

Conflicting Rules and Paradox

This paper seeks to understand various paradoxes as cases of conflicting rules. In particular, the ambition is to outline a new perspective on and response to the Liar – though it will take us a while to get that far. We begin in section one with an account of simple rule confliction. Section two then brings this account to bear on a paradox, the Secretary Liberation Paradox, which is readily seen to involve conflicting rules. Finally in section three I suggest that the Liar can also be seen as a case of rule confliction, and outline briefly how that perspective provides for Liar arguments to be blocked.

Section 1

1.1

Let's consider a couple of straightforward examples of the general phenomenon of conflicting rules.

Two children arrive in a field to play a game of cricket. Before play begins the children agree that the batsman wins if he hits a six and that the batsman loses if he hits the ball into nettles (to hit a six in cricket is to hit the ball over the edge of the pitch without its bouncing). The game then proceeds merrily until the batsman hits a big six into a previously unnoticed clump of nettles. The obvious argument ensues: the batsman claims he has won because he has hit a six, and the bowler claims that the batsman has lost, and so not won, because the ball has been hit into nettles. Now we may have here, I take it, a genuine conflict of rules. There may be no case for saying that when the children agreed both that batsman wins if he hits a six and that the batsman loses if he hits into nettles two rules were not successfully laid down. Had the batsman hit a six not into nettles he would have won, and he would have lost if he had hit into nettles within the boundary. And nor may there be any case for saying that there are two rules here, but these rules are not both quite as the children agreed them. There may be no basis either for saying that the one rule was really that the batsman wins if he hits a six unless that six goes into nettles, or for saying that the other rule was really that the batsman loses if he hits the ball into nettles unless the hit is a six. Rather, to repeat, we may have here two rules in genuine conflict.ⁱ

A second example is due to Graham Priest. Priest writes:

Suppose that there is a certain country which has a constitutional parliamentary system of government. And suppose that its constitution contains the following clauses:

In a parliamentary election:

- (1) no person of the female sex shall have the right to vote;
- (2) all property holders shall have the right to vote.

We may also suppose that it is part of common law that women may not legally possess property. As enlightenment creeps over the country, this part of common law is revised to allow women to hold property. We may suppose that a *de facto* right is eventually recognised as a *de jure* one. Inevitably, sooner or later, a woman, whom we will call 'Jan', turns up at a polling booth for a parliamentary election claiming the right to vote on the ground that she is a property holder. (Priest 2006 pp.184-5)

By one law, Jan has the right to vote, by another she does not. Both laws are in force and neither, we may suppose, has an implicit restriction or 'exceptive clause' making the conflict merely apparent:

A case may easily arise where both of the contradiction-producing laws are of equal rank, where both laws were produced at the same time, where the intentions of the legislators are lost in the mists of time, are moot, or are downright inconsistent, where there is no precedent for waiving one law rather than another, and so on. In short, there is nothing objective to underpin the claim that one law has an implicit exceptive clause in virtue of which it is not applicable. (Priest 2006 p.187)

The possibility of rules conflicting in the way exemplified by these two cases is, I take it, uncontroversial. From this possibility, however, Priest draws a highly controversial conclusion. By ‘contradiction-producing’ he means exactly that. Jan both has and does not have the right to vote. The batsman both wins and loses – and so wins and does not win (a loser is not a winner). Priest’s reasoning to this conclusion is rather short. He argues that if it is a law of the land that no woman has the right to vote, then as Jan is a woman she does not have the right to vote. And if it is a law that property holders have the right to vote then Jan the property holder has that right. But this reasoning can and should – I think – be blocked. There is an inference being made here that Priest does not make explicit. Jan’s having the right to vote follows from her being a property holder and all property holders having the right to vote. It does not follow from her being a property holder and its being the law that all property holders have the right to vote. In order to obtain a contradiction from the conflict of property rules, we need therefore to move from its being a rule that property holders have the right to vote to its being the case that property holders have the right to vote. And this move can be resisted. Indeed, this move *has* to be resisted if we don’t want to follow Priest to the contradiction.^{ii, iii}

The point here is of course quite general. If it is agreed that there can be rules in force which give incompatible pronouncements on a certain case, then as the pronouncement of a rule is a consequence of the rule’s statement together with relevant facts of the case, it follows on pain of contradiction that a rule can be in force without its statement being true. It is a rule both that the batsman wins if he hits a six and that he loses if he hits into nettles, but it is not the case both that the batsman wins if he hits a six and that he loses if he hits into nettles. It is a law both that property holders have the right to vote and that women do not, but it is not the case both that property holders have the right to vote and that women do not. That there can in this way be a gap between a rule being in force and its statement being true is, I take it, an interesting first lesson of the possibility of rule confliction.

1.2

I shall expand on this lesson below. Before doing that, though, I want to make a clarificatory comment on what I has just been taken to be uncontroversial, and then also to make two constitution claims. The first constitution claim concerns a rule’s having a certain verdict – that is, its being the case that by rule R, p – and the second concerns such things as a person’s having the right to vote, or a batsman’s winning. Both claims will be of central importance through the paper.

It is uncontroversial, I claimed, that there can be such a thing as a conflict of rules. Rules can come into (genuine) conflict. In saying this, we should however be clear, no commitment is made to the controversial position that *any* set of rules may be in force no matter how they might conflict – to the position that, as it were, *any* conflict of rules is possible. One might well think that it could not be a rule of a cricket game both that the batsman wins if he hits a six and that he loses if he hits a six: a game is something playable and there is no such thing as playing by both of these rules. (One could equally not play by the rules that the batsman scores one if he hits the ball, must change ends if he scores one and does not score if he changes ends.) It is plausible, this is to say, that the potential for conflict had by certain rules tells against the possibility of their being in force together. But this is perfectly consistent with the equally plausible claim that certain other rules which contain a potential for conflict *can* be in force together and *can* together be played by. Which is just to say, again, that there can be such a thing as a conflict of rules.

