Ergo AN OPEN ACCESS JOURNAL OF PHILOSOPHY

Swyneshed Revisited

ALEXANDER SANDGREN
Australian National University

I propose an approach to liar and Curry paradoxes inspired by the work of Roger Swyneshed in his treatise on insolubles (1330-1335). The keystone of the account is the idea that liar sentences and their ilk are false (and only false) and that the so-called "capture" direction of the T-schema should be restricted. The proposed account retains what I take to be the attractive features of Swyneshed's approach without leading to some worrying consequences Swyneshed accepts. The approach and the resulting logic (called "Swynish Logic") are non-classical, but are consistent and compatible with many elements of the classical picture including modus ponens, modus tollens, and double-negation elimination and introduction. It is also compatible with bivalence and contravalence. My approach to these paradoxes is also immune to an important kind of revenge challenge that plagues some of its rivals.

1. Self-falsification

Roger Swyneshed's proposed solution to the liar paradox, developed in his treatise on insolubles (1330-1335),[1] is one of a family of proposals according to which liar sentences like (1) are only false (*Ins.* 15, 16), i.e. false, not true, and not both true and false.

(1) (1) is false.

I will often drop the 'only.' By 'false' and 'true' without qualification, I mean only false and and only true respectively, where 'only' excludes all other truth values including combinations of truth values.

1. Throughout I refer to Read's English translation of the Swyneshed's *Insolubilia* (henceforth, in citations, '*Ins.*') from Spade (1979), available at the Logic Museum (http://www.logicmuseum.com/wiki/Authors/Roger_Swyneshed/Insolubilia).

Contact: Alexander Sandgren <alexander.sandgren@gmail.com>

https://doi.org/10.3998/ergo.7961

According to Swyneshed, a sentence is false if it "falsifies itself" either directly or indirectly (*Ins.* 4, 5, 17, 18) and a sentence is true only if it does not self-falsify (*Ins.* 14; Spade 1983: 106)

A sentence falsifies itself if it implies that it is false (*Ins.* 4, 5, 7, 17).[2] As I will use the term 'implies,' A implies B just in case the conditional 'if A then B' is true. Put another way, the idea is that a sentence self-falsifies if its truth requires that it is false.

It is a somewhat complex question what, exactly, this kind of implication is. However, the implication is clear in many familiar cases. (1) intuitively implies that it itself is false, as Swyneshed (*Ins.* 5) says, directly. We will consider more indirect self-falsification below.

Swyneshed effectively proposes an extra way sentences can be false, and imposes an extra constraint on truth. According to Swyneshed, for p to be true, things have to be as p says they are (p must "signify as it is") *and* p must not self-falsify (*Ins.* 14). Compare the simple capture principle that says that for any sentence p, p implies that p is true.

Capture: $p \models T(p)$

'T(x)' is the truth predicate and '$\models$' indicates semantic entailment. Unless otherwise specified, lower-case letters like 'p' and 'q' designate truth-apt sentences, either as constants denoting particular sentences or as variables ranging over truth-apt sentences (like when they occur in schemata like Capture).

We can capture Swyneshed's central idea in terms of a restricted capture principle. Standard release remains.

Restricted Capture: $(p \wedge \neg SF(p)) \models T(p)$

Release: $T(p) \models p$

'$\wedge$' indicates conjunction, '$\neg$' indicates negation, 'SF(x)' is the 'self-falsifies' predicate, and 'T(x)' is the truth predicate. This reflects the idea that self-falsifying sentences are false even if things are, in some sense, as they say they are. This blocks a central step in liar-paradoxical reasoning. If (1) self-falsifies, it is false, even if things are as it says they are, thus blocking the inference from (1)'s being false to its being true. Self-falsifyingness is a *semantic* feature of some sentences; a sentence self-falsifies because of what it says about how things are (what it "signifies").

This paper is not intended as Swyneshed exegesis. The aim is to draw on and overhaul Swyneshed's central idea and develop an improved Swyneshed-inspired

2. One might worry that every falsehood will imply that it is false. I discuss this challenge in §3.2.

approach to truth and semantic paradoxes.[3] To have a label, call the proposal the "Swynish" approach to paradox. I will also attempt to extend Swyneshed's idea to yield what I take to be an attractive solution to Curry paradoxes (of both "flavors" (Beall & Murzi 2013)). In the process, I will begin to formulate a non-classical logic I'll call "Swynish Logic" that stems from the proposed Swyneshed-inspired treatment of truth and semantic paradoxes.

I will often contrast the Swynish approach with some components of a classical approach. This is to highlight what is distinctive about the Swynish approach. I will also be proceeding from a bivalent starting point, mainly to make the presentation of the solutions smoother. However, the question of how other truth values might be linked into the Swynish approach will come up. As will become clear, I am not particularly interested in keeping the Swynish approach as classical or bivalent as possible and the reader is welcome to combine their favourite deviations from classical logic with the Swynish ones outlined here.

2. Promise and Compromise

What are some appealing features of a Swynish approach to the liar and Curry paradoxes? To begin with, it does not require that liar sentences or Curry sentences do not express what they seem to express. There is no need to claim that the kind of self-reference apparently exhibited by (1) is impossible, as so-called "restrictionists" do (see, for example, Skyrms (1970)). Nor is there any need to claim that liar sentences like (1) do not really say anything, do not express propositions (Kripke 1975), or are somehow meaningless. On the Swynish proposal, sentences like (1) can signify exactly what they seem to signify about how things are.

On a related note, the proposal is also immune to a kind of revenge challenge that some other solutions to the liar face (Beall 2007; Shapiro 2011) in the face of which some deliberately limit their ability to express the tools used in their approach to paradox. Defenders of the Swynish approach can clearly delineate (and express) the properties that the sentences have to have for them to self-falsify without opening the door to revenge.

3. Another proposal according to which liar sentences are false is due to Bradwardine. Bradwardine's work almost certainly influenced Swyneshed's. However, Bradwardine's proposal to come with what seems to be unnecessary semantic baggage, including what Read (2002) and Restall (2008) call "Bradwardine's Axiom" and associated claims about semantics and entailment. Roughly, Brawardine's Axiom says that every sentence signifies everything which follows from it. As Restall (2008) discusses, this claim needs a good deal of unpacking if it is to be both fit for purpose and non-trivial. Swyneshed's solution is simpler and better suited to the kind of overhaul I propose below, so I will focus on developing a Swyneshed-inspired account.

