a randomizing one, will also suffer from exactly the same deficiencies. The announcement makes any projections on the prisoner’s part, as to the potential date of execution, unstable.

Now one might suppose that the symmetry between Saturday and Sunday exists only from the prisoner’s perspective, and can easily be broken by the judge, since, in fact, making the announcement appears to presuppose that the symmetry can be broken. This is not so. To consider the matter again from the judge’s point of view, suppose she decides against Sunday on grounds that the prisoner, because present at the inscription ceremony, will always have a basis for projecting that on Saturday evening, if still alive, he will anticipate Sunday execution. And this becomes a reason for finding against Saturday too. Since there are now considerations against both days, the original asymmetry between Saturday and Sunday is broken. When she is favourably disposed to ‘Saturday’, confident that the prisoner will not on Friday evening be possessed of sufficient grounds for formulating a projection concerning a Saturday execution, this favourable disposition itself becomes a reason against Saturday. So even for the judge symmetry is no sooner shattered than it is restored. Considerations of reciprocal projections, while they cannot rationally be ignored, are at the same time (at least in this case) to the disservice of the aims for which these very projections are sought. And this is true on both sides of the announcement. It is true of the judge, just as it is of the prisoner, that she cannot on the basis of projections alone formulate an anticipation concerning the day of execution, even though she will make the announcement, or even has made it but forgotten the date established for it. And this also is a datum that any theory treating this problem must somehow accommodate.
IV. CONFLICT

The gods are just, and of our pleasant vices
Make instruments to plague us (King Lear v iii)

The lessons I shall draw rest on this simple and uncontroverted proposition: that many beliefs — many more than one would think offhand — rest on anticipations of other individuals’ reasoning processes. And that some of these beliefs are reciprocally dependent, because they involve anticipations of each projecting individual by the other. I shall refer to beliefs which exhibit this type of mutual dependence as common-perspectival. Common-perspectival beliefs are the result of tacit, and typically also subconscious, intellectual negotiation — to make use of already well established terminology of strategy theory. Formation of common-perspectival beliefs cannot occur in a sterile, intellectually sealed environment in which each expectant self is blind to projections others may make by anticipating its reasoning processes, or in an environment in which anticipations of one individual’s reasoning processes by another can have no impact on the projections of the individual whose reasoning processes are being anticipated. Rather, formation of common-perspectival beliefs occurs only among individuals of whose intellectual lives certain aspects are known to be, and known to be known to be, open to public inspection — individuals whose intellectual lives, as well as their public lives, are, just as they aspire to be, intertwined.3 Aspirations that others should be vividly aware of certain of one’s anticipations need not be altruistic; for public access to certain aspects of one’s intellectual life is under normal circumstances in the service of self. And in fact one might even say, and not without precedents, that rationality is a species of anticipatability.

Why is the prisoner (in particular) unable to bring deliberations to a stop? My explanation will be simply that the deliberations he aspires to consummate require the possibility of negotiating common-perspectival beliefs; that negotiation requires compatible aims; but that the aims of prisoner and judge are incompatible.

Projections are typically formulated against a background of aims or purposes, some more remote than others. Otherwise there would be no reason to favour projecting concerning one potential episode in human affairs rather than another. I shall say that two aims come into accidental conflict when both can be satisfied together in some logically possible set of

3 This idea is brilliantly treated by Thomas Schelling, The Strategy of Conflict (Harvard UP, 1960).
circumstances, but cannot both be satisfied in the actual circumstances. (It is logically possible, for example, that each of our aims to win a million dollars in a lottery be satisfied, but not that each aim be satisfied under circumstances in which we enter the same lottery with only one million-dollar prize.) And I shall say that two aims are in direct conflict when, logically speaking, both cannot be satisfied together under any circumstances.

The prisoner of the surprise execution story acquires the aim of anticipating the day of his execution, upon coming to learn his sentence. The judge, in passing sentence, has the aim that the prisoner shall not anticipate the day of execution, knowing that the announcement will nevertheless foster in the prisoner the aim to anticipate the day. It is in fact with the aim that the prisoner be unable to anticipate the day that she passes sentence as she does, in his presence. It is not possible that both aims be satisfied together. Thus they are in direct conflict. The judge herself creates this conflict by (a) establishing her own aim that the prisoner shall be taken by surprise, and (b) announcing the sentence in the way she does, thereby fostering in the prisoner the aim of anticipating the day of execution. Finally, the sentence itself, and the prisoner’s presence at its proclamation in the judge’s direct sight, create conditions which favour achievement of the judge’s aim, but not the prisoner’s. And the fact that the prisoner can come to acknowledge all that we have said here does not improve his situation in the minutest degree.

Thus we may explain the prisoner’s difficulties in terms of conflicting aims, which cause the fabric of common-perspectival beliefs to unravel. These snags in the fabric of common-perspectival beliefs often favour the satisfaction of one aim over another. In the surprise execution, the judge’s aim is favoured, since her aim is not that anyone concerned (even herself) should have grounds for anticipating the day of execution, but rather that the prisoner in particular should not be possessed of such a ground. And she cannot undermine the prisoner’s potential foundations for such a belief without undermining her own as well.

