On paradox without self-reference

NEIL TENNANT

Stephen Yablo [4] asks

> Why are some sentences paradoxical while others are not? Since Russell the universal answer has been: circularity, and more especially self-reference.

This is not strictly true. The proof-theoretic diagnosis of paradoxicality in [3] did not make self-reference either necessary or sufficient for paradoxicality. The purpose of this note is to show how that proof-theoretic diagnosis takes care of Yablo's ingenious example of a sequence of sentences giving rise to paradox without self-reference:

(S_1) for all $k>1$, S_k is untrue,
(S_2) for all $k>2$, S_k is untrue,
(S_3) for all $k>3$, S_k is untrue, ...

The proof-theoretic diagnosis of paradoxicality seemed, at the time it was put forward, to be borne out by all the familiar logical paradoxes. Yablo's paradox is an interesting new test for that proof-theoretic account.

The basic idea behind that account is that we need to distinguish straightforward inconsistency, or self-contradiction, of a set of assumptions, from paradoxicality. Both involve proofs of absurdity ($\perp$). But the proofs of absurdity in connection with straightforward contradictions are *normalizable*, whereas those in connection with paradoxes are not.

Proofs are normalizable when they can be brought into normal form by a finite sequence of applications of reduction procedures. These reduction procedures are designed to get rid of unnecessary prolixity. Such prolixity can arise, most importantly, by applying an introduction rule for a logical operator and then immediately applying the corresponding elimination rule. The result is a sentence occurrence within the proof standing as the conclusion of an application of the introduction rule and as the major premiss of an application of the corresponding elimination rule. Reductions get rid of such 'maximal' sentence occurrences, which stand as unwanted 'knuckles' in the proof. The reduction procedures for the logical operators are designed to eliminate such unnecessary detours within proofs.[1]

[1] For a full statement of the reduction procedures for all the usual logical operators, see [2].

ANALYSIS 55.3, July 1995, pp. 199–207. © Neil Tennant

So are other abbreviatory procedures σ, which have the general form of 'shrinking', to a single occurrence of A, any logically circular segments of branches (within the proof) of the form shown below to the left:

$$\begin{array}{c} A \\ \hline B_1 \\ \vdots \\ B_n \\ \hline A \end{array} \quad \overset{\sigma}{\Rightarrow} \quad A$$

One thereby identifies the top occurrence of A with the bottom occurrence of A, and gets rid of the intervening occurrences of $B_1, \ldots, B_n$ that form the filling of this unwanted sandwich. Logically, one should live by bread alone.

Once a proof is in normal form, none of the reduction procedures can be applied to it any more. The normal form proof will contain no maximal sentence occurrences, and no logically circular branch segments: no unwanted knuckles and no sandwiches. Normal form proofs can be thought of, informally, as non-meandering, non-roundabout, non-question-begging demonstrations of a conclusion's *resting on* certain assumptions. When the conclusion is absurdity ($\perp$) then we have a non-meandering, non-roundabout, non-question-begging demonstration of the *straightfoward inconsistency* of the set of undischarged assumptions.

Not so, however, with paradoxicality. What one finds, upon examination of the proofs of absurdity in connection with paradoxes (both well-known and not-so-well-known) is that *they are not in normal form* and *they cannot be normalized.*[2] Moreover, they seem to establish absurdity, in some cases (namely, the pure paradoxes, such as The Liar) 'out of thin air'. It is as though they show the incoherence of the very language itself, rather than the inconsistency of any particular (non-empty) set of statements within the language. By the time one reaches absurdity as the conclusion of a proof for a pure paradox, one has discharged all assumptions. Furthermore, if one applies the reduction procedures, the resulting reduction

[2] As I put it in [3], at p. 269: 'While every paradox is associated with a proof of [absurdity], *the proofs concerned do not reduce to ones in normal form.*' (Emphasis in the original.) Later, at p. 271, I wrote 'I propose precisely the test of non-terminating reduction procedures.' After examining familiar paradoxes, however, and finding that non-termination in those cases was due to looping, I inadvisedly strengthened my claim in the completeness conjecture at p. 283: 'A set of sentences is paradoxical relative to M iff there is some proof of [absurdity] from Θ(M), involving those sentences in *id est* inferences, that has a looping reduction sequence.' I should instead have written: '... that has a non-terminating reduction sequence.' This much is now clear in the light of Yablo's example. But nowhere in my account in [3] did I lay the blame for paradoxicality on self-reference.

sequence does not terminate. Indeed, in all cases I was able to examine in [3] (such as The Liar Paradox, Grelling's paradox, the postcard paradox etc.) the reduction sequence enters a loop. After several reduction steps, one encounters a proof that has already appeared earlier in the reduction sequence – usually, the very proof (not in normal form) with which one started.

