IMPROPER SELF-REFERENCE IN CLASSICAL LOGIC AND THE PREDICTION PARADOX

M. J. O'CARROLL

INTRODUCTION

Many classical paradoxes arise from statements resembling S : ~S (here : is used in defining a statement and should be read "which reads"). Mackie (1) has classified some improper self-referring statements. This paper classifies all self-referring statements.

For self-referring statements involving contingencies, it may not be easy to decide *a priori* whether the statement is proper. However we have the alternative that either the statement is improper or some contingent statement is a tautology. The logic may then be applied as if the statement were proper and some conclusion may be reached. If it turns out, *a posteriori*, that the statement was proper, then the conclusion is valid. Otherwise the logic has been applied outside its field of valid application and the conclusion is not necessarily true.

The prediction paradox (2) has received much attention but has not previously been reduced to its simplest terms nor even correctly formulated. In this paper (section 3) it is reduced to one of the improper forms classified.

1. CHARACTERISATION OF IMPROPER SELF-REFERENCE

Consider an axiomatic two-valued non-levelled logic. A self-referring statement

S : f(S)

is consistent within the logic if and only if the relation

$S \equiv f(S)$

does not violate the axioms.

The general function f(S) is

(S.1) v (~S.2)

where 1 and 2 are arbitrary constants. The statement S : f(S) is then proper if and only if 1 v ~2, in a logic which does not exclude non-definitive statements like A : A.

The following cases are easily demonstrated:

(i) If 1 . ~2, then S is non-definitive [f(S) reduces to S].
(ii) If 1 . 2, then S is true [f(S) reduces to S v ~S].
(iii) If ~1 . ~2, then S is false [f(S) reduces to ~S . (1 v 2)].
(iv) If ~1 . 2, then S is liar-paradoxical (1) [f(S) reduces to ~S].

If improper statements S : f(S) were admitted to the logic, then from S v ~S a deduction about 1 and 2 could be made, where 1 and 2 are arbitrary. This can violate the axioms of the logic. So for a consistent (two-valued, non-levelled) logic, some self-referring statements must be forbidden: namely, any of our statements S : f(S) in which ~1 . 2 is not always false.

What view should be taken with improper self-referring statements? There have been descriptions in the literature to the effect that the speaker utters the appropriate sounds but fails to make the statement. I feel that a better view is that the speaker makes a statement which is outside the field of application of classical logic.

Can we expect to find a logic which both applies to the field of all human statements and has a "correct" sense of implication? Can we determine such a thing as a "correct" sense of implication? Look at physics. Physical behaviour is observable only approximately. We can never determine exact physical behaviour. Consequently we can never find exact physical theories. Perhaps we have a similar situation with logic.

2. Paradoxical self-reference discussed by Prior (3) and Mackie (1)

Many improper self-referring statements in well known paradoxes become liar-paradoxical (i.e. like S : ~S which if true is false and if false is true) unless a pragmatic statement is universally true (or false).

The Cretan statement "All Cretan statements are false" is of the form

S : ~S . 2

where 2 is "any other Cretan statement is false". This is our general statement

$$S : (S.1) \vee (\sim S.2)$$

where 1 is universally false. Thus S is liar-paradoxical unless 2 is false. That is, unless there is another Cretan statement which is true.

Mackie characterises the improper self-reference in the special cases

$S : S.A$ whence either $\sim A$, or S is non-definitive
$S' : S' \vee A'$ (same as $S : S.A$ with $S' \equiv \sim S$ and $A' \equiv \sim A$)
$S : \sim S. \sim A$ whence either A, or S is liar-paradoxical
$S' : \sim S' \vee \sim A'$ (same as $S : \sim S. \sim A$ with $S' \equiv \sim S$ and $A' \equiv \sim A$)

which are all included in the general case by suitable choice of constants 1 and 2.

However the paradox of Buridan is more general than these cases of Mackie. Essentially it is

A : any statement.
B : Of A and B, one is false and one is true.

Thus B is

$$B : (A. \sim B) \vee (\sim A.B)$$

which is of the form

$$S : (S.1) \vee (\sim S. \sim 1).$$

Either S is liar-paradoxical or 1 is true. That is either B is liar-paradoxial or A is false.

3. The Prediction Paradox

The literature on this paradox (2) has not produced a correct formulation.

A teacher announces

1 : there will be an examination on one and only one afternoon next week, and

2: this examination will be unexpected.

Students argue the last day is out, since the exam would be expected by them. So is the last but one, etc. The teacher maintains that an exam on Wednesday cannot be expected.

This paradox has been stated formally as follows:
Let p, q, ..., etc., be "The exam will be on Monday and on no other day", etc. Let p′, etc. be "The exam will not be before Monday" etc... This is fact students would have observed on the morning of the examday. The teacher's statements then are (for a five-day week)

$$1: p \vee q \vee r \vee s \vee t$$
$$2: (\pi)\,[(D \supset \pi) \supset \sim\pi]$$

where π runs over p, q, ... and D is the allowable premiss for deducing the exam day. What has previously been considered is

$$D \equiv 1.2.\pi'.$$

But this is not paradoxical. It supports the students' argument, and D is inconsistent. Then the teachers' argument breaks down.