However, it is not the very fact that the beliefs in question are common-perspectival which is the problem, but instead that the aims on which they must be negotiated are directly in conflict. The execution case may be contrasted with the case of two friends, one of whom says to the other ‘I shall pay you a visit this weekend on a day such that you will be able to anticipate my visit the day before’. The promiser makes the announcement with the aim that the promisee should be able to anticipate the visit. The promisee also has this aim, or could have it. There is no conflict, since both know that if the promiser stays away on Saturday, both will formulate a
projection, and anticipate reciprocal projections, of a Sunday visit. There is no instability, because there is no direct conflict.

The problem with the surprise execution is not, therefore, the presence of an element of self-reference in the announcement. The contrasting case of the two friends also shows this to be the case, since even if we concede there is self-reference there, that does not lead to impossibilities of projection. The problem springs from the existence of directly conflicting aims, whose presence prevents any successful negotiation of expectations which, if negotiation were possible, could lead to stable beliefs. The prisoner can prevent the tortured reasoning we have witnessed only by either withdrawing from, or else refusing to enter into, the enterprise of projecting. And he does either at the cost of having no opinion whatever about the day of execution. And the same goes for the judge. This is the supreme achievement of the judge’s announcement. We can explain, therefore, how the anticipatory announcement precipitates the situation it foretells, where in its absence there might have been no difficulties, nor incentives to project.

According to the present view, therefore, the prisoner of the original story goes wrong in only one way: he concludes on the basis of otherwise impeccable reasoning that he cannot be executed in satisfaction of the sentence. His only mistake is to suppose, understandably, that once seven days have been eliminated, none remains. What we have seen is that there cannot be an end to eliminations, although there can be a beginning and a continuing. To put the point metaphorically, the prisoner’s cup is bottomless with days to be eliminated, although these are always seven in number. His cup of days is filled and refilled with the same days; but, as in the typical American restaurant, it is never empty. This bottomlessness of the prisoner’s cup is an accomplishment of the judge’s announcement. The judge achieves her aim that the prisoner be executed in a state of surprise, and prevents the prisoner achieving his own aim to anticipate the day of execution, which is itself brought into existence by the proclamation. There is nothing wrong with the prisoner’s reasoning as it pertains to eliminations. But he is conspired against by the judge, who makes certain arrangements to ensure that these eliminations cannot come to an end.

V. AN OBJECTION

I have purported to explain the prisoner’s difficulties through bringing to attention a pair of conflicting aims — in other words, by declaring that the

---


© The Editors of *The Philosopher* (January), 1997
difficulties which he is experiencing spring from the ordinary conflicts of
day-to-day life. But, the objection might go, is it not enough to cause the
very same trouble that the prisoner should simply believe the anticipatory
proclamation correct, for whatever reason? Is not the trouble caused simply
by the logical nature of the judge’s statement? So have I not unnecessarily
added to the story that the judge is trying to bring the proclamation about
by means of the proclamation itself?

No, for two reasons. First, while I have indeed brought the focus of
attention to the judge’s aims, I have not thereby added to the story. For the
judge’s aims have been part of the story all along, since, as I say, projections
such as those the prisoner aspires to achieve are normally founded on
anticipations of other individuals’ reasoning processes; but up until now the
judge’s aims have been only an implicit part of the story. It is in fact
precisely the judge’s aims, implicitly understood, which make the prisoner’s
difficulties of reasoning so robust. For if we are allowed to assume that the
judge has no aims to bring the proclamation about through its very
announcement, then we can tell a surprise execution story in such a way
that there are no difficulties whatever of reasoning.

Suppose the prisoner learns accidentally of the announcement (that he is
to be executed one day next week, but is not to be in a position to anticipate
the day in advance), for example, by overhearing a conversation between
jailer and executioner; but he does not learn also from this source the day
for which the execution is arranged. In this scenario, the prisoner, even if he
has no basis for anticipating the day of execution, nevertheless has no
grounds for eliminating any eligible day, not even the last. But he can, once
again, be certain that, if alive on the evening of day 6, he will know that he
is to be executed the following day. And this is precisely the sense in which
the prisoner’s coming to learn of his sentence through being present at
sentencing gives him no information from which to work out the day, and
(on the contrary) leads him to aim at an unattainable goal.