Consider next the idea of a rule's statement. By the statement of the rule that p I mean, simply the statement that p . The statement of the rule that the batsman wins if he hits a six is the statement that the batsman wins if he hits a six. Or slightly more carefully, it is the statement that a circumstance in which a batsman playing the game hits a six is a circumstance in which that batsman wins the game. I claimed above that the verdict of a rule is a consequence of the rule's statement together with relevant facts of the case. This should be clear enough. Jack the batsman hits a six, and the rule's verdict in this case – the verdict that Jack wins – is a consequence, given that Jack has hit a six, of the statement that the batsman wins if he hits a six. It should also be clear, I think, that the converse claim, suitably restricted, is true. If it follows from a certain cricket rule's statement together with the circumstances of a certain play that the batsman wins, then it is the verdict of that rule that the batsman wins. Indeed, the obvious but important thought here is that the rule statement's having the consequence *constitutes* the rule's having the verdict. If in a certain circumstance it is a consequence of a rule statement that the batsman wins, then this fact constitutes the rule's having the verdict that the batsman wins.

The second key constitution claim draws on the notion of a set or body of rules having a verdict. This notion will be centrally important through the paper, for I shall hold that what determines whether the batsman wins, or whether someone has the right to vote, are not individual rules as such but rather whole bodies of rules. If we want to know whether someone has the right to vote, we must consider the whole body of relevant law: an answer cannot be had to that question by considering a single law only. To see how a body of rules may provide an answer to such a question, however, we need to define a notion of such a body's having a verdict. This may be done – for the purposes and examples of this paper – as follows: it is necessary and sufficient for a set of rules to have a certain verdict that one or more members of the set has that verdict, and no member of the set has an incompatible verdict. This notion then goes to work through the claim that if a relevant set of rules has the verdict that p , the set's having this verdict constitutes its being the case that p . If in a certain circumstance it is the verdict of the body of rules of the cricket game that the batsman wins, then the batsman does indeed win, his winning being constituted by the fact that it is the verdict of the body of rules that he wins. Winning is, in this way, a matter of the verdict of the rules of the game.

1.3

Let's return now to our result that a rule can be in force even though its statement isn't true. This prompts various questions. The line of inquiry I want first to consider notes that whilst cases of rule confliction may show that rule statements are not invariably true, they are certainly *standardly* true. If it's a rule that p , then standardly, at least, it will be the case that p . So how, if there can be a gap between a rule being in force and its statement being true, is this standard truth to be explained? And how does its explanation allow for the statements of conflicting rules not both to be true?

Our second constitution claim provides for a relatively straightforward response to these questions. Suppose that it is a rule of the cricket game that the batsman loses if he is caught, and suppose further that the batsman is caught. Given that the batsman is caught, it is the verdict of the rule that the batsman loses if he is caught that the batsman loses. Standardly, this will mean that it is the verdict of the body of rules of the game more generally that the batsman loses. And if it is the verdict of the body of rules of the game that the batsman loses, then the batsman does indeed lose, this being constituted by its being the verdict of that body that he loses. The batsman loses if he is caught because if he is caught it is the verdict of the game's set of rules that he loses. A similar story may be run regarding the right to vote. Suppose it is a law of the land that all men over the age of 50 have the right to vote. If a person is both male and over 50, this law rules that they have the right to vote. And if from here we may conclude that if a person is male and over 50, the body of law of the land rules that they have the right to vote, then it follows that a man over the age of 50 does indeed have the right to vote, his having that right being constituted by that's being the ruling of the land's body of law.

This, I suggest, is how rule statements of the sort we have had in view are standardly true. And it should be clear how this account of the standard truth of rule statements allows for statements of conflicting rules not to be true. In making it a rule that the batsman loses if he is caught, the children make it the case – in the circumstances – that if the batsman is caught it is the verdict of the set of rules of the game that the batsman loses, and so they make it the case that if the batsman is caught he loses. This phenomenon might be called ‘truth by decree’ – to decree that p is to make it a rule that p . It is, however, not true by decree that the batsman wins if he hits a six, for in laying down this rule – in issuing this decree – the children do not in the circumstances make it the case that if the batsman hits a six it is the verdict of the set of rules of the game that he wins. If the batsman hits a six but that six goes into nettles the conflicting verdicts from the different rules mean that there is no verdict from the game’s body of rules on the batsman’s winning.

1.4

We have then an explanation of how, if it is a rule that p , it will standardly be the case that p – an explanation I have labelled ‘truth by decree’. Importantly, this explanation allows that statements of conflicting rules are not both true. We do not have here, however, a full account of the phenomenon of rule confliction. For whilst what has been said allows that the statements of conflicting rules are not both true, it does not imply a positive position on the truth or untruth of such statements. Do property holders have the right to vote?

As a preliminary to answering this question, let’s sharpen up the ‘standardly’ in the claim that rule statements are standardly true. For the kind of rules we have had in view, a rule’s statement will fail to be true by decree if there is another rule with which it is in potential conflict. Equally though, such a rule’s statement will fail to be true by decree *only* if there is another rule with which it is in potential conflict. If it is a rule that the batsman loses if he is caught, then if the batsman is caught it will be the verdict of the set of rules of the game that the batsman loses unless there is another member of that set which may in such a circumstance deliver a conflicting verdict. And this, we can note, provides a constraint on theories of the truth or untruth of statements of conflicting rules, for what we have in view here is that if a rule statement is not true by decree, this is only *because* of a possible conflict. The failure of the rule statement that the batsman wins if he hits a six to be true by decree is explained by nothing other than the existence of another rule with which the six rule is in potential conflict: this conflict is *the only reason* why it is not true by decree that the batsman wins if he hits a six. Again: had no such other rule been in force, then by this fact alone the statement that the batsman wins if he hits a six would have been true by decree.

Now at least one of the two conflicting cricket rule statements doesn’t merely fail to be true by decree: at least one of them must on pain of contradiction fail to be true *simpliciter*. But which, and how? Well, our answer to the ‘how’ half of this question is constrained. If it is not true that the batsman wins if he hits a six (say), there will be no explanation of this untruth independent of the possible conflict of rules, for an independent explanation of why it is not true that the batsman wins if he hits a six will at once be an independent explanation of why it is not true by decree that the batsman wins if he hits a six, but there is, we have just claimed, no such explanation. It is not true that the batsman both wins if he hits a six and loses if he hits into nettles, but there is no explanation of the untruth of this conjunction independent of possibility of conflict between the rules it states.

