TRUTH-TELLERS IN BRADWARDINE'S THEORY OF TRUTH

Greg Restall*

Philosophy Department
The University of Melbourne
restall@unimelb.edu.au

september 29, 2008
Version 0.99

Abstract: Stephen Read's work on Bradwardine's theory of truth is some of the most exciting work on truth and insolubilia in recent years [5, 6]. Read brings together modern tools of formal logic and Bradwardine's theory of signification to show that medieval distinctions can give great insight into the behaviour of semantic concepts such as truth. In a number of papers, I have developed a model theory for Bradwardine's account of truth [7, 8]. This model theory has distinctive features: it serves up models in which every declarative object (any object signifying *anything*) signifies its own truth. This leads to a puzzle: there is a good argument to the effect that if anything signifies its own truth—if anything is a '*truth-teller*'—it is *false* [6], and that this feature of a theory of truth—that not every declarative need signify its own truth—is what distinguishes Bradwardine's account from Buridan's [5, 6]. What are we to make of this? If the argument fails, what distinguishes *problematic* truth-tellers (such as a sentence that explicitly says of itself that it is true) from *benign* truth tellers? It is my task in this paper to explain this distinction, and to clarify the behaviour of truth-tellers, given my the contemporary formal treatment of Bradwardine's account of signification.

1 bradwardine's theory

Bradwardine's theory of truth, as set out by Stephen Read in a series of papers [5, 6], defines truth in terms of a prior notion of signification. This notion

*Thanks to Catarina Dutilh Novaes and to the audience at the 1st gpmr Conference in Medieval and Applied Logic for feedback on some of the ideas in this paper, and especially to Stephen Read for introducing me to Bradwardine's theory of truth, and for confronting me with the puzzle in this paper. The comments of anonymous referees also helped me clarify my thinking and expression of this material. Any remaining errors and infelicities are, of course, mine. This research is supported by the Australian Research Council, through grant dp0343388, and the Penguin Café Orchestra.

is expressed grammatically by neither a *predicate* nor an *operator*. The concept is expressed in claims of the form 'x says that p' or 'x signifies that p'. The phrase 'says that' is syntactically and semantically a hybrid. To form a sentence, we can substitute a name or referring expression in place of 'x' and another sentence in place of 'p.' Expressed by a 'connecticate' uniting a singular term with a sentence, signification is the core notion from which truth is defined. I will write the connecticate of signification with the simple infix colon ":". Whenever t is a singular term and p is a sentence,

$$t : p$$

is another sentence, to be read 't signifies that p', or simply just that 't says that p.' According to Read's analysis, Bradwardine uses signification to define the truth predicate: t is true if and only if it signifies something, and everything it signifies is the case. It is false if it signifies something that is not the case. So truth and falsity are defined notions, where the definitions utilise signification and what we have come to call propositional quantification.

At this point we must clarify a distinction between Bradwardine's terminology and modern vocabulary: quoting Read [5, page 191] we have

> DEFINITION 1: *A true proposition is an utterance signifying only as things are.*
>
> DEFINITION 2: *A false proposition is an utterance signifying other than things are.*

For Read and for Bradwardine, what is called a 'proposition' is what is denoted by the referring expression in the first part of the claim of the form '$x : p$.' The proposition is the utterance or sentence or other *object* of which we are predicating truth. To confuse matters, we now use 'proposition' to describe quantification into the *second* position, in which the sentence is mentioned but not used. To keep matters as clear and precise as I can, I will describe the objects of which we will predicate truth and falsity *sentences* (thinking primarily of tokens, and not types, though we may sometimes predicate truth of types if the sentences do not vary too much in significance from context to context), or *utterances* or other such things. I will avoid talking about propositions as much as possible in what follows, except when talking of what is now called propositional quantification. Using this vocabulary, then, DEFINITIONS 1 and 2 become:

- Ts is defined as $(\exists p)(s : p)$ & $(\forall p)(s : p \to p)$
- Fs is defined as $(\exists p)(s : p\ \&\ \neg p)$

The object s is true when it signifies something, and all it signifies is the case, and s is false, when it signifies something that is not the case. (In the formal syntax ':' binds more tightly than '&' so '$(\exists p)(s : p\ \&\ \neg p)$' is not to be read as '$(\exists p)s : (p\ \&\ \neg p)$'.) Note that these formal definitions do not utilise any notion of being "the case"—there is no separate notion of truth playing a role in Bradwardine's own definitions of the notion of truth, except for any notion of truth implicit in treating these marks as making assertions. When rendered into English prose, to be sure, we use predications such as 'is the case' or 'holds' when reading propositionally quantified expressions, but this need not mean that we must use a second notion of truth to define the first.[1]

According to Read, Bradwardine's account of truth and signification relies on what we call *Bradwardine's axiom*:

DEFINITION [BRADWARDINE'S AXIOM] *Every proposition signifies or means contingently or necessarily everything which follows from it contingently or necessarily* [5, p. 209].