Second, the objection that I have changed the problem presupposes,
correctly, that the difficulties of reasoning faced by oracle victims result from
activities of deduction. However, the activity of deduction, according to this
objection, is a strictly syntactic affair, governed exclusively by rules of
inference; deduction is nothing but a mechanical procedure of applying per-
missible rules of inference. This is false. For one thing, deduction involves a
screening of premises as to consistency and possible truth, which mere

7 For example Nerlich writes (p. 513) of the oracle’s announcement-maker: ‘That [i.e., the
prisoner’s difficulties of reasoning] is precisely what [she] wants to achieve’; and cf. Martin
Gardner’s popularization of the puzzle in The Unexpected Hanging and Other Mathematical Diver-

© The Editors of The Philosophical Quarterly, 1997
application of inference rules does not. In fact, an assessment of the consistency of a certain sentence with realities of which he is painfully aware is precisely what the prisoner of our story is at pains to conduct. More importantly, deduction is goal-orientated, aimed at deducing a particular conclusion, not merely permissible ones. So why is our prisoner unable to come to make a satisfactory conclusion in his efforts to assess the consistency of the sentence with his knowledge of it? Sorensen is on to something when he suggests that it is the prisoner’s epistemic position vis à vis his sentence which causes the trouble. But Sorensen does not grasp sufficiently clearly the strategic element in the species of belief which the prisoner is at pains to achieve.

VI. WAYS OF PARADOX RECONSIDERED

A paradox, according to the view that deduction is mere application of inference rules to propositions, is a piece of argumentation consisting of a chain of apparently impeccable reasoning, from premises thought to be true, or definitions thought to be unproblematic, to a proposition that is either inconsistent or surprising. Quine distinguished three species of such argumentations.8 The first, which he called ‘veridical paradoxes’, reveal a surprising truth; the second, which he called ‘falsidical paradoxes’, expose a falsehood, previously unrecognized, in our presuppositions; and the third, which he called ‘antinomies’, call for far-reaching revision in extensive networks of conceptions or presuppositions. However, on this taxonomy, certain of the so-called ‘pragmatic paradoxes’ (for example, one of the most famous specimens, Moore’s statement ‘It is raining, but I do not know it’) do not count as paradoxes, since what is said (the proposition expressed) entails nothing that is either surprising or contradictory. But surely these puzzles too require philosophical treatment, which, it is to be hoped, will shed light on certain philosophical matters, or test philosophical proposals, or provoke philosophical enquiry. Typically, a pragmatic paradox involves a single source of information that seems to impeach itself in some way, leaving us not knowing what to make of things. We would like to have a philosophical account of such impeachments, and of the means by which they can be achieved.

In the oracle paradox, apparent self-impeachment is achieved through a statement whose announcement is evidence that the impossible has been achieved. Now a statement of the form ‘Expression of this statement is an impossible achievement’ is not a contradiction. Such a statement might very

well be true, even if what is expressed is something that cannot, if expressed, be true. The interesting thing about the oracle is that the statement announced to victims does not, unlike the statement ‘Expression of this statement is an impossible achievement’, entail a statement whose very expression entails its falsehood. So these statements are not inconsistent with their announcement.

But self-impeachment is not a species of inconsistency. Things are just the other way round: inconsistency is a species of self-impeachment. Thus the more celebrated paradoxes are species of a larger class of problem involving the pragmatics of life, of which the less celebrated specimen of paradox is the more representative. This suggests that we cannot amend Quine’s taxonomy simply by adding more categories. We must instead insert Quine’s limited taxonomy into a more encompassing one. This idea is quite suggestive; for might it not be true that some of the celebrated semantic paradoxes admit of solution in much the same way as pragmatic paradoxes do? If so, there might be fewer antinomies than previously thought.

As we have seen, the failure of the prisoner’s evaluation process to come to a halt is not due to failure of consistency: the swinging phenomenon (as close as we get to inconsistency) is not the cause, but instead a symptom or manifestation of the problem. The problem, as we have seen, is that certain expectations, which typically form the basis of action, cannot be negotiated. This is a result of actions taken by the judge – actions that must be seen as presupposing certain reasoning processes, not yet undertaken by the prisoner, but nevertheless anticipated. There is a grain of truth in the prisoner’s ultimate conclusion, for indeed there is an impossibility in his situation, which previous friends of this puzzle have failed to diagnose. It is the impossibility of negotiating reciprocal expectations under the circumstances created by the judge. But this impossibility is not a species of inconsistency. Impossibilities of this sort are abundant. Where there is direct conflict, it will always be impossible for all involved to achieve their aims.

The presumption that every paradox involves a problem in the domain of concepts, propositions and their entailment relations involves a presupposition that all paradoxes are treatable. The proper treatment for a pathological conception or semantic principle, like the unrestricted comprehension principle with which Russell’s paradox finds fault, is to dismiss it and enquire after a more virulent or better refined species of conception or principle. And the proper treatment for misapplication of principles, concepts or predicates is to disallow the applications. But there may be no such cure for a situation which makes certain aims impossible to achieve. For while we may be able to immunize against infection by tainted semantic

© The Editors of The Philosophical Quarterly, 1997
devices, principles or applications through banishing them from civilized society, what can be done about a constellation of conditions such as those we have been discussing, which, when they come together, do not tolerate the existence of (for example) a rational anticipation concerning one's fate? We who encounter oracular statements under controlled and sterile laboratory conditions can only mutter 'There, but for the grace of God, go I'.

State University of New York at Buffalo