I shall illustrate this rather abstract description with the example of the Liar sentence and the looping reduction sequence triggered by the proof of absurdity associated with it. We have the following two rules of inference for the truth predicate (disquotation and quotation):

$$\frac{T\phi}{\phi} \qquad\qquad \frac{\phi}{T\phi}$$

and we have the following two definitional rules, or *id est* inferences for the Liar sentence λ that says that λ is false:

$$\frac{\lambda}{\neg T\lambda} \qquad\qquad \frac{\neg T\lambda}{\lambda}$$

(I shall omit corner quotes throughout, since context will always make clear where these should be placed.) Consider now the following proof Σ of $\bot$ (absurdity):

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overline{T\lambda}^{(1)} \quad \dfrac{\dfrac{\overline{T\lambda}^{(1)}}{\lambda}}{\neg T\lambda}}{\bot}{\scriptstyle (1)}}{\neg T\lambda}}{\lambda}}{T\lambda} \qquad \dfrac{\dfrac{\overline{T\lambda}^{(2)} \quad \dfrac{\dfrac{\overline{T\lambda}^{(2)}}{\lambda}}{\neg T\lambda}}{\bot}}{\neg T\lambda^{*}}{\scriptstyle (2)}}{\bot}$$

Note that it contains no undischarged sentential assumptions. Note, moreover, that it is not in normal form, for the asterisked occurrence of $\neg T\lambda$ is maximal. That is, $\neg T\lambda$ stands there as the conclusion of an application of negation introduction, and as the major premiss of an application of negation elimination.

In order to try to normalize this proof, we have to apply the procedure $\neg R$ of negation reduction:

$$\dfrac{\begin{matrix}\Theta \\ P\end{matrix} \qquad \dfrac{\begin{matrix}\overline{P}^{(i)} \\ \Xi \\ \bot\end{matrix}}{\neg P}{\scriptstyle (i)}}{\bot} \qquad \overset{\neg R}{\Rightarrow} \qquad \begin{matrix}\Theta \\ [P] \\ \Xi \\ \bot\end{matrix}$$

When we apply this reduction procedure to the proof Σ, we obtain the following proof Σ_1:

$$\dfrac{\dfrac{\dfrac{\dfrac{\overline{T\lambda}^{(2)} \qquad \dfrac{\dfrac{\overline{T\lambda}^{(2)}}{\lambda}}{\neg T\lambda}}{\bot}\,{}^{(2)}}{\neg T\lambda}}{\dfrac{\lambda}{T\lambda}} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overline{T\lambda}^{(1)} \qquad \dfrac{\dfrac{\overline{T\lambda}^{(1)}}{\lambda}}{\neg T\lambda}}{\bot}}{\neg T\lambda^{*}}\,{}^{(1)}}{\lambda}}{T\lambda}}{\dfrac{\lambda}{\neg T\lambda^{*}}}}{\bot}$$

Notice, however, how the train of reasoning from the top asterisked occurrence of $\neg T\lambda$ to its bottom asterisked occurrence is prolix. These two occurrences of $\neg T\lambda$ ought to be identified. Once we shrink the branch in this way, we obtain a shorter proof: namely Σ, the proof we started with! Thus the reduction sequence loops after only two steps:

$$\Sigma \qquad \begin{array}{c}\neg R\\ \Rightarrow\\ \\ \Leftarrow\\ \sigma\end{array} \qquad \Sigma_1$$

Note also that the Liar sentence does not stand as an undischarged assumption of the proof Σ of absurdity. It merely mediates the convoluted moves we make within the semantically closed language that sustains the *id est* inferences above. The Liar is not a premiss (undischarged assumption) of a *reductio ad absurdum*. No inconsistency of belief has been established; only the occasional incoherence of a certain means of expression.