This point will resurface in later discussion. Rather than dwell on it now, however, let’s look to give a suitably constrained account of the truth and untruth of the statements of conflicting rules. There are, on the face of it, three possible lines available here. The first begins with the thought that if the batsman hits a six into nettles there are conflicting verdicts from different rules of the game as to whether he wins, and so no verdict from game’s rules considered as a set. But as there is nothing more to winning or failing to win than the verdict of the game’s rules, there is ‘no fact of the matter’ as to whether the batsman wins. Similarly for Jan: the rule confliction means that the body of law of her land does not give a verdict on her case, and so there is no fact of the matter as to whether or not she has the right to vote.

The second line is Priest's: by one rule Jan has the right to vote and by the other she does not, so she both has and does not have the right to vote. And a third line, rejecting both dialetheias and absences of facts of the matter, suggests that there is more than one way in which the batsman can fail to win the game. The batsman can fail to win the game because it is the verdict of the set of the game's rules that he doesn't win, or he can fail to win the game because body of rules of the game does not issue verdict on whether he has won. A person can not have the right to vote either because it is the pronouncement of the land's body of law that she doesn't have that right or because that body does not pronounce on whether she has that right.

My strong preference is for the third of these proposals. If different rules of the game issue conflicting verdicts on whether the batsman wins, then it is not the case that the batsman wins. The batsman wins if, and only if, that is the verdict of the set of rules of the game. One reason for this preference is methodological. Endorsing either contradictions or ideas of 'no fact of the matter' would seem to require classical logic to be given up, and given the alternative I do not see any good reason here to do this.^{iv} But this is not the place to explore such methodological matters. Indeed, there is not space here to explore fully the non-methodological issues surrounding these proposals. What I can do before moving on, rather, is outline how I would begin to defend the proposal I am preferring against an obvious objection.

It is a consequence of my preferred proposal that the batsman who hits a six into nettles neither wins nor loses, and so the cricket rule statements are both false. (As the batsman wins if, and only if, that is the verdict of the rules of the game, so too the batsman loses if, and only if, that is the verdict of the rules of the game.) This sounds, I think, a perfectly happy result. The result in the case of Jan's right to vote may, however, sound rather less happy. In this case it is a consequence of the preferred proposal that Jan does not have the right to vote. The rule statement that property holders have the right to vote will then be false and the statement that women do not have the right to vote will be true (although not true by decree – in making it a law of the land that women do not have the right to vote the legislators do not make it the case that the body of law of the land rules of any woman that she does not have the right to vote). But does this result not trample on an important symmetry? Can it really be right to say that Jan does not have the right to vote – are things not, as it were, *even* on that matter? Is there not a symmetry here such that the suggestion that Jan does not have the right to vote could amount at best only to a *stipulation* of some kind – to an attempt, as it were, to make it a further law of the land that if other laws conflict, there is no right.

A first response to this objection suggests that at least some of its plausibility derives from a failure to distinguish clearly between the theoretical question of whether someone has or does not have the right to vote and the practical question of whether someone should be allowed to vote or disallowed from voting. For there is indeed a symmetry at the level of this latter, practical concern. A person should be allowed to vote if, and only if, the body of law of the land gives them the right to vote, and they should be disallowed from voting if, and only if, the body of law of the land denies them the right to vote. In the case of Jan, the body of law of the land neither gives nor denies her the right to vote: it simply does not pronounce. So it is neither the case that she should be allowed to vote nor the case that she should be disallowed from voting. There is nothing the person running the polling booth should do on account of the law of the land in response to Jan's request for a ballot paper.

This first response may of course not quell the objection: someone may fully digest it – they may, that is, recognise that the practical symmetry should not be confused with a theoretical symmetry which it does not imply – and yet insist they weren't involved in any such confusion. Issues of who should be allowed to vote set clearly to one side, there is, it might be pressed, no more reason to say that Jan does not have the right to vote than to say that she does. And so we must say either both or neither: someone who says one but not the other could only be attempting – misguidedly – to stipulate the matter. Here the recommended account must hold the line. And it will do this, I would suggest, with the thought that having the right to vote (winning the game, losing the game) is, as it were, a positive attribute which by

default you don't have: it is an attribute which must be given to you by a body of rules if you are to have it. If the rules of the cricket game do not as a set make you the winner, then it is not the case that you are the winner. That is not a stipulation: it is not an implicit laying down of a further rule that if the other cricket rules do not as a set make you the winner then you don't win. Rather, it is the position that for any game with rules for winning, one wins that game only if it is the verdict of those rules that one wins. If a land's body of law does not give you a certain legal right, then you do not have that right even if that body also fails to deny it to you.

Section 2

2.1

Let's turn now to the paradox half of the paper's title. As indicated, I want to suggest that certain paradoxes can be seen as cases of conflicting rules. My key preparatory example of such a paradox will be Charles Chihara's Secretary Liberation Paradox.^v

Chihara invites us to imagine a land in which various clubs are established. Certain club secretaries of this land, the tale then proceeds, find that the eligibility rule of the club to which they are secretary excludes them from membership. In protest, these excluded secretaries band together and form a new club, the Secretary Liberation Club or SLC for short, in whose constitution they write: 'A person is eligible for the SLC if, and only if, they are secretary to a club for which they are ineligible'. The new club prospers and grows and after a while appoints its own secretary, Mr Finline, who as it happens is not secretary to any other club. Now this story may at first glance seem innocuous – but this impression will disappear if it is asked whether Mr Finline is eligible for the SLC. For it may be reasoned that if Mr Finline is eligible for the SLC then there is no club for which he is ineligible and for which he is secretary, and so he is ineligible for the SLC. But if Mr Finline is ineligible for the SLC then there is a club for which he is ineligible and for which he is secretary, namely the SLC itself, and so he is eligible for the SLC. We then conclude by negation introduction that Mr Finline is both eligible and ineligible to join the Secretary Liberation Club.

The premises of this paradox are:

- (a) A person is eligible for the SLC if, and only if, they are secretary to a club for which they are ineligible, and
- (b) Mr Finline is secretary to, and only to, the SLC.