According to the Swynish approach, there is no need to deny that a language can define its own truth predicate. Famously, Tarski's (1936) proposed solution to the liar paradox relies on this denial. Nor is there need to posit a hierarchy of truth predicates or languages, or to insist that languages must be semantically open. What is more, the truth predicate can be expressed and interpreted in a natural way.

The central elements of Swyneshed's proposal are also compatible with principles that some are interested in upholding such as Bivalence and Contravalence.[4]

Bivalence: Every truth-apt sentence is either true or false.

Contravalence: No sentence is both true and false.

The Swynish proposal does not force us to countenance a third truth value, as is required by other proposed solutions to the liar (Kripke 1975). However, there are options for incorporating a third truth value into a broadly Swyneshed-inspired approach, some of which will be discussed in §9.

If the proposed solutions to Curry paradoxes can be made to work, another desirable feature of the Swynish approach is that it lets us handle the liar paradox and Curry paradoxes using the same tools.

The solutions to both the liar and Curry paradoxes to be advanced here are compatible with the validity of Modus Ponens.

Modus Ponens: $(p, p \rightarrow q) \models q$

Depending on some choices about how to build a Swynish semantics for the conditional, Modus Tollens can also be vindicated.

Modus Tollens: $(\neg q, p \rightarrow q) \models \neg p$

For more discussion of this, see §5.1. Some other solutions to liar and Curry paradoxes compromise on the validity of Modus Ponens or Modus Tollens, so inasmuch as it is good not to threaten the validity of these inference patterns, the Swynish approach has a comparative advantage.

If we adopt the approach to the liar and Curry paradoxes defended here, there is no need stemming from these paradoxes to revise the structural rules of entailment or implication by rejecting, for instance, transitivity (Ripley 2013; Weir 2005; 2015), reflexivity (French 2016), or contraction (Petersen 2000; Zardini 2011; Beall & Murzi 2013; Ripley 2015).

4. This is a point of departure from Swyneshed, as he rejects Bivalence (93, 99), but this is not crucial for present purposes.

Finally, it seems to me that the Swynish approach to liar and Curry paradoxes is elegant. The Swynish changes to the theory of truth are straightforward and modest, and the associated claims about negation and the conditional are principled though non-classical. I invite the reader to agree with me on this point of elegance, but I suppose there is no accounting for taste!

If the history of the debate about the semantic paradoxes has revealed anything, it is that no genuine solution to these paradoxes will be entirely satisfactory. In the face of these paradoxes, some give up consistency, others give up their ability to express semantic values, yet others claim that apparently meaningful sentences do not really have meaning after all, yet others reject initially plausible structural rules for entailment, and so on. When proposing solutions to semantic paradoxes, no one gets to hold onto everything they might have wanted to hold onto. Instead, when comparing solutions we must evaluate trade-offs between different motivations, theoretical virtues, and so on. In other words, those engaged in this debate are in the compromise business.

In this spirit, I submit the Swynish approach as a promising line of compromise. Depending on how wedded one is to Bivalence, Contravalence, transitivity, expressive power, or what have you, the supposed advantages just mentioned will have more or less weight. On the other hand, the extent to which one is wedded to classical logic, transparent theories of truth, the standard treatment of negation, and so on, will affect whether the compromises I suggest here seem worth making. All the same, in what follows I hope to make the case that the Swynish approach is at least interesting and worthy of exploration.

3. Three Claims

Swyneshed (*Ins.* 25-27) makes three claims that he takes to follow from his approach.

(2) Things are as some false sentence says they are. (*Ins.* 25)
(3) There are valid arguments with only true premises and a false conclusion. (*Ins.* 26)
(4) There are pairs of contradictory sentences such that both of their members are false. (*Ins.* 27)

These claims are illustrative so I will consider them in turn.

3.1. Truth and Correspondence

The core idea behind (2) is that there are two kinds of conditions on truth; what we might think of as two dimensions of truth. One kind of condition concerns whether things *correspond* (in some crucial sense) to how the sentence says things are (Herzberger 1973).

Swyneshed's requirement that the sentence does not self-falsify is a special case of a wider range of extra conditions proposed by Bradwardine, Buridan and others. What is common to these approaches is the claim that the truth of a sentence is not just a matter of things corresponding to what the sentence signifies; there must be correspondence *and* sentences must satisfy a further condition. In Swyneshed's case, that condition is that sentences must not self-falsify.

(2) is a consequence of both Swyneshed's proposal and other proposals that involve a non-correspondence condition on truth. Correspondent sentences that do not meet a further condition on truth will be false even when things are, in an intuitive sense, as they say they are. So if accepting (2) is a cost, the defender of this kind of proposal must pay it.

3.2. Validity and Indirect Self-falsification

To motivate (3), Swyneshed (*Ins.* 26) presents a sentence like (5).

(5) The conclusion of this argument is false, therefore the conclusion of this argument is false.

(5) seems valid and I will take it as valid for now.

The conclusion is false according to Swynish approach. But the premise seems true; it says that the conclusion is false and the conclusion *is* false. So we have a valid argument that takes us from truth to falsity, or so Swyneshed claims. Swyneshed takes this to be an interesting discovery about validity, I take it to show that some part of his reasoning goes wrong. The fact that Swyneshed takes (3) to be an interesting discovery makes salient exegetical questions on which I will not dwell here: what exactly is Swyneshed's account of validity?, if validity does not preserve truth, what *do* valid arguments preserve? For a recent discussion of these questions, see Read (2020a).

How can a defender of an otherwise Swynish view avoid accepting (3)? Note that some sentences falsify themselves indirectly. Take a familiar liar cycle.

(6) (7) is true.
(7) (6) is false.

(6) is true only if it is false. (6) says that (7) is true (and is true only if (7) is true) and (7) says that (6) is false (and is true only if (6) is false). In this way, (6) falsifies itself indirectly via (7). The chain of requirements in virtue of which a given sentence

self-falsifies may be relatively long and can depend on the semantic features (and associated truth conditions) of other sentences.[5] In Swyneshed's (*Ins.* 5-7) terms, some sentences are "insufficiently relevant" to inferring themselves to be false. The sense in which, for example, (6) is "insufficiently" relevant to inferring itself to be false is that its self-falsifying depends on contingent facts about the semantic features of another sentence, namely (7).