We could formalise the condition like this:

$$\text{If } t : p \text{ then if (if } p \text{ then } q) \text{ then } t : q$$

Rearranging, we might have another formulation

$$\text{If (if } p \text{ then } q) \text{ then (if } t : p \text{ then } t : q).$$

The crucial issue in understanding Bradwardine's axiom is what form of conditional expression might be used in formulating it. We can consider *material* conditionals

$$\text{If } (p \supset q) \text{ then } (t : p \ \supset \ t : q)$$

where we can infer $\ulcorner p \supset q \urcorner$ either from $\ulcorner \neg p \urcorner$ or from $\ulcorner q \urcorner$. These conditionals are very weak, and it is most likely that they are not the right way to get to the core of Bradwardine's notion of truth. After all, we have

$$\text{If } \neg p \text{ then } (t : p \ \supset \ t : q). \qquad \text{If } q \text{ then } (t : p \ \supset \ t : q).$$

which together almost trivialise signification. First, if $t : p$ and $\neg p$, then $t : q$—so if t signifies something (namely $\ulcorner p \urcorner$) such that $\neg p$, then t signifies everything. So anything signifying something not the case signifies everything and anything. Second, if q, then if t signifies that p (if t signifies *something*) then t signifies that q. That is, anything that signifies something signifies everything that is the case. We therefore would have three options: (1) an object can signify nothing, or (2) if it signifies something, it signifies everything that is the case—and (3) if it also signifies something that is *not* the case, it signifies everything. An objects status with regard to signification can be reduced to two pieces of information. Step 1: Is it *declarative*?—does it signify anything? If

[1] Whether this is a mark against propositional quantification is a bone of contention. I think that a coherent story—which neither takes a notion of propositional truth as primitive, nor confuses use and mention of propositional expressions—is possible, but it is not my place to argue this here [1, 2].

not, it signifies nothing. If so, then it signifies at least everything that is the case. Step 2: Does it signify anything that is not the case? If so, it is *false*, and it signifies *everything*. If not, it is *true*. So, everything is either *nondeclarative*, or if it is declarative, it is either *true* or *false*. If Bradwardine's axiom is read in a material fashion, signification collapses into three cases: nondeclarativeness, truth and falsity.

So, if we want a richer analysis of signification, we must explore different options. One different option is to consider *modal* conditionals:

$$\text{If } (p \Rightarrow q) \text{ then } (t : p \ \Rightarrow \ t : q)$$

where $\ulcorner p \Rightarrow q \urcorner$ means that $\ulcorner p \urcorner$ *entails* $\ulcorner q \urcorner$. This kind of reading is considered in "Modal Models for Bradwardine's Theory of Truth" [8]. I showed there that you could *model* Bradwardine's theory in a 'possible worlds' model by representing taking signification as another modal operator—we treat $\ulcorner t : p \urcorner$ in the same vein as $\ulcorner \Box p \urcorner$, where the necessity is relatvised to the object doing the signifying. In this way, if $\ulcorner p \urcorner$ entails $\ulcorner q \urcorner$, then anything signifying $\ulcorner p \urcorner$ signifies $\ulcorner q \urcorner$. If $\ulcorner p \urcorner$ is t-*necessary*, then $\ulcorner q \urcorner$ is also t-necessary.

The details of the model theory are not important, for the puzzle I will discuss here is a problem even in the simpler non-modal case, and its diagnosis is independent of modal matters. For more detail on the modal reading of Bradwardine's axiom, see my earlier papers [7, 8].

The crucial feature of Bradwardine's theory of truth is found in its analysis of the insolubilia: a liar statement, for example, is something that signifies its own untruth. It is an object l such that $l : \neg Tl$. What can we say about l?