From these two, the contradiction that Mr Finline is both eligible and ineligible for the SLC follows by standard, classically valid inferences. So assuming in line with the last section that classical logic is not to be disrupted, our first move in responding to the paradox will be to reject one or other premise. Rejecting premise (b) is, however, not an open possibility: it is a given within the story that Mr Finline is secretary to, and only to, the SLC.^{vi} We shall therefore have to reject premise (a). Premise (a) is inconsistent with the true premise (b) and so must be false.^{vii}

So far so good. The next move I want to make is to claim that despite being false, premise (a) nonetheless states a rule of the SLC. It is a rule of the SLC that a person is eligible if, and only if, they are secretary to a club for which they are ineligible. This claim is easily motivated. Writing an eligibility clause in a constitution is a standard way of laying down an eligibility rule of the content of that clause. Had the secretaries instead written down 'A person is eligible for the SLC if, and only if, they are a secretary', then their doing so would have made just this a rule of the club. So if it is held that the secretaries failed in their intention of laying down a rule of the content of their eligibility clause, a reason needs to be given explaining why. Furthermore, we may consider the behaviour of the SLC admissions committee. Several applications were made over the years to join the SLC. In response to each application, the admissions committee asked of the applicant whether they were secretary to a club for

which they were ineligible, and then accepted or rejected the application as this question received a positive or negative answer respectively. Now it seems straightforward to describe this behaviour as a following by the committee of the rule stated in premise (a). But if this description is accepted it follows – unless one were to hold further, but then bizarrely, that the committee’s following that rule is a mistake – that premise (a) does indeed state a rule of the club.

Why might one hesitate to endorse this apparently well motivated claim? Well, there are two related points of possible concern. The first of these has been noted: premise (a) is false. We have also earlier noted, however, that this is not in itself sufficient reason to deny that premise (a) states a rule of the club: a rule may be in force without its statement being true. Indeed, I shall press below that premise (a) may be seen to be a false rule statement of the same kind as certain false rule statements crossed above regarding cricket victory and the right to vote. Before going there, however, let’s consider a second point of possible concern. This point takes rise from the observation that the SLC constitutional eligibility clause is a conjunction of two quantified conditionals, namely that a person is eligible for the SLC if they are secretary to a club for which they are ineligible, and that a person is ineligible for the SLC if they are not secretary to a club for which they are ineligible. The would-be SLC eligibility rule may therefore be viewed as composed of two sub-rules, one of the content of each of the conjuncts. More, we can note by following the two arms of the paradox reasoning that these two sub-rules give conflicting pronouncements on the case of Mr Fineline. (The statement that a person is SLC eligible if they are secretary to a club for which they are ineligible has the consequence that Mr Fineline is SLC eligible, so the stated rule has the verdict that Mr Fineline is SLC eligible.) To claim that premise (a) states a rule of the club is thus to imply that when the constitution was written the founding secretaries laid down a pair of rules which contain between them a possibility of conflict. And with this in view we can now state the second point of concern. Consider the possibility that instead of writing what they did, the secretaries wrote instead: ‘A person is eligible for the SLC if they are a secretary and ineligible for the SLC if they are a secretary’. In this case, they would surely have failed to lay down a rule of the content of that sentence. Why? Well, because of the potential for conflict between the two would-be sub-rules. But if this is so, the concern is, then their actual writing will fail for the same reason to make what was actually written a rule of the club.

The obvious response to this second concern begins by recalling the point made in section 1 that there is a difference between the claim that there can be rules in force which contain a possibility of conflict and the claim that *any* rules can together be in force no matter how they would conflict. Just because the latter is false – and one could indeed cite as a counterexample here the rules that a person is SLC eligible if they are a secretary and that a person is SLC ineligible if they are a secretary – this does not mean that the former is false. Merely pointing out that the SLC secretaries would in another circumstance have failed to lay down a pair of rules because of their potential for conflict does not therefore entail that they actually so failed. Indeed, we can be rather more robust here. For so long as it is agreed that rules can be in force which contain a possibility of conflict – so long, that is, as one thinks there can be such a thing as a rule confliction –, then the SLC is surely as good an example of the phenomenon as one could hope to construct. If rule confliction is granted, this is to say, there is no reason concerning such confliction not to hold that premise (a) states a rule of the club. The secretaries *do* successfully lay down a rule of the content of premise (a), and this rule is – just as the objector suggests – composed of two sub-rules which give conflicting verdicts on Mr Fineline.

More strongly, there is then no reason at all not to hold that premise (a) states a rule of the club. In particular, the first concern that premise (a) is false is seen to be empty, for viewing the SLC as a case of rule confliction, premise (a) is seen to be a conjunction of statements of conflicting rules, and such conjunctions are always false. It is a rule of our cricket game both that the batsman wins if he hits a six and that he loses if he hits into nettles, but it is false that the batsman both wins if he hits a six and loses if he hits into nettles. Similarly, it is a rule of the SLC both that a person is SLC eligible if they are

secretary to a club for which they are ineligible and that a person is SLC ineligible if they are not secretary to a club for which they are ineligible, but premise (a) is nonetheless false.^{viii}

2.2

The perspective on the SLC paradox as a case of conflicting rules allows for the further work of section 1 to be brought to bear, providing us with a full response to the paradox. We can, this is to say, explain why premise (a) appears to be true, explain why what makes premise (a) appear true does not entail that it is true, and also give in place of premise (a) a positive account of SLC eligibility.^{ix} Let's take these three in turn.

Premise (a) appears to be true because it appears to be true by decree. The founding secretaries laid down the rules stated by premise (a), and to lay down the rule that p (to decree that p) is, in the way we have seen, standardly sufficient for making it the case that p. This is why the SLC scenario presents a paradox. Despite appearances, however, premise (a) is not true by decree: its inconsistency with premise (b) means it can't be true and so *a fortiori* can't be true by decree. But to say merely this is not yet to explain *how* premise (a) is not true by decree: it is not yet to explain how it is that in laying down the SLC rules the secretaries did not make premise (a) true. Our earlier work, however, provides for an explanation. Premise (a) is not true by decree because, indeed *only* because, of the possibility of conflict between the rules it states. Had the secretaries laid down just one of the two eligibility rules, the statement of that rule would have been true by decree. As things are, however, both rules are laid down, and their possible conflict has the disjunctive consequence that either a person may be secretary to a club for which they are ineligible and yet not be pronounced eligible by the set of SLC rules (Mr Fineline is ineligible), or a person may be secretary to no club for which they are ineligible and yet not be pronounced ineligible by the set of SLC rules (Mr Fineline is eligible). Thus at least one of the two rule statements will fail to be true by decree.