As Kripke (1975: 692) observes, some paradoxical sentences are paradoxical because some extrinsic feature of the world makes them so:

> an adequate theory must allow our statements involving the notion of truth to be risky: they risk being paradoxical if the empirical facts are extremely (and unexpectedly) unfavorable.

By saying something about the truth of a sentence other than itself, (6) makes itself vulnerable to being paradoxical if the world does not cooperate. There is no *intrinsic* feature of (6) that, on its own, makes (6) paradoxical. For instance if (7) just said that Frank Jackson wrote *Epiphenomenal Qualia*, (6) would not self-falsify. So we can say things like "given the contingent fact that (7) says that (6) is false, (6) implies that it, (6), is false."

Returning to (5), I propose that if (5) is valid, the premise indirectly falsifies itself via the argument. This would mean that the premise is false and we could avoid accepting (3) for the reasons Swyneshed gives.

Supposing (5) is valid, the only way for the conclusion (*C*) to be false, as the premise (*PR*) says it is, is for *PR* to be false as well. Below is a table specifying the possibilities. The grey row is excluded when we suppose that the argument is valid. The 'I' in the middle column indicates invalidity.

PR	$\models$	*C*
T		*T*
T	*I*	*F*
F		*T*
F		*F*

The only remaining situation in which things are as the premise says they are is one in which the premise is false. So if the argument is valid, the premise implies

5. This way of categorising indirect self-falsification is a departure from Swyneshed's taxonomy of self-falsifying sentences. Swyneshed (*Ins.* 4, 79, 97) implies that indirect self-falsification requires that the sentence falsifies something that falsifies it, as in cases where pairs of sentences both say that the other is false (*Ins.* 4). But (6) does not falsify (7), on the contrary. So if we were sticking closely to Swyneshed's taxonomy of direct and indirect self-falsification, we would say instead that (6) directly self-falsifies but differently from how (1) directly self-falsifies. I prefer my taxonomy; the way that (6) self-falsifies seems indirect in the sense that it self-falsifies *via* another sentence.

that the premise is false and it is, therefore, false. It is false even if things are as it says they are.

In general, a sentence (p) self-falsifies if

i) p says that a sentence (k) is false,
ii) k is a conclusion of a valid argument (η), and
iii) p is η's sole premise.

Lower-case Greek letters like 'η' are variables ranging over arguments. If (5) is valid, the premise of (5) satisfies all three of these conditions. If it is not valid then the ii) condition fails for the premise of (5).

One nice question is whether (5) is valid. This is somewhat controversial (hence my leaving the middle column of the above table mostly blank), and we need not settle this question for now. In either case, we can avoid having to accept (3). If (5) is valid, its premise is false because it self-falsifies. If (5) is not valid, it is not a valid argument taking us from truth to falsity.

3.3. A Relevant Wrinkle

There is a serious wrinkle that you may have noticed. By appealing to chains of implication like that which make the premise of (5) a self-falsifier via the argument as a whole, a concern arises that all falsehoods will be characterised as self-falsifying. If the kind of implication in play in the characterisation of self-falsification is material, any false sentence will self-falsify. Take (8) and (9),

(8) (9) is true.
(9) Frank Jackson is the Pope.

Supposing we are talking about material implication, and given the contingent fact that Frank Jackson is not the Pope, the claim that (9) is true implies that Frank Jackson is the Pope which, in turn, implies that (8) is false (along with everything else). The same goes for the rest of the non-paradoxical falsehoods. (8) self-falsifies *if* the kinds of chains of implication that give rise to indirect self-falsification are understood as chains of *material* implication. But plausibly (8) does not self-falsify, it is just plain false.

The Swyneshedian must appeal to some other kind of implication. There will need to be some limitations on what some falsehoods imply, in the relevant sense of 'imply'. The Swyneshedian needs to characterise (6) as self-falsifying without characterising (8) as self-falsifying. Both (6) and (8) are *risky* in Kripke's sense; whether they are paradoxical depends on contingent extrinsic features of the world. But the world does not impose paradoxicality on (8) in the same way it imposes it on (6). It seems that there is some important difference between (6) and (8), but it is a nice question what, exactly, the difference amounts to.

It is natural to appeal to relevance. Indeed Swyneshed's talk about how self-falsifiers are "relevant" to inferring themselves to be false suggests this move. Whether (8) is true or false is, intuitively, not relevant to evaluating (9). (9) does not say anything about (8)'s truth or say anything about the truth value of some other sentence that, down the line of implication, says something about (8)'s truth. On the other hand, whether (6) is true is directly is relevant to whether (7) is true. On this sort of story, self-falsification requires a chain of a certain *relevant* implication relations from a sentence to its own falsehood. The notion of relevant implication could then be spelled out in terms of a modal, hyperintensional counterfactual dependence, or some other way that captures the above judgements about relevance.

Restall (2008) has developed an interesting modal model of implication for Bradwardine's theory of truth that is similar to the kind of account the Swyneshedian will need here. Restall's work (see also Restall [1995]) could shed light on a more complete account of the implication the Swynish theorist needs. Another source of insight here is Herzberger's (1973) work on developing a model for Buridan's theory of truth that distinguishes a sentence's broader truth conditions from the conditions on things corresponding to what the sentence says about how things are. This kind of model might provide the Swyneshedian tools useful in building out their conception of the implication involved in self-falsification.

This is a fascinating issue. Notes promising future work usually go near the end of papers, but I am afraid I will have to issue one here. I hope to sidestep this issue for the purposes of this paper; the sentences characterised as self-falsifiers in the rest of this paper plausibly do not self-falsify by implying something that implies everything. Rather, the chains of implication in play will be based on judgments about what sentences say and what they say it about. The self-falsifying sentences in question will be like (6) and not like (8), however that distinction between the two should be understood on closer inspection.