As already remarked, similar 'looping' reduction procedures are to be found with all the other well-known semantic paradoxes. The reduction sequences of the proofs of $\bot$ in the case of these paradoxes *do not terminate*; and similar remarks hold about how the sentences involved in them do not stand, individually or collectively, as assumptions being reduced to absurdity.

This feature struck me as the essence of paradoxicality. My proposal was that we had, in effect, the materials for an *axiomatization of paradoxicality* (in Kripke's sense in [1]) even if not a decision method, for any semantically closed language. Simply enumerate proofs of absurdity; start normalizing those that are not in normal form; and check to see whether the reduction sequences ever enter loops, or manifest any other conclusive evidence that they will not terminate. As soon as a reduction sequence does enter a loop, or manifest such evidence, one can check off the proof concerned as a 'paradoxical' proof.

As such, *these proofs were harmless*. For the only thing we should fear and try to avoid would be a *straightforward* proof of contradiction among statements we tried to make about the world. These statements would have to stand, collectively, as a set of undischarged assumptions of a *reductio* proof in normal form. Only then would it have been shown that absurdity lurked in one's representations *as these concerned the world*, and not in the convolutions of sentential reference and semantic predication within the semantically closed language to which they belonged. (Note that sentential reference is not necessarily self-reference. Yablo's example involves no self-reference, but it certainly does involve sentential reference – or, to be precise, quantification over sentences.)

I still find this proof-theoretic distinction between straightforward inconsistency and paradoxicality an attractive one. What is so interesting about Yablo's chain paradox, however, is that it shows that, if the proposed proof-theoretic test for paradoxicality is to pass muster, we have to find ways of detecting non-normalizability in cases other than the most obvious ones that involve looping reduction sequences. The novel feature of the Yablo sequence is that the non-normalizability of the proof in question does *not* consist in its initiating a looping reduction sequence of proofs, but rather in its initiating a non-terminating reduction sequence of a different *but easily recognizable* kind. Whether the means of such recognition could be smoothly generalized is what interests me.

To see how the proof of absurdity in connection with Yablo's paradox is indeed non-normalizable, consider his sequence of sentences S_0, S_1, ... involving the truth predicate T. I shall render the generic member of this sequence formally and uncontroversially as follows:

$$S_n\text{:}\quad \forall k{>}n\ \neg T(S_k)$$

We have, as before, the two rules for the truth predicate:

$$\frac{T\phi}{\phi} \qquad\qquad \frac{\phi}{T\phi}$$

and we have in this case the following two definitional rules, or *id est* inferences for the *n*th Yablo sentence that says that all succeeding sentences of the Yablo sequence are false:

$$\frac{S_n}{\forall k{>}n\ \neg TS_k} \qquad\qquad \frac{\forall k{>}n\ \neg TS_k}{S_n}$$

These last two rules are justified by the very meaning of each sentence in the Yablo sequence. For each sentence S_n *says that* $\forall k{>}n\ \neg TS_k$. Understanding S_n involves recognizing one's entitlement (and the entitlement of others) to the use of the *id est* rules just given.

Using our rules for disquotation and *id est* inference, we can construct the following proof $\Theta(m)$ of $\bot$ from the assumption TS_m. Note that $\Theta(m)$ is parametric in m: that is, any numeral could be uniformly substituted for m and the result would be a proof.

$$\Theta(m)\qquad \dfrac{\dfrac{\dfrac{\dfrac{TS_m}{S_m}}{\forall k{>}n\ \neg TS_k}}{\neg TS_{m+1}}\ \forall E \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{TS_m}{S_m}}{\forall k{>}m\ \neg TS_k}}{\forall k{>}m{+}1\ \neg TS_k}\ \text{by arithmetic}}{S_{m+1}}}{TS_{m+1}}}{\bot}$$

Now consider the following proof which I shall call $\Sigma(n)$.

$$\Sigma(n)\qquad \begin{array}{c} \dfrac{}{TS_m}\,(1) \\ \Theta(m) \\ \dfrac{\dfrac{\dfrac{\bot}{\neg TS_m}\,(1)}{\forall k{>}n\ \neg TS_k}\ m \text{ parametric for } \forall I}{S_n} \\ \overline{[TS_n]} \\ \Theta(n) \\ \bot \end{array}$$