It is easy in the cases of the cricket game and the right to vote to see not just that at least one of the conflicting rule statements fails to be true by decree, but further that they both so fail. Indeed, this is visible in advance of sorting out their truth or falsity. Matters are, however, a little more delicate in the case of the SLC, and they are delicate in such a way that we cannot determine the failure of an individual rule statement to be true by decree in advance of sorting out who is, and who is not, SLC eligible. (The basic complication here is that, unlike in the cricket game and right to vote cases, whether a person meets the condition set out in the rule statement's antecedent may be not be an external matter determined independently of the SLC rules.) The overall picture remains nonetheless quite straightforward. In particular, the straightforward way in which we determined the circumstances of cricket game victory and of a person's having the right to vote carries over unproblematically to the SLC.

Running a story parallel to that provided for these earlier examples, we shall say that a person is SLC eligible if the set of SLC eligibility rules pronounces in their favour, and a person is SLC ineligible either if the set of SLC eligibility rules pronounces against them, or if that set fails to pronounce. (If the rules of SLC eligibility do not make you eligible for the club, then it's not the case that you're eligible for the club.) So, if Mr Smith is secretary to a women only tennis club then he is eligible for the SLC. If Mr Jones is not secretary to any club then he is ineligible for the SLC. And the failure of the set of SLC eligibility rules to issue a verdict on Mr Fineline means that he is ineligible for the SLC. From here the following results fall out. First, as it is the case both that the set of SLC eligibility rules do not pronounce on Mr Fineline, and that he is secretary to a club for which he is ineligible (namely the SLC), it is not true by decree that a person is SLC eligible if they are secretary to a club for which they are ineligible. Indeed, this half of premise (a) is false: Mr Fineline is ineligible for a club to which he is secretary and yet ineligible for the SLC. On the other hand, however, if a person is not secretary to a club for which they are ineligible, then the set of SLC rules will pronounce that person SLC ineligible. So it is not only true but furthermore true by decree that a person is ineligible for the SLC if they are not secretary to a club for which they are ineligible.

2.3

Before moving on to the Liar, let's pause first to make a comment on Mr Finline's SLC ineligibility, and second to digest more fully the result that premise (a) fails to be true by decree only because of the possibility of conflict between the rules it states.

As with the claim that Jan lacks the right to vote, one might feel uncomfortable at the claim that Mr Finline is ineligible for the SLC. Are things not even as to his being or not being SLC eligible? And should we not therefore be suggesting either that he is both eligible and ineligible or that there is 'no fact of the matter' as to his SLC eligibility? Must not any declaration 'breaking the symmetry' amount to a stipulation effectively changing the object of discussion? Here as before I want to remark that there is indeed *something* of importance hereabouts on which things are even, namely the practical matter of whether an application by Mr Finline to join the SLC should be accepted or rejected. The set of rules of the club neither pronounces Mr Finline eligible nor pronounces him ineligible, and so it is neither the case that his application should be accepted nor the case that his application should be rejected. Things are not, however, even on the matter of his SLC eligibility. In order to be eligible for the SLC you need to be made so by the set of rules of the club. This is not a matter of stipulation but reflects rather the general truth that club eligibility is a property conferred by club rules; if the rules don't give it to you, you don't have it. And the SLC eligibility rules do not give it to Mr Finline.

Next let's consider that premise (a) fails to be true by decree only because of the possibility of conflict between the rules it states. This entails that the falsity of premise (a) is not explicable independently of the potential conflict: an independent explanation of the falsity of premise (a) would be at once an independent explanation of why premise (a) is not true by decree. (Note that the explanation we have offered of the falsity of premise (a) is indeed not independent of the conflict.) But the possibility of conflict between the rules stated by premise (a) is constituted by (such inconsistencies as) the inconsistency of premise (a) with premise (b); the rules giving conflicting verdicts on Mr Finline is constituted by their statements having contradictory consequences regarding his SLC eligibility given the facts of his case. (This was the first constitution claim of section 1.2 above.) It follows, then, that the falsity of the rule stating paradox premise (a) is not explicable independently of the paradox; it is explicable only with reference to what is constituted by the paradox.

This is, I take it, an interesting result. To bring out how, consider the following passage from Susan Haack. Haack requires of a paradox solution:

that it should ... indicate which apparently unexceptionable premise or principle of inference must be disallowed ... ; and that it should, in addition, supply some explanation of why that premise or principle is, despite appearances, exceptionable. It is hard to make precise just what is required of such an explanation, but roughly what is intended is that it should be shown that the rejected premise or principle is of a kind to which there are independent objections – objections independent of its leading to paradox, that is. It is important, though difficult, to avoid supposed 'solutions' which simply label the offending sentences in a way that seems, but isn't really, explanatory. (Haack 1978 p.139)

There are, on the face of it, two possible motivations for insisting that a paradox solution provide a ground independent of the paradox for rejecting the false or invalid point of argument. The first of these is that there must be such a ground, and so a full solution will tell us what it is. The second, and this is Haack's concern, is that if no such ground is provided no genuine explanation will have been given of why the rejected point of argument is to be rejected.

There is indeed a danger that if one does not show a rejected paradox premise to be subject to an independent failing, one's would-be explanation of the premise's falsity will reduce to the claim that this premise is false because it must be false because it is inconsistent with the other true premises. This is,

however, only a danger. Our account of the falsity of premise (a) is genuinely explanatory, it does not reduce to mere finger pointing, and yet it does not show premise (a) to be subject to an objection independent of the paradox. One interesting lesson of the SLC example will therefore be that Haack is too quick in suggesting that if a putative explanation of why a paradox premise fails to be true is to be a genuine explanation, it must meet this independence requirement. A second interesting lesson of the example follows from the fact that the falsity of premise (a) is not *explicable* independently of the paradox. (It is explicable only with reference to the rule confliction – which is constituted by the paradox.) For one might in a broadly realist frame of mind think that paradoxes are, as it were, *our* problems: reality itself is constituted independently of such puzzles. This thought fails for SLC eligibility. Solving the SLC paradox does not involve correctly representing a reality of SLC eligibility and ineligibility constituted independently of the paradox: there is no such reality.