One more thing to note about what, exactly, the Swynish theorist needs here: they must deny what are sometimes called the 'negative paradox of material implication' (NPMI) and the 'positive paradox of material implication' (PPMI).

$$\textbf{NPMI: } \neg p \rightarrow (p \rightarrow q)$$

$$\textbf{PPMI: } q \rightarrow (p \rightarrow q)$$

Those who reject PPMI on these grounds must supply a non-standard semantics for implication that, among other things, does not rely primarily on the claim that truth preservation is sufficient for implication (Restall 1995).

3.4. False Contradictory Pairs

Moving on to (4), consider (1) and (10).

(1) (1) is false.

(10) (1) is not false.

(1) and (10) seem contradictory. Yet on Swyneshed's proposal, (1) and (10) are both false. If (1) is false, as Swyneshed contends, then (10) is false as well since (10) says that (1) is not false. If (1) and (10) are contradictory, their both being false would be a violation of what is sometimes called the "rule of contradictory pairs" (RCP).

RCP: For any pair of contradictory sentences, one is true and one is false.

For a recent discussion of RCP and how it relates to the rest of Swyneshed's views, see Read (2020b).

This result suggests a non-standard treatment of negation. In the case of (1) and (10), we have a false sentence whose negation is false; so there will need to be another "row" for negation.

$\neg$	p
T	F
F	T
F	F

A way into the bottom row is if p self-falsifies. This is more or less what Swyneshed (*Ins.* 27) admits in his argument for (4). (1) and (10) are both false and p and $\neg p$ are both false whenever p self-falsifies. However, $\neg p$ is not false because it self-falsifies, it is false because things are not as it says they are. This means that $\neg p$ does not, in general, inherit self-falsifingness from p. According to this treatment, the value of negations are not determined without remainder by the value of the negated sentences. However, the truth value of $\neg p$ is determined without remainder by p's semantic features (including whether it self-falsifies). The proposed treatment of negation is incompatible with what is sometimes called "Negation Introduction."

Negation Introduction: $F(p) \models \neg p$

According to the Swynish proposal, there are ways for sentences to be false such that their negation is also false.

The proposed treatment of negation is compatible with some natural inference rules involving negation and falsehood. The principles I'll call "Smash" and "Grab" are compatible with the proposed treatment of negation.

Smash: $\neg p \models F(p)$

Grab: $p \models F(\neg p)$

The idea behind Smash is that if things are as $\neg p$ says they are then, things are not as p says they are (whether or not p self-falsifies) so p is false. The idea behind Grab is that if things are as p says they are, then things are not as $\neg p$ says they are (regardless of whether p self-falsifies). These principles tie negation to falsehood in a natural way and may be of some comfort for those concerned about the loss of Negation Introduction.

The proposed treatment of negation is also compatible with double negation elimination and introduction.

Double Negation Elimination: $\neg\neg p \models p$

Double Negation Introduction: $p \models \neg\neg p$

More broadly, if p self-falsifies, $\neg p$, $\neg\neg p$, $\neg\neg\neg p$, and so on can all be false, alternating between self-falsifying (as in the case of $\neg\neg p$) and things not being as the sentence says they are (as in the case of $\neg p$).

The Swynish approach to negation is compatible with Disjunctive Syllogism.

Disjunctive Syllogism: $(\neg p, p \vee q) \models q$

'$\vee$' indicates disjunction. If $\neg p$ is true, as one of the premises says it is, then p is false (by Smash). If the disjunctive premise is true, it does not self-falsify. But if $p \vee q$ does not self-falsify, $\neg p$ and $p \vee q$ will indeed imply q. Given that $\neg p$ is true, the only way for $p \vee q$ to be true is if things are as q says they are and q does not self-falsify.

Despite being compatible with Bivalence, the Swynish treatment of negation allows for counterexamples to the so-called "Law of Excluded Middle" (LEM).

LEM: $\models (p \vee \neg p)$

Consider (11).

(11) Either (11)'s first disjunct is false or (11)'s first disjunct is not false.

The second disjunct is the negation of the first. The first disjunct self-falsifies and is false. At the same time, things are not as the second disjunct says they are so it is false as well. So (11) is false and we have a counterexample to LEM but there needn't be any propositions assigned a "middle" value. Perhaps the conventional name for this principle is misleading here.

4. Revenge and Self-VOTification

The "revenge" move involves constructing a bespoke liar sentence by attempting to use tools of a proposed solution to the original liar paradox to construct a more stubborn paradoxical sentence (Beall 2007; Shapiro 2011). Take the response to the liar that says that liar sentences like (1) are neither true nor false and have a third truth value, call it Third. Consider (12).

(12) (12) is either false or Third.

Suppose (12) is either true, false, or Third. If it is false, then we can reason in the standard way and we are back where we started. If (12) is Third, it seems like it should also be true (by Capture). If (12) is true, (12) is either false or Third (by Release), so we arrive at the paradoxical conclusion that (12) is both true and either false or Third. So those who make this proposal cannot adequately state their own solution to the liar, since if they could we could construct a liar sentence like (12) that the proposed solution cannot handle.

Is Swyneshed's approach to the liar subject to this kind of revenge challenge? It seems to be. Consider (13).

(13) (13) is not true.

Does (13) self-falsify? Perhaps not. If there is a third truth value, one way for (13) to be not true is for it to have the third value.

I propose that instead of appealing to self-falsification, the Swynish theorist should appeal to self-value-other-than-trueification, "self-VOTification" for short. A sentence self-VOTifies if it implies (directly or indirectly) that it has a value other than true and all sentences that self-VOTify are false. Like self-falsifyingness, self-VOTifyingness is a semantic feature of sentences. All self-falsifiers are self-VOTifiers but not all self-VOTifiers are self-falsifiers. Even if (13) does not self-falsify, it self-VOTifies. Attempted revenge challenges will, it seems, similarly involve sentences that self-VOTify and can, therefore, be handled in a similar way.

Self-VOTifiers and self-falsifiers can interact with negation in different ways. For some self-VOTifiers, their negation is true. For example, take (14) and (15), where "Third" picks out a purported third truth value:

(14) (14) is Third.

(15) (14) is not Third.

(14) self-VOTifies and is, therefore, classified as false on the proposed story. But (15) is plausibly true, since it says that (14) is not Third and (14) is not Third. The lesson is that negating self-VOTifiers sometimes but not always results in a falsehood. Things are also not as (14) says they are ((14) is false and not third) and this means that the proposed relationship between negation and self-VOTification is compatible with Grab.