In fact $1.t' \supset t$ and so we have $2 \supset \sim t$. Then $1.2.s' \supset 1.\sim t.s' \supset s$, and so we have $2 \supset \sim s$. And so on, until eventually $2 \supset \sim 1$, so that D is inconsistent. The students can "predict" any statement π by a proof $1.2.\pi' \supset \pi$, even though they can also produce a counter proof $1.2.\pi' \supset \sim\pi$ (actually $2 \supset \sim\pi$). But they have no valid prediction for their premiss is inconsistent. There is no paradox. The statement 2 is merely equivalent to ~ 1 (for any choice of D):

$$\begin{aligned} 2 &\equiv (\pi)\,[(D \supset \pi) \supset \sim\pi] \\ &\equiv (\pi)\,[\sim(\sim D \vee \pi) \vee \sim\pi] \\ &\equiv (\pi)\,[(D.\sim\pi) \vee \sim\pi] \\ &\equiv (\pi)\,(\sim\pi) \\ &\equiv \sim 1. \end{aligned}$$

The teacher rightly maintains that (say) r cannot be predicted validly but wrongly assumes $1.2.r' \supset r$ is a valid prediction.

However, what is meant but not written in the literature, is to require D to be consistent. The effect is achieved by defining an unexpected exam as one which cannot be deduced (from $1.2.\pi'$)

without there being a counter-deduction. That is, the teacher says

$$1: p \vee q \vee r \vee s \vee t$$

$$2: (\pi)\,[\{(1.2.\pi' \supset \pi).\sim(1.2.\pi' \supset \sim\pi)\} \supset \sim\pi].$$

Neither the teacher's nor the students' argument can be carried out in this case. But there is a paradox. The statement 2 is simply equivalent to $\sim 1 \vee \sim 2$ as follows:

$$\begin{aligned} 2 &\equiv (\pi)\,[\{(D \supset \pi).\sim(D \supset \sim\pi)\} \supset \sim\pi] \\ &\equiv (\pi)\,[\sim(\sim D \vee \pi) \vee (\sim D \vee \sim\pi) \vee \sim\pi] \\ &\equiv (\pi)\,[(D.\sim\pi) \vee \sim D \vee \sim\pi] \\ &\equiv (\pi) \sim(D.\pi) \\ &\equiv \sim 1 \vee \sim 2 \vee [(\pi)(\sim\pi)] \\ &\equiv \sim 1 \vee \sim 2 \end{aligned}$$

This is the same as the improper self-reference

$$S : \sim S.1$$

where S is ~ 2. Hence 2 is liar-paradoxical unless 1 is false.

In words, the *conclusion* is: *either* it is not true that there is an exam on one and only one afternoon "next week" *or* the teacher's statement 2 falls outside the field of valid application of two-valued, non-levelled logic.

4. An imitation of the Goedel statement

Nerlich (2) discusses in this context an imitation of the Goedel statement.

If we try to reproduce in a two-valued logic, the statement

G : It is not provable that G,

G must state from what premiss G is to be not provable.

In this logic G is provable as follows:

If $\sim$G, then G is provable, then G is true. Therefore G is true. Now the fact that G is provable implies G is false. Thus G is paradoxical.

Here the premiss for provability of G is trivial (that is, a tautology). If we write G now as

$$G : \sim(t \supset G)$$

this is immediately

$$G : \sim G$$

which is of course liar-paradoxical.

This imitation is weak for we have no way of saying "provable" which is different from simply saying "true".

University of New England,
Armidale, N.S.W., Australia M. L. O'CARROLL

REFERENCES

[1] J. L. MACKIE, *Notre Dame J.F.L.*, vol. 2, p. 236, 1961.
[2] D. J. O'CONNOR, *Mind*, vol. 57, p. 358, 1948.
L. J. COHEN, *Mind*, vol. 59, p. 86, 1950.
P. ALEXANDER, *Mind*, vol. 59, p. 538, 1950.
M. SCRIVEN, *Mind*, vol. 60, p. 403, 1951.
P. WEISS, *Mind*, vol. 61, p. 265, 1952.
W. V. QUINE, *Mind*, vol. 62, p. 65, 1953.
R. SHAW, *Mind*, vol. 67, p. 382, 1958.
D. KAPLAN and R. MONTAGUE, *Notre Dame J. F. L.*, vol. 1, p. 79, 1960.
G. NERLICH, *Mind*, vol. 70, p. 503, 1961.
B. MEDLIN, *American Philosophical Quaterly*, vol. 1, p. 66, 1964.
[3] A. N. PRIOR, *Notre Dame J. F. L.*, vol. 2, p. 16, 1961.