Section 3

3.1

These last remarks are rather short. It is not, however, my ambition in this paper to pursue their concerns in detail. What I want to do in this final section, rather, is suggest that the SLC can provide a model for thinking about certain other, philosophically more significant paradoxes.

Consider the Grelling paradox. This has as its key premise:

H: A predicate is heterological if and only if it does not apply to itself

To respond to the Grelling as we have responded to the SLC we would need to regard H as a rule statement. This would appear to be easily done: H states the rules for heterologicality laid down in an act of definition much as premise (a) states the rules for SLC eligibility laid down in the constitution. It may successfully be defined – stipulated, if you like – that a predicate is heterological just in case it does not apply to itself, just as it may successfully be decreed that a person is SLC eligible if and only if they are secretary to a club for which they are ineligible. As seen by its generation of paradox, however, premise H is false. This is due to the possibility of conflict between the rules it states – a possibility constituted by its generation of paradox.

A similar treatment is, on the face of it, available for the Liar. First here we need a position on truth bearers. For the purposes of this paper I shall take these to be sentences (and not worry too much what exactly a sentence is). Given, then, a sentence S which says that p, consider:

R_S: S is true if, and only if, p

R_S, the suggestion will be, is a rule statement: it states the rules for the truth of S, or the rules for truth in the case of S. For most sentences S, R_S will be true. There are, however, sentences the statement of whose truth rules is false. This is due to conflicts between truth rules, such conflicts being constituted by the Liar and related paradoxes.^x

Let's add a little flesh to this skeleton for the Liar by looking at a couple of examples. As indicated, the strategy will closely parallel that deployed for the SLC (and so also that for the cricket game and right to vote scenarios). Consider first the sentence:

J: Jack loves Jill

The truth rules for J are:

R1j: If Jack loves Jill, then J is true

R2_j: If Jack does not love Jill, then J is not true

Suppose that Jack loves Jill. Then the rule R1_j pronounces J true and the rule R2_j does not pronounce. Thus the set of truth rules for J pronounces J true, and this constitutes its being the case that J is indeed true. Alternatively, suppose that Jack does not love Jill. In this case R2_j pronounces J not true and R1_j does not pronounce. So the set of truth rules for J pronounces J untrue, and this constitutes its being the case that J is untrue. The *statements* R1_j and R2_j are, we can therefore see, both to be endorsed, the grounds for their endorsement comparing straightforwardly to our earlier discussions of ‘truth by decree’.

Next, consider a Liar sentence:

L: L is not true

It is no part of the response to the Liar being mooted to doubt that L says what it appears to say. The rules of truth for L are therefore taken as:

R1_L: If L is not true, then L is true

R2_L: If L is true, then L is not true

We readily see that L is true by the rule R1_L, and that it is not true by the rule R2_L. The pair of truth rules for L thus fails to pronounce on L, and so we conclude in line with our earlier examples that L is not true.^{xi} (If the rules for truth do not make a sentence true, then it is not true.) With this conclusion we then endorse the rule statement R2_L and reject the rule statement R1_L, thereby blocking the paradox argument.

To be clear: the position here is that what L says to be the case is indeed the case (L says that L is not true, and L is indeed not true), but L is nonetheless not true. This is of course odd, but we should not by now be surprised. Mr Fineline is, after all, secretary to a club for which he is ineligible, and yet ineligible for the SLC. The batsman hits a six but doesn’t win the game.^{xii}

3.2

These suggestions for the Grelling and the Liar have the payoff of providing a straightforward explanation of the paradoxes’ existence. H appears to be true because it states the rules for heterologicality, and rule statements are standardly true. We take R_s to be true because it states the rules for the truth of S. (And if S is a regular sentence such as ‘Snow is white’ then R_s is indeed true.) The obvious fact should be emphasised, however, that the suggestions made are not more than a preliminary to a full response to the Liar. Two points of underdevelopment stand out. First, an expansion is called for on what is meant by ‘rule for truth’. And second, a general strategy for determining the truth or untruth of sentences needs to be provided. There is not, unfortunately, space here for either of these goals to be properly pursued. Some quick comments may nonetheless be useful.

Regarding what is meant by ‘rules for truth’, I should be clear that these are not being understood as ‘semantic rules’ – they are not being understood, that is, as rules for the use of a sign in a certain language. The truth rules are no more rules of English than the cricket rules – and it was, of course, no part of the cricket story that the children are Anglophones! The rules regard, rather the property of truth. Just as the SLC rules are not in the first instance rules for the use of a certain English predicate but rather rules fixing whether a person has a certain property (the rules belong not to language but to the property of SLC eligibility), so the truth rules, as I would understand them, are not linguistic rules for the use of the English predicate ‘x is true’ but rules determining (nothing less than) whether a sentence is true.^{xiii, xiv}

Moving (swiftly) on to the question of how this determination works in general, certain further ‘case studies’ can easily be offered. Consider for example the Truth Teller:

T: T is true

Here we have the rules:

R1_T: If T is true, then T is true

R2_T: If T is not true, then T is not true

Neither R1_T nor R2_T pronounces on T's truth, and so nor does the set of rules with members R1_T and R2_T. T is thus untrue and the statements of R1_T and R2_T are both to be endorsed (though not as true by decree). Another interesting sentence is:

Y: If Y is true then Y is true

With rules:

R1_Y: If (if Y is true then Y is true) then Y is true

R2_Y: If (if Y is not true, then Y is true) then Y is not true

Here R2_Y does not pronounce but R1_Y pronounces Y true – and so Y is true.

Merely considering these two further examples does not, however, suggest a clear route to a general strategy. For a start, nothing so far said provides straightforwardly for a treatment of the pair of sentences:

U: V is true

V: U is not true

(We cannot reason here in parallel to J that if V is true R1_U pronounces U true and R2_U does not pronounce, for where Jack's loving or not loving Jill is a matter determined independently of the truth rules for J, the truth or untruth of V is not determined independently of the truth rules for U.) I nonetheless think that a full general treatment, including a provision for sentences such as U and V, is readily developable from the above work and suggestions. (By way of an indication: the broad result of such a treatment would be that the true sentences are those that fall in the extension of 'true' at Kripke's maximal intrinsic fixed point, with all other sentences being straightforwardly untrue.) As said, however, laying out even the outline of such a general treatment would take much more space than we have left.^{xv} What we can do instead to take us to the paper's end is see one way in which the general treatment would sophisticate what has come before, and then use that work to respond to the Curry paradox.