Next, consider (16).

(16) Things are not as (16) says they are.

I propose that (16) should be characterised as false. It follows from (16) that things are not as it says they are. But things being as a sentence says they are is a requirement on truth, and if things aren't as a sentence says they are, that sentence is false (recall Smash). So it follows from (16) that (16) is false and so (16) self-falsifies.[6]

Broadly, the truth, falsity, and being self-VOTifying predicates can be given their natural interpretation.

(17) (17) self-VOTifies.

(17) implies that (17) implies that (17) has a value other than true and so it self-VOTifies and is false. (17) self-VOTifies, in some sense, via itself. There is, as far as I can see, no threat of revenge that flows from this ability to ascribe the relevant properties. According to the Swynish approach, "false," "true," and "self-VOTifies" can be taken at face value.

5. Curry Paradoxes

5.1. The Standard Curry

The response to the liar paradox discussed so far lights the way to an attractive solution to Curry's paradox (Curry 1942; Löb 1955; Bimbó 2006). One simple

6. This is another departure from Swyneshed's stated views. Swyneshed (*Ins.* 93, 99) characterises sentences like (16) as neither true nor false. I won't get into exegesis here, but Stephen Read has suggested to me that the reason Swyneshed goes this way ties into his interest in insolubles that do not directly concern truth, like 'this does not express a proposition' and 'this is not known'.

expression of Curry's paradox involves a conditional whose antecedent says that the conditional is true.

(18) If (18) is true, then Frank Jackson is the Pope.

For ease of reference I will label the false sentence 'Frank Jackson is the Pope' "*FP*." Next consider the antecedent of (18), which I will call "*A*." If *A* is true, we can infer that the conditional is true (by Release since *A* says that (18) is true). From there we can infer *FP* by modus ponens. But on the material conception of the conditional all conditionals with false antecedents are true. So if *A* is false, (18) is true. But *A* says that (18) is true (and doesn't say anything else), so it seems that if (18) is true, so is *A*. Then we can again infer *FP* by modus ponens. This is paradoxical since *FP* is false and we could repeat the process to infer any false sentence from a corresponding sentence with the same structure.[7]

I propose that the antecedent of (18) self-falsifies and is, therefore, false even if (18) is true. We can then block the step in the paradoxical reasoning that takes us from the conditional to the truth of its antecedent. The antecedent implies that the conditional is true, but since *FP* is false, (18) implies that *A* is false. So, chaining those together, *A* implies that (18) is true but (18) implies that *A* is false, so *A* self-falsifies. Then even if (18) is true, and *A* says that (18) is true, *A* is still false, since it self-falsifies. So we do not get to infer *FP* and the same goes for all Curry sentences with false consequents.

What about Curry sentences with true consequents like (19)?

(19) If (19) is true, then Frank Jackson wrote *Epiphenomenal Qualia*.

There is a choice point here for a defender of the Swynish approach; should sentences like (19) be characterised as true?

Some suggest that it is paradoxical enough to be able to derive a contingent truth from a Curry sentence (Zardini 2021; Oms 2023). The worry is that, on many treatments of conditionals, contingent truths (like the consequent of (19)) can be derived a priori from the logic of truth and Curry sentences and that this is problematic even if we are unable to derive falsehoods from Curry sentences. Call this the "positive Curry paradox."

7. I have presented the paradox this way because it makes salient the part of the reasoning (sometimes contained in a broader appeal to the rules of conditional proof) that takes us from the Curry conditional to the truth of its antecedent, but nothing much hangs on this presentation. However one presents Curry's paradox, there will be a step that involves inferring the truth of the antecedent from the conditional. It is that step that the Swynish story allows us to block.

It might appear as if the Swynish approach does not tell us much about how to handle the positive Curry paradox. However, the Swynish approach opens up some dialectical moves here. Consider (20).

(20) If (20) is true, then there are an even number of stars.

For the Swynish theorist, whether (20)'s antecedent self-falsifies depends on how many stars there are. As discussed above in §3.2, the Swynish theorist should agree with Kripke's claim that whether a given sentence is paradoxical can depend on contingent features of the world, extrinsic to the sentence itself. When the antecedent says that the conditional is true, the Swynish theorist can block the inference from the conditional to the antecedent by restricting Capture; even if the conditional is true, we cannot infer that the antecedent is true unless the antecedent does not self-falsify. In turn, whether an antecedent of a Curry sentence self-falsifies depends on the truth value of the consequent (as outlined above). What this means is that if we do not know whether the consequent is true, we do not know whether the Curry antecedent self-falsifies.

To derive that there are an even number of stars from (20) (in the way that is supposed to be problematic), we need both the conditional and its antecedent as premises. By restricting Capture and making the truth of the antecedent depend on the truth value of the consequent, we have the antecedent as a premise only if the consequent is not false; if the consequent were false, then the antecedent would self-VOTify. If we do not know whether there are an even number of stars, we do not know whether the antecedent premise is true. But in that case we are not able to infer a contingent truth *a priori*, since the inference requires a premise whose truth depends on the world being a certain contingent way. According to the Swynish approach, when the consequent is contingent, the Curry antecedent is, in a sense, contingent as well. This may take some of the sting out of the positive Curry paradox. It might make the claim that we can derive contingent truths from Curry sentences more palatable since we can qualify that we can only do so under the supposition that the contingent consequent is true. On this sort of story, these derivations are a posteriori, at least in the way just described. Perhaps, then, it is not so problematic that sentences like (19) are characterised as true.

There may be a lingering worry here. It might be concerning enough that a contingent truth can be derived through Curry sentences, even in cases where the truth of the consequent is established as true. I am not sure how to address that concern and having to stare down those with this concern may be a mark against a Swynish approach that characterises Curry sentences with true consequents as true.

Doing full justice to the question of how to treat the positive Curry paradox must be left for future work. I will proceed as if sentences like (19) are true.

Next, should a defender of the proposed Swynish approach characterise Curry conditionals with false consequents as false? There are two salient options.

One option, ''Option 1,'' is to characterise (18) and like sentences as true. A number of consequences flow from Option 1, not least concerning of which is that it allows for counterexamples to Modus Tollens (and parallel rules concerning contraposition).