A comment is required on the application of universal introduction here. The conclusion $\neg TS_m$, which is the premiss for this application, is parametric in m – that is, m does not occur in any assumption that is undischarged at that stage. Thus we *could*, if we had so wished, have inferred the stronger conclusion $\forall k \neg TS_k$ by universal introduction. But it suits our purposes at this stage to infer the weaker conclusion $\forall k{>}n\ \neg TS_k$ instead, for that is what gives rise, via the next *id est* inference, to S_n. In this proof the conclusion $\bot$ appears to come 'out of thin air'. Expanded so as to show the graftings, indicated by square brackets, onto the assumption occurrences of TS_n within the proof $\Theta(n)$, the proof $\Sigma(n)$ becomes the following:

$\Sigma(n)$

$$
\begin{prooftree}
\AxiomC{$\overline{\mathrm{TS}_m}^{(1)}$}
\noLine
\UnaryInfC{$\Theta(m)$}
\noLine
\UnaryInfC{$\bot$}
\RightLabel{(1)}
\UnaryInfC{$\neg\mathrm{TS}_m$}
\UnaryInfC{${}^*\forall k{>}n\ \neg\mathrm{TS}_k$}
\UnaryInfC{S_n}
\UnaryInfC{$[\mathrm{TS}_n]$}
\UnaryInfC{S_n}
\UnaryInfC{${}^*\forall k{>}n\ \neg\mathrm{TS}_k$}
\UnaryInfC{$\neg\mathrm{TS}_{n+1}$}
\AxiomC{$\overline{\mathrm{TS}_m}^{(2)}$}
\noLine
\UnaryInfC{$\Theta(m)$}
\noLine
\UnaryInfC{$\bot$}
\RightLabel{(2)}
\UnaryInfC{$\neg\mathrm{TS}_m$}
\UnaryInfC{$\forall k{>}n\ \neg\mathrm{TS}_k\ {}^{**}$}
\UnaryInfC{S_n}
\UnaryInfC{$[\mathrm{TS}_n]$}
\UnaryInfC{S_n}
\UnaryInfC{$\forall k{>}n\ \neg\mathrm{TS}_k\ {}^{**}$}
\UnaryInfC{$\forall k{>}n{+}1\ \neg\mathrm{TS}_k$}
\UnaryInfC{S_{n+1}}
\UnaryInfC{TS_{n+1}}
\BinaryInfC{$\bot$}
\end{prooftree}
$$

Note how the asterisked occurrences of the same sentence sandwich an unnecessary filling. These can be bypassed by identifying the sides of each such sandwich and leaving out the filling; that is, by applying an abbreviatory reduction:

$\Sigma_1(n)$

$$
\begin{prooftree}
\AxiomC{$\overline{\mathrm{TS}_m}^{(1)}$}
\noLine
\UnaryInfC{$\Theta(m)$}
\noLine
\UnaryInfC{$\bot$}
\RightLabel{(1)}
\UnaryInfC{$\neg\mathrm{TS}_m$}
\UnaryInfC{${}^*\forall k{>}n\ \neg\mathrm{TS}_k$}
\UnaryInfC{$\neg\mathrm{TS}_{n+1}$}
\AxiomC{$\overline{\mathrm{TS}_m}^{(2)}$}
\noLine
\UnaryInfC{$\Theta(m)$}
\noLine
\UnaryInfC{$\bot$}
\RightLabel{(2)}
\UnaryInfC{$\neg\mathrm{TS}_m$}
\UnaryInfC{$\forall k{>}n\ \neg\mathrm{TS}_k\ {}^{**}$}
\UnaryInfC{$\forall k{>}n{+}1\ \neg\mathrm{TS}_k$}
\UnaryInfC{S_{n+1}}
\UnaryInfC{TS_{n+1}}
\BinaryInfC{$\bot$}
\end{prooftree}
$$

We see here that the single asterisked occurrences of '$\forall k{>}n\ \neg\mathrm{TS}_k$' in the left subproof have become one, as have its double asterisked occurrences in the right subproof. In the resulting proof the rightmost occurrence of '$\forall k{>}n\ \neg\mathrm{TS}_k$' is now otiose; and the leftmost occurrence of '$\forall k{>}n\ \neg\mathrm{TS}_k$' is maximal, in that it stands as the conclusion of universal introduction and as the major premiss of universal elimination. By universal quantifier reduction and deletion of the otiose occurrence of '$\forall k{>}n\ \neg\mathrm{TS}_k$' we obtain