Bradwardine's reasoning concerning liars is straightforward. We may start by asking: is l true? That is, do we have Tl? Tl is, by definition, $(\exists p)(l : p)$ & $(\forall p)(l : p \to p)$. For the first conjunct, $(\exists p)(l : p)$, we can point to $l : \neg Tl$. Here is something that l signifies, by assumption. What about the second conjunct? Is it the case that $(\forall p)(l : p \to p)$? Well, if this is he case, then in particular since $l : \neg Tl$, we could conclude $\neg Tl$. So, we have shown that if Tl, it follows that $\neg Tl$. By *reductio ad absurdum*, we may conclude $\neg Tl$.

So, we have proved that a liar statement is not true. Does it follow that it is also *true*? The fact that we find the reasoning to Tl compelling is what makes liar paradox a genuine *paradox*. We are tempted to reason as follows: we have proved that the l is not true. This is *just what the liar says*, so it must be true. This last move is the step that Bradwardine resists, according to Read. Yes, *something* that the liar says is not the case. It does not follow that l is true, for this would require concluding that *everything* that the liar says is the case: Tl entails $(\forall p)(l : p \to p)$. Can we prove this? No, we cannot. Only a further assumption, such that the liar signifies exactly one thing, or that everything that the liar signifies is entailed by $\neg Tl$, or some other such conclusion would help us derive the claim that the liar is also *true*. In the absence of this assumption, we may not conclude that the liar is true. It is simply false. It must signify *something* that is not the case.

What does the liar signify that is not the case? According to Read, Bradwardine's response is that the liar signified its own truth: that $\mathsf{l} : \mathsf{T}\mathsf{l}$. As far as I can see, this does *not* follow from his explicit assumptions, strictly speaking. I have argued elsewhere [7] that, on Bradwardine's account—as understood by Read—some liar statements might signify their own truth, others might not. What they have in common is that they signify their own *untruth*, and because of that, each liar statement signifies *something* that is not the case, and hence, are untrue. However, that 'something else' can vary from one liar statement to another.

Nonetheless, it is quite possible to construct models of this formalisation of Bradwardine's theory in which not only do *insolubilia* such as liar statements say of themselves that they are true, but in which *all* declarative objects declare their own truth. The construction of the "Modal Models" paper [8] delivers such models. The fact that Bradwardine's theory has models in which *every* declarative object declares its own truth gives rise to the puzzle to be considered in this paper.

2 THE PUZZLE OF TRUTH-TELLERS

Consider the following statement:

$$2 + 2 = 4 \text{ and } (\mathsf{r}) \text{ is true} \qquad (\mathsf{r})$$

Taking the signification of this sentence at face value, we may conclude that $\mathsf{r} : (2 + 2 = 4 \ \& \ \mathsf{T}\mathsf{r})$, and hence, since (r) signifies anything that follows from what (r) signifies, we have

$$\mathsf{r} : 2 + 2 = 4$$
$$\mathsf{r} : \mathsf{T}\mathsf{r}$$

The statement (r) signifies that $2+2 = 4$ and it also signifies its own truth. Now, what can we say about (r)? Is it true, or is it false? It seems that the theory we have seen so far does not tell us one way or another. When we inquire after (r)'s truth, we are forced to ask: is $2 + 2 = 4$? Hopefully our arithmetical knowledge extends that far, and we can answer in the affirmative. Now we ask: is (r) true? To answer *this*, we are left back where we started. It seems that it could be true — at least, the assumption that (r) is true does not conflict with anything that we have seen so far. If (r) is true, then everything that (r) says is the case. (r) says that $2 + 2 = 4$, and that (r) is true. This is consistent and coherent. But then, it seems that (r) could equally well be false — at least, the assumption that (r) is false does not conflict with anything that we have seen so far. If (r) is false, then something that (r) says is not the case. (r) says that $2 + 2 = 4$, and that (r) is true. This is consistent and coherent — that (r) is true is something that (r) says that is not the case.

So, for all we know, the truth of (r), and its falsity are both possible. Do we have any reason to prefer one over another? Stephen Read thinks that we

do [6]. Read takes (r) to be false, since there is nothing to make it be *true*. The cycle of checking continues for ever, and there is no way to *ground* the truth of (r) in anything outside it. So, it is *false*.
Now consider another sentence

$$2 + 2 = 4 \tag{s}$$

The sentence (s) says that $2 + 2 = 4$. Does the sentence (s) also signify its own truth? In my model theory for Bradwardine's account of truth, *everything* declarative says of itself that it is true. Could this be the case? If it were the case, for (s) we would have

$$s : 2 + 2 = 4$$
$$s : Ts$$

This looks familiar. We have said the same sort thing concerning what (s) says that we have said concerning what (r) says. Does not the same reasoning apply? Should we not *also* conclude that (s) is false, for in the process to evaluate the truth of (s) we can ask: does $2 + 2 = 4$? Yes. Is (s) true? Well, for that we need to check everything that (s) says. What does it say? Well, it says that $2 + 2 = 4$ (which is the case), and it says that (s) is true. Well, is that the case? Is (s) true? . . . The process never stops.