Modus Tollens: $(\neg q, p \to q) \models \neg p$

Take (21).

(21) If the antecedent of (21) is false, Frank Jackson is the Pope.

The negation of (21)'s antecedent, (22), is false.

(22) The antecedent of (21) is not false.

But, given Option 1, (21) is true and so is the negation of its consequent. So we have a counterexample to Modus Tollens.

Another option, "Option 2," is to characterise (18) and similar sentences as false. This would allow us to avoid that kind of counterexample to Modus Tollens. But it does have its own complexities. To begin with, it would imply that some false conditionals have false consequents. As discussed in §3.2, the Swynish theorist will have to make this claim in any case, so even if this is a cost, it is not an extra cost.

The idea behind Option 2 is that, just as Swyneshed proposes an extra way for sentences to be false other than things not being as it says they are, there is another way for a conditional to be false other than its having a true antecedent and a false consequent and (18) is an example of a sentence that is false in that way. If the antecedent self-falsifies and the consequent is false, then the conditional is also false.

The Swynish conditional understood in line with Option 2 is not truth-functional in the sense that the truth values of the antecedent and consequent do not determine without remainder the value of the conditional (Edgington 1995; 2020).

I am inclined to take Option 2; I am keen to maintain Modus Tollens and, for the Swynish theorist, the truth-functionality of conditionals will have to be sacrificed in any case. Another advantage is that if one takes Option 2 *and* claims that Curry sentences with true consequents are true, we can say that Curry conditionals simply take the truth value of their consequents. Your mileage on these advantages may differ, but I will assume Option 2 in what follows.

5.2. The Validity Curry

Next consider a different version (or "flavor") of Curry's paradox often called the ''validity Curry" (Beall & Murzi 2013). As an illustration, take (23).

(23) (23) is valid, therefore Frank Jackson is the Pope.

The premise (*PM*) is true only if the argument is valid. This leads, via the familiar Curry reasoning, to our apparently being able to derive *FP*. If *PM* is true, we can infer that the argument is valid (by Release), then we can infer *FP*. But if *PM* is false, it looks like we can infer that (23) is valid. But then we can, given unrestricted Capture, infer *PM* since it says that (23) is valid, so we can again infer *FP*.

I propose that under the supposition that *FP* is false, *PM* self-falsifies. Thus even if (23) is valid (as *PM* says it is), *PM* is false. We are thus unable to infer *FP* (and we are unable to infer any other false sentences in the way described above).

Given that (23) is valid and *FP* is false, the only way for the sole premise of (23) to be true, is for it to be false. Three things are going on in the case of (23): the conclusion of (23) is false, the premise of (23) says that (23) is valid, and the premise of (23) is true only if all the premises of (23) are true. These together imply that the only way for *PM* to be true is for it to be false. *PM* is true only if all the premises of (23) are true (*PM* is (23)'s sole premise). So for *PM* to be true, (23) must be valid (as *PM* says it is) *and* (23)'s sole premise must be true. But if *FP* is false, then the only way that all the premises of (23) can be true, is if (23) is invalid.

The table below represents the possibilities.

PM	$\models$	*FP*
T		*T*
T	I	*F*
F		*T*
F		*F*

When we suppose that *FP* is false, we rule out the grey rows. The only remaining situation in which things are as *PM* says they are is one in which *PM* is false. On that row, at least one of (23)'s premises is false. In that case *PM* is false, since it is true only if all (23)'s premises are true. If *FP* is false, for *PM* to be true, *PM* must be false, so *PM* self-falsifies.

In summary, even if (23) is valid and its only premise says it is valid, that premise is still false if *FP* is false.[8]

8. If the Swyneshed-inspired approach to truth I defend here helps with the validity Curry, then Ripley's (2014) suggestion that truth has an alibi for the paradoxes of self-reference (French 2016) looks increasingly shaky.

In general, a sentence (p) self-falsifies if

iv) p implies that a given argument (η) is valid,
v) a conclusion of that argument (k) is false, and
vi) p is η's sole premise.

In the case of (23), iv) and vi) go together. To see how vi) can come apart from iv), consider (24) and (25).

(24) Grass is green and the moon is bigger than the sun, therefore the moon is bigger than the sun.

(25) (24) is valid.

(25) is true. The conclusion of (24) is false. But (25) does not self-falsify, since (25) is not (24)'s sole premise.

6. Swynish Truth

Let's pause to take stock. I have proposed a Swyneshed-inspired solution to liar and Curry paradoxes based on some claims about truth and the logic of truth. This solution has implications for how we ought to understand the semantics of connectives, most obviously negation and the conditional.

The central proposal about truth is that the truth of a given sentence requires that things are as it signifies that they are and that it does not VOTify itself. Accordingly, I propose replacing capture with what I call "VOT-restricted Capture."

VOT-restricted Capture: $(p \wedge \neg SV(p)) \models T(p)$

'SV(x)' is the 'self-VOTifies' predicate. Release is unchanged.

This revised approach to truth seems quite natural. For a sentence to be true, things must be as the sentence says they are. If a sentence self-VOTifies, for things to be as it says they are, they have to fail to be, at least entirely, as it says they are. Self-VOTifying sentences imply that they have a value other than true and we are assuming that the values other than true exclude truth. By making its truth require its having a value other than true, what (1) says makes it impossible, in a strong sense, for it to be true. When a sentence self-VOTifies it seems to sabotage its ability to be true. If this is right, it is natural to characterise sentences that self-VOTify as false. This obviously constitutes a departure from so-called "naive" theories of truth characterised by the unrestricted T-schema. The defender of the Swynish proposal will simply have to appeal to other advantages that might counterbalance any costs that arise from rejecting a simpler theory of truth.

7. Grain

One challenge that faces any approach to paradox according to which liar sentences are false is, ironically, the threat of the theory itself self-falsifying. Consider (1) and (26).

(1) (1) is false.

(26) (1) is false.

(26) seems to be a central part of the proposed approach to paradox. The worry is that (26) and (1) say the same thing about the same thing so, on some accounts of semantic features and the individuation of sentences, they should have the same truth value. So either the centerpiece of the proposed solution (the claim that liar sentences are false) is false or the liar is true. Neither possibility is comfortable for the Swynish theorist.