3.3

As so far understood, for a set of rules to have a certain verdict is for one or more members of that set to have that verdict and for no member of the set to have an incompatible verdict. But this understanding ignores that rules sometimes yield verdicts only in concert with other rules. To make good on this shortcoming (and this will be necessary to deal with such sentences as U and V) we need to introduce the notion of a *rule-combination* having a verdict. The natural definition to make here is that the combination of rules $\alpha_1 \dots \alpha_n$ regarding, say, the right to vote has the verdict that a certain person has or does not have the right to vote if, and only if, that person's having or not having the right to vote is a consequence of the statements of $\alpha_1 \dots \alpha_n$ together with relevant facts of the case. The definition of a set of rules having a certain verdict is then complicated in the obvious way. For a set of rules to have a certain verdict is for one or more members or combinations of members of that set to have that verdict, and for no member or combination of members of that set to have an incompatible verdict.^{xvi}

Consider now a Curry sentence:

C: If C is true then grass is orange.

This has as its truth rules:

R1_C: If (if C is true then grass is orange), then C is true

R2_C: If (it is not the case that if C is true then grass is orange), then C is not true

C is of interest because from the statements of R1_C and R2_C one may reason to the conclusion that grass is orange without going via a contradiction. Our treatment is as follows. Suppose C is true. Then from the statement R2_C it follows that if C is true then grass is orange. So it follows from the statement R2_C that if C is true then grass is orange. From here we can go two ways: we can either bring to bear the fact that grass is not orange and conclude that by the rule R2_C C is not true, or we can bring the statement R1_C to bear and conclude that by the R2_C-R1_C rule combination C is true. There is, then, no overall verdict from the set of rules {R1_C, R2_C}. C is therefore not true, and the statement R1_C is false. (The statement R2_C is true, but not by decree.)

Now it is notable that what we say about the given Curry sentence C, and so what we say about the statements R1_C and R2_C, depends upon the fact that grass is not orange. If grass were orange, then neither R1_C nor R2_C would pronounce by themselves on C, but the R2_C-R1_C combination would pronounce C true. In this case, then, C would be true as would the statements of both R1_C and R2_C. This highlights an important aspect of what is going on here, an aspect worth reemphasising, namely that the paradox generating truth rule statements that are being found false are not being so found for reasons independent of the paradox they generate. The story, rather, is as follows. As with cricket victory, SLC eligibility and the rest, we take the rules relevant to a particular case and see how those rules pronounce or fail to pronounce given the circumstances of the case. From here we reach a conclusion first on the case in view (i.e. on whether the sentence is true, on whether a certain type of person is SLC eligible etc.), and then second on the truth or untruth of the rule statements. In cases of paradox, this procedure involves deciding the status of the case, and so then also of the rule statements – which are paradox premises – on the basis of the rule confliction: on the basis, that is, of what is constituted by the paradox.

This last point is, I take it, of intrinsic interest. The truth or untruth of a sentence is not in general a matter independent of the Liar and related phenomena: truth has, in this sense, no paradox-independent nature. Additionally, though, it is important because it suggests that the response offered in this paper to the Liar is immune to revenge problems. It is a common feature of attempts to block the Liar paradox that the material they introduce to effect the block may be used to generate a related paradox for which the block is ineffective. If, however, one is rejecting a paradox premise on the ground – crudely – that it leads to paradox, then it is hard to see where the revenge might come from. The block would appear to be universally applicable.^{xvii, xviii}

References

- Azzouni, J. The Inconsistency of Natural Languages: How We Live with It. *Inquiry* 50:6, 590-605
- Azzouni, J. *Tracking Reason: Proof, Consequence, and Truth*. Oxford: OUP
- Burgess J. and Burgess A. (2011). *Truth*. Princeton: Princeton University Press

- Chihara, C. (1979). The Semantic Paradoxes: A Diagnostic Investigation. *Philosophical Review* 88(4), 590-618
- Eklund, M. (2002a). Deep Inconsistency. *Australasian Journal of Philosophy* 80, 321-331
- Eklund, M. (2002b). Inconsistent Languages. *Philosophy and Phenomenological Research* 64, 251-275
- Haack, S. (1978). *Philosophy of Logics*. Cambridge: CUP
- Patterson, D. (2009). Inconsistency Theories of Semantic Paradox. *Philosophy and Phenomenological Review* 79, 387-422
- Priest, G. (2006). In *Contradiction*. Oxford: OUP
- Scharp, K. (2007). Replacing Truth. *Inquiry* 50:6, 606-621
- Tarski, A. (1983). The Concept of Truth in Formalized Languages. In Corcoran ed., *Logic, Semantics, Metamathematics*. Indianapolis: Hackett

ⁱ I am grateful to Keith Allen for this example.

ⁱⁱ I assume in this paper that we should indeed not follow Priest to the contradiction. In assuming this I am not being dismissive of dialetheism – this paper is simply not the place to argue against it.

ⁱⁱⁱ I have had the response to these examples of rule confliction that there is in fact no contradiction in the offing here – there is nothing contradictory either to embrace (Priest) or to avoid (me). This is because a meaningful claim that X wins the game always makes reference, implicitly or explicitly, to a rule or rules by which X is claimed to win. There is no such thing as winning (or having the right to vote) *simpliciter*; there is only winning by rule(s) R. So there is no meaningful Priestian contradiction that the batsman both wins *simpliciter* and does not win *simpliciter*.

I do not think such a position is tenable. Consider the rule statement that the batsman wins if he hits a six. If there is no winning *simpliciter*, then this statement, more fully spelled out, will have to be that if the batsman hits a six, he wins by rule(s) R. But by what R? By the same rule again? If so, the rule statement will be that if the batsman hits a six, he wins by the rule that (if the batsman hits a six, he wins). But then this will need further spelling out as the statement that if the batsman hits a six, he wins by the rule that (if the batsman hits a six, he wins by the rule that (if the batsman hits a six, he wins)) – and we have a regress. Alternatively, the rule(s) R implicitly referred to in the rule statement that the batsman wins if he hits a six is a different rule. But which? And even if a well motivated answer can be given to this, R will have to be a rule which speaks of winning, in which case we shall still have a regress. The obvious solution here – and something I take to be clear in any case – is that there *is* such a thing as winning a game *simpliciter*, as well as such a thing as winning by rule R (that is, as rule R having the verdict that one wins *simpliciter*), and that the winning spoken of in our cricket rule statement is, as it seems to be, just such winning *simpliciter*.