$\Sigma_2(n)$

$$\dfrac{\dfrac{\begin{array}{c}\overline{TS_{n+1}}^{(1)}\\ \Theta(n+1)\\ \bot\end{array}}{\neg TS_{n+1}}{\scriptstyle (1)} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}\overline{TS_m}^{(2)}\\ \Theta(m)\\ \bot\end{array}}{\neg TS_m}{\scriptstyle (2)}}{\forall k>n+1\ \neg TS_k}}{S_{n+1}}}{TS_{n+1}}}{\bot}$$

But in $\Sigma_2(n)$ the occurrence of '$\neg TS_{n+1}$' is maximal, in that it stands as the conclusion of negation introduction and as the major premiss of negation elimination. Negation reduction yields

$$\begin{array}{c}\dfrac{\dfrac{\dfrac{\begin{array}{c}\overline{TS_m}^{(1)}\\ \Theta(m)\\ \bot\end{array}}{\neg TS_m}{\scriptstyle (1)}}{\forall k>n+1\ \neg TS_k}\ m \text{ parametric for } \forall I}{S_{n+1}}\\ \hline [TS_{n+1}]\\ \Theta(n+1)\\ \bot\end{array}$$

This is just the variant $\Sigma(n+1)$ of our original proof $\Sigma(n)$, with $n+1$ in place of n in the lower half. It is clear that repetitions of our procedures applied thus far will in due course yield the variant $\Sigma(n+2)$ of our original proof, with $n+2$ in place of n in the lower half; and then the variant $\Sigma(n+3)$ of our original proof, with $n+3$ in place of n in the lower half....; and so on. This is a paradigm of a recognizably non-terminating reduction sequence, even though it is not, strictly speaking, a looping one. It is, rather, a spiralling one:

$$\Sigma(n) \overset{\sigma}{\Rightarrow} \Sigma_1(n) \overset{\forall R}{\Rightarrow} \Sigma_2(n) \overset{\neg R}{\Rightarrow} \Sigma(n+1) \Rightarrow \ldots \Rightarrow \Sigma(n+2) \Rightarrow \ldots \Rightarrow \Sigma(n+3) \ldots$$

Yablo's sequence would appear, therefore, to bear out my original diagnosis of paradoxicality. Its apparent *lack of self-reference*, which Yablo rightly took to be its most interesting feature, is reflected, on my account, in the non-looping character of the (nevertheless non-terminating) reduction sequence for the non-normal proof of absurdity that results 'from' Yablo's sentences. Note, however, that – as was the case with The Liar – his sentences do not stand as undischarged assumptions of the proof of absurdity in question. They merely mediate the convoluted moves we make within the semantically closed language that sustains the *id est* inferences above. Yablo's sentences are not premisses (undischarged assumptions) of a *reductio ad absurdum*. Once again, no inconsistency of belief has been established; only the occasional incoherence of a certain means of expression.

And what of self-reference? What is its distinguishing feature? It need not be obvious when self-reference is indeed involved, and is responsible for the paradoxicality of the set of sentences in question. I shall make so bold as to suggest that it is precisely when the non-terminating reduction procedures enter loops that self-reference is involved. And when they don't enter loops – as with Yablo's example – then self-reference is not involved.

It is high time, I think, for theorists of paradox to look at the resources available from proof theory in coming to grips with their subject matter. Proof theory is at the service of epistemology in semantically closed languages as well as in semantically open ones.[3]

The Ohio State University
Columbus, OH 43210-1365, USA
ntennant@magnus.acs.ohio-state.edu

References

[1] S. Kripke, 'Outline of a theory of truth', *Journal of Philosophy* 72 (1975) 690–716.
[2] N. Tennant, *Natural Logic* (Edinburgh: Edinburgh University Press, 1978; 2nd edn. 1990).
[3] N. Tennant, 'Proof and paradox', *Dialectica* 36 (1982) 265–96.
[4] S. Yablo, 'Paradox without self-reference', *Analysis* 53 (1993) 251–52.

[3] I would like to thank Peter Smith for his helpful editorial suggestions.