This challenge brings out that a defender of a Swynish approach should adopt an account of semantic features according to which (26) and (1) come out as having different semantic features. Depending on how appealing one finds coarse-grained accounts of semantic features, this might not be perceived as much of a cost of the approach.

I am fairly confident in the viability of sufficiently fine-grained conceptions of semantic features. Consider (27) and (28).

(27) Hesperus is a planet.

(28) Phosphorus is a planet.

These seem to have different semantic features, and it seems possible to affirm one without affirming the other and yet, intuitively, they say the same thing about the same thing.

Similarly for (29) and (30).

(29) The second sentence of *On Denoting* is false.

(30) "Thus a phrase is denoting solely in virtue of its form" is false.

Again, we have a pair of sentences saying the same thing about the same thing but which appear to have distinct semantic features.

If Deirdre says "I am happy" and I say "Deirdre is happy," it seems plausible that we have uttered different sentences (or at least it is plausible that our utterances have different semantic features). In this vein, we might contend that truth conditions might be inherently perspectival (Perry 1979; Lewis 1979) or tied to the particular thinker, sentence, or speaker.

On the Swynish account, (1) and (26) have different semantic features. (1) implies that it, (1), is false but (26) does not entail that it, (26), is false. One might ask, what feature of these sentences gives rise to this difference? I think how we answer this question will turn on how truth conditions, indexicality, names, meaning, and context tie together. I will not attempt to dive into that debate here, but suffice to say that the Swynish theorist is on the hook here. They must carve semantic features at least this finely and owe a well-motivated story about why this carving is appropriate.

Some extant contextualist responses to semantic paradoxes will allow for the needed distinctions. In broad terms, according to the contextualist, the semantic status of sentences like the liar vary across contexts because there is some important role for context in the proper ascription of truth or semantic defectiveness to sentences (Parsons 1974; Burge 1979; Koons 1992; Glanzberg 2001; 2005). On this sort of story, one might argue that (26) and (1) are to be evaluated in different contexts and that this shift is what explains how one can be true and the other false. I don't want to commit to this sort of view at this stage, but appealing to context here is one option available to Swynish theorists for drawing the distinctions they need.

8. Compound Sentences and Self-VOTification

Just as there are simple liar sentences like (1), there are liar-like compound sentences that exhibit similar semantic features that make them paradoxical. As Swyneshed (*Ins.* 60) suggests, we can construct self-falsifying conjunctions.

(31) Frank Jackson wrote *Epiphenomenal Qualia* and (31) is false.

(31) self-falsifies since it implies its second conjunct which, in turn, implies that it, (31), is false. This implication does not depend on the truth value of the other conjunct.

These self-VOTifying conjunctions should, essentially, be treated along the lines of the Swynish treatment of simple liar sentences. Just as the Swynish theorist suggests that Capture fails in the case of the liar, things being as all the conjuncts say they are (as in the case of (31)) does not entail the conjunction is true. Some conjuncts, like the first conjunct of (31), make conjunctions liar-like and render the conjunction self-VOTifying. Even though, in some crucial sense, things correspond to how (31) says they are, (31) is false because it fails to satisfy the other condition on truth that is not obviously reducible to correspondence.

As Swyneshed (*Ins.* 62) points out, a disjunction can self-falsify.

(32) Either Frank Jackson is the Pope or (32) is false.

Given that the first disjunct is false without self-VOTifying, the disjunction implies its second disjunct (Disjunctive Syllogism) and, in turn, the second disjunct implies that the disjunction is false and so (32) self-falsifies. Neither disjunct self-VOTifies but the disjunction does. As in the case of conjunction, (32) is false even if, in some intuitive sense, things are as it says they are.

Similarly there are liar-like conditionals such as (33).

(33) If Frank Jackson wrote *Epiphenomenal Qualia* then (33) is false.

Given that its antecedent is true, (33) implies its consequent. In turn its consequent implies that it, (33), is false. So (33) self-VOTifies. The consequent of (33) appears to be true; things are as it says they are and it seems not to self-VOTify. Some sentences are such that their being the consequent of a conditional renders the conditional a self-VOTifier, even if that consequent is true.

9. A Nearby Road

Some posit an extra truth value for reasons not stemming from the liar or Curry paradoxes, for instance as part of a treatment of vagueness or presupposition failure. Much of what I have proposed here is compatible with there being a third truth value. But if there are more than two truth values, one might reasonably ask, why not claim that self-VOTfifying sentences have the third truth value? Perhaps that would allow us to retain more of the standard account of negation and implication within a broadly Swynish approach. The idea would be to characterise self-VOTifying sentences as having a third truth value, again let's call it "Third."

One important choice point for defenders of this approach; should they classify the negations of Third sentences as true, false, or Third? It seems unnatural to characterise them as false. According to the idea in question, things are as (10) says they are (since (1) is indeed not false).

Characterising the negations of Third sentences as true leads to contradiction. Consider (34) and (35).

(34) (34) is either false or Third.

(35) (34) is neither false nor Third.

(34) self-VOTifies so, on this proposal, it would be Third. (35) is the negation of (34). So under the supposition that negating Third sentences always results in a truth, (35) is true. But if (35) is true, we can infer (by Release) that (34) is neither false nor Third, which contradicts the characterisation of (34) as Third.

The option to characterise negations of Third sentences as uniformly Third is perhaps the most promising. However, (10) appears not to self-VOTify, either directly or indirectly. On this three-valued story, it seems like (10) should be classified as true, since things are as it says they are, (1) is indeed not false, it is Third. Yet taking this option would force the claim that (10) is Third. Those who take this option must argue that, appearances notwithstanding, sentences like (10) are not true (as the three-valued story prima facie implies) and are not false (as the bivalent version of the Swynish proposal would have it), but are instead Third.

Given the background of a Swynish approach, it seems unnatural to characterise sentences like (10) as anything other than false. On the other hand, truth-functionality for negation is desirable, so some defenders of a broadly Swynish approach may wish to sacrifice what I take to be plausible claims about sentences like (10) at the alter of truth-functionality.

Those who assign a third truth value to self-VOTifying sentences would also have to develop a corresponding treatment of the conditional, conjunction and disjunction. The literature on three-or-more-valued logic is rich and there are many giants upon whose shoulders they might try to stand.