^{iv} It may be that ideas of ‘no fact of the matter’ can be developed within a classical frame, and indeed that this is a promising strategy in the case of vagueness. More generally, then, my methodological preference here is for simplicity: in order to characterise what is going on with Jan or with the cricket game there is on the face of it no need either to drop classical logic or to appeal to a theoretically delicate idea of ‘no fact of the matter’. Let’s pursue, rather, the straightforward option.

^v See Chihara 1979 pp.593-5

^{vi} Premise (b) may, if necessary, be supported as follows. First, there is a club called the Secretary Liberation Club. This club was founded, has a constitution, has meetings and engages in various other

club-like activities; it is listed in the national register of clubs. Second, Mr Finline is a secretary to this club. It is sufficient for being a secretary to a club that one perform the function of a secretary to it, and Mr Finline performs the function of a secretary to the SLC. Third, Mr Finline is not secretary to any other club.

^{vii} Chihara agrees, claiming that premise (a)'s inconsistency with certain "empirically determinable" facts – namely those asserted by premise (b) – means that it must be rejected (Chihara 1979 p.594). Priest, on the other hand, views the SLC paradox as giving rise to a rule based dialetheia – that Mr Finline is both eligible and ineligible for the SLC – just as do the laws of Jan's state (see Priest 2006 p.200).

^{viii} We don't of course *have* to 'divide up' the SLC rule to see the SLC as a case of rule confliction. It is a consequence of the 'undivided' rule statement both that Mr Finline is eligible and that he is ineligible, and so the undivided rule gives the two conflicting verdicts that he is eligible and that he is ineligible. The undivided SLC rule conflicts with itself in the case of Mr Finline. This noted, I shall stick below for simplicity of exposition to the 'two rules' perspective – though nothing of substance depends upon it.

^{ix} It is notable that Chihara does not do this. More specifically, he neither explains how what makes premise (a) appear true does not entail that it is true (he says only that its inconsistency with premise (b) means that it can't be true), nor says anything about what if not premise (a) is true regarding SLC eligibility. Roy Sorenson briefly mentions the paradox in his 2001 (see pp.76-77), but he too does not get beyond the claim that premise (a) is/must be false.

^x This suggestion bears an obvious similarity to that of 'inconsistency theorists' such, recently, as Jody Azzouni, John and Alexis Burgess, Matti Eklund, Douglas Patterson and Kevin Scharp (see Azzouni (2006) and (2007), Burgess and Burgess (2011), Eklund (2002a) and (2002b), Patterson (2009) and Scharp (2007)). There is also an important basic difference; for slightly more detail see note 13 below.

^{xi} L is also not false. Paralleling the truth rules, the falsity rules for a sentence S which says that p are:

RFS: S is false if, and only if, not-p

In the case of L, which says that L is not true, the rules are: L is false if, and only if, L is true. But it is independently established that L is not true. So nor is it false.

^{xii} I am implicitly rejecting the assertion rule that one should author an assertoric sentence only if that sentence is true – I write 'L is not true' even though that sentence is not true. The version of the rule I prefer, along with many deflationists, is that one should author an assertoric sentence which says that p only if p.

^{xiii} This is, I think, a point of serious divergence from both Chihara and the more recent 'inconsistency theorists' mentioned in note 10 above. These writers all, I believe, want to press some version of Tarski's suggestion that the Liar indicates *natural language* to be in some way inconsistent (see Tarski (1983) pp. 164-5). So Chihara writes that '[a]n accurate statement of what 'true' means' will be false' (Chihara 1979 p.607)), and Eklund that 'the meaning of 'true' is such that 'S and "S is true" come out equivalent without S and "S is true" actually being equivalent' (2002b p.259)). There is, as said, an obvious similarity here to my claim that certain truth rules have false statements, but also an important difference: I'm concerned not with semantic rules for the word 'true' but with rules for the property of truth.

^{xiv} Setting aside the drift of my stage-setting examples, I'm motivated to insist that I'm not concerned with language here by the fact that that position – assuming it can be made good – protects me from a couple of objections. If one is thinking in terms of rules for the use of words, then it may be natural to see the Liar paradox as a result equally of the classical rules for the logical expressions as of the rules for the use of 'true'. And so my focus on the truth rules might look unmotivated: Why say that the truth rules are in force but have false statements rather than saying that the classical logical rules are in force but have false statements. Surely one should, given the perspective being adopted, think that everyone has equally dirty hands here. (This is Eklund's position: see, e.g., Eklund (2002a).) A second challenge which I would face if I were to think of truth rules as linguistic would be to my view that truth is a positive status which by default one does not have. 'x is not true' is, one might press, just as good a predicate as 'x is true'. So why should falling under the one predicate be a positive status whilst falling under the other be a default – are they not both equally positive statuses? My rejection that the truth rules are linguistic denies these two challenges any obvious purchase.

^{xv} What is crucially needed, amongst a few other things, is an account of which rules are to be applied to which sentences. In the examples we have run through, we have applied only its own rules to the sentence in view. But this is not always what is to be done. In the case of U and V, for example, the general treatment will have all four truth rules applied to U and all four to V.

^{xvi} This redefinition is of course retrospective. It does not, however, bring into question any of our earlier results.

^{xvii} Consider as a would-be revenge sentence:

(Z) The rules of truth for Z either pronounce against Z or do not pronounce on Z

Suppose Z is true, then given the statement $R2_Z$ it follows that the rules of truth for Z either pronounce against Z or do not pronounce on Z, and so that Z is not true. Thus by $R2_Z$ Z is not true. If we suppose that Z is not true, on the other hand, then given the statement $R1_Z$ the rules of truth neither pronounce against Z nor fail to pronounce. So the rules must pronounce in favour and Z is true. Thus by $R1_Z$ Z is true. As a set, then, the rules of truth for Z do not pronounce on Z, and so Z is not true and the statement $R1_Z$ is false.

^{xviii} Thanks to Steinvör Árnadóttir and Matti Eklund for very helpful discussion.