10. What Next?

I have developed a Swyneshed-inspired solution to liar and Curry paradoxes, along with some of the associated implications for the corresponding logic, theory of truth, and semantics for connectives. Where to from here? There is work to be done to refine the proposal and explore more of its consequences; particularly as regards the relationship between validity, truth, and falsity and how to understand the corresponding logic of connectives. As flagged in §3.3, the Swynish theorist will also need a theory of implication that renders the distinction between self-VOTifing falsehoods and non-self-VOTifing falsehoods non-trivial. Another interesting avenue for further work is a discussion of how the proposed solution to the liar and Curry paradoxes relates to the truth-teller and positive Curry paradoxes.

Acknowledgments

Thanks to Ben Blumson, Szymon Bogacz, Clare Due, Paul Egre, Rohan French, Alan Hájek, Edwin Mares, Daniel Nolan, Graham Priest, Stephen Read, Ellie Ripley, Shawn Standefer, Koji Tanaka, Sara Uckelman, an audience at the 2023 Australian Association of Philosophy conference, an audience at the Australian National University, and to two anonymous reviewers for Ergo for feedback on the material in this paper.

References

Beall, J. C. (2007). *Revenge of the Liar: New Essays on the Paradox.* Oxford University Press.
Beall, J. C. and Murzi, Julien (2013). Two Flavors of Curry's Paradox. *Journal of Philosophy*, *110*(3), 143–165.
Bimbó, Katalin (2006). Curry-Type Paradoxes. *Logique et Analyse*, *49*(195), 227–240.
Burge, Tyler (1979). Semantical Paradox. *The Journal of Philosophy*, *76*(4), 169–198.
Curry, Haskell B. (1942). The Inconsistency of Certain Formal Logic. *Journal of Symbolic Logic*, *7*(3), 115–117.
Edgington, Dorothy (1995). On Conditionals. *Mind*, *104*(414), 235–329.
Edgington, Dorothy (2020). Indicative Conditionals. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2020 ed. Metaphysics Research Lab, Stanford University.
French, Rohan (2016). Structural Reflexivity and the Paradoxes of Self-Reference. *Ergo, an Open Access Journal of Philosophy*, *3*.
Glanzberg, Michael (2001). The Liar in Context. *Philosophical Studies*, *103*(3), 217–251.
Glanzberg, Michael (2005). Truth, Reflection, and Hierarchies. *Synthese*, *142*(3), 289–315.
Herzberger, Hans G. (1973). Dimensions of Truth. *Journal of Philosophical Logic*, *2*(4), 535–556.
Koons, Robert C. (1992). *Paradoxes of Belief and Strategic Rationality.* Cambridge University Press.
Kripke, Saul (1975). Outline of a Theory of Truth. *Journal of Philosophy*, *72*(19), 690–716.
Lewis, David K. (1979). Attitudes De Dicto and De Se. *The Philosophical Review*, *88*(4), 513–543.
Löb, Martin H. (1955). Solution of a Problem of Leon Henkin. *Journal of Symbolic Logic*, *20*(2), 115–118.
Oms, Sergi (2023). Some Remarks on the Notion of Paradox. *Acta Analytica*, *38*(2), 211–228.
Parsons, Charles (1974). The Liar Paradox. *Journal of Philosophical Logic*, *3*(4), 381–412.
Perry, John (1979). The Problem of the Essential Indexical. *Noûs*, *13*(1), 3–21.
Petersen, Uwe (2000). Logic Without Contraction as Based on Inclusion and Unrestricted Abstraction. *Studia Logica*, *64*(3), 365–403.
Read, Stephen (2002). The Liar Paradox From John Buridan Back to Thomas Bradwardine. *Vivarium*, *40*(2), 189–218.
Read, Stephen (2020a). The Rule of Contradictory Pairs, Insolubles and Validity. *Vivarium*, *58*(4), 275–304.
Read, Stephen (2020b). Swyneshed, Aristotle and the Rule of Contradictory Pairs. *Logica Universalis*, *14*(1), 27–50.

Restall, Greg (1995). Four-Valued Semantics for Relevant Logics (and Some of Their Rivals). *Journal of Philosophical Logic*, *24*(2), 139–160.
Restall, Greg (2008). Modal Models for Bradwardine's Theory of Truth. *Review of Symbolic Logic*, *1*(2), 225–240.
Ripley, David (2013). Paradoxes and Failures of Cut. *Australasian Journal of Philosophy*, *91*(1), 139–164.
Ripley, David (2014). Review of Kevin Scharp, Replacing Truth. *Notre Dame Philosophical Reviews*.
Ripley, David (2015). Comparing Substructural Theories of Truth. *Ergo: An Open Access Journal of Philosophy*, *2*.
Shapiro, Lionel (2011). Expressibility and the Liar's Revenge. *Australasian Journal of Philosophy*, *89*(2), 297–314.
Skyrms, Brian (1970). Notes on Quantification and Self-Reference. In Robert L. Martin (Ed.), *The Paradox of the Liar* (67–74). Yale University Press.
Spade, Paul V. (1979). Roger Swyneshed's "Insolubilia": Edition and Comments. *Archives d'histoire doctrinale et littéraire du Moyen Âge*, 177–220.
Spade, Paul V. (1983). Roger Swyneshed's Theory of Insolubilia: A Study of Some of His Preliminary Semantic Notions. *History of Semiotics* (Foundations of Semiotics, vol. 7, 105–113). John Benjamins.
Swyneshed, Roger (1330-1335). *Insolubilia* (S. Read, Trans). The Logic Museum. https://www.logicmuseum.com/wiki/Authors/Roger_Swyneshed/Insolubilia.
Tarski, Alfred (1936). Der Wahrheitsbegriff in den Formalisierten Sprachen. *Studia Philosophica*, *1*, 261–405.
Weir, Alan (2015). A Robust Non-Transitive Logic. *Topoi*, *34*(1), 1–9.
Zardini, Elia (2011). Truth Without Contra(di)ction. *The Review of Symbolic Logic*, *4*, 498–535.
Zardini, Elia (2021). Substructural Approaches to Paradox: An Introduction to the Special Issue. *Synthese*, *199*(3), 493–525.