It seems that the reasoning, if it is good in the case of (r), is also good in the case of (s). But (s) was nothing special. Any true claim would have done just as well. In my models for Bradwardine's theory, every declarative object is a truth teller, and it seems as if—if Read is right—we should say that they are all false.

This is the puzzle: on the one hand, Read's argument in 'Symmetry and Paradox' [6] explains why truth-tellers should be taken to be *false*. On the other hand, there is a model construction that gives us interpretations in which every declarative tells its own truth. Is this view of declaratives consistent with Bradwardine's position? It would seem to follow from these premises that declaratives are all false. But that is absurd, since there are *some* truths. So, some assumption must go: either we are to reject the thesis that non-paradoxical declaratives can declare their own truth (and hence, reject my model construction), or we are to reject Read's reasoning, or finally, we are to reject our formalisation of Bradwardine's theory of truth. Which of these is to be rejected?[2]

3 WHAT GOES ON IN THE MODEL?

In this section I will examine the model construction in a little more detail, so we can understand how this construction treats sentences like (s) and those like (r) differently. In this way, we will be able to use that construction to provide some insight into the different ways that sentences can act as truth-tellers. This

[2]Thanks to Stephen Read for presenting this puzzle over a long lunch in Bonn after the close of the GPMR conference.

will require attention to a little more technical detail, in order to clearly see the difference between these different truth-tellers. I will attempt to involve *just* the amount of technicality required to make the philosophical point, and no more. For the full details of the model construction, I refer you elsewhere [8]. Then in the final section of the paper I will attempt to apply what we have learned from that construction, and to point to further philosophical questions concerning signification and Bradwardine's account of truth.

To make the reasoning precise, I will fix on particular examples of the sentences with the features of our (r) and (s) from the previous section. The model construction takes any formal theory with a possible worlds model and propositional quantification, and adds the logic of signification in such a way as to satisfy Bradwardine's axiom. It *also* ensures that, if the theory in question provides a means of quotation, then the object $\ulcorner A \urcorner$ says that A. The condition that $\ulcorner A \urcorner : A$ ensures that the truth-elimination axiom holds

$$T\ulcorner A \urcorner \supset A$$

since $T\ulcorner A \urcorner$ entails $(\forall p)(\ulcorner A \urcorner : p \supset p)$, and $\ulcorner A \urcorner : A$ delivers A as a conclusion. The converse—$A \supset T\ulcorner A \urcorner$—does *not*, in general hold. A liar sentence l is not true. It doesn't follow that it's true that it's not true.

Now, how can we construct sentences with the features of (r) and (s)? Since (s) is an arithmetic sentence, let our base model for our construction be a model of *arithmetic*, and let our language be the language of arithmetic, including $+$, $\times$, 0, 1 and the usual language of first-order predicate logic, supplemented with signification and propositional quantification (and hence, Bradwardine's truth and falsity predicates which are definable in terms of signification and propositional quantification).

The model construction is *iterative*, in the manner familiar in current model-theoretic treatments of the paradoxes [3, 4, 9]. At every STAGE of the model construction, let the model interpret arithmetic in a completely standard fashion. The *domain* of the model is the collection of natural numbers, and addition and multiplication are given their natural interpretation. For signification, we need to just choose an *encoding* of the sentences of the language, in the manner of Gödel. For each sentence, we need some effective and natural way to choose a number which encodes it—its Gödel number. If the Gödel number of a formula A is the number n, then we will let $\ulcorner A \urcorner$ be a term in the language that denotes that number ($1 + 1 + \cdots + 1$, the n-fold addition of 1 will do). The only thing to vary from one STAGE to the next is the interpretation of signification.

To interpret signification, at STAGE 0 of the construction, we will say that the numbers that are *not* Gödel numbers of any formula signify nothing. (They are non-declarative). The numbers that *are* Gödel numbers—at the start—signify *everything* (they are false). At the start of the construction, nothing is true, according to the model's 'internal' account of truth.

At STAGE $n + 1$, the interpretation of signification is defined in terms of the behaviour of the model at STAGE n. The crucial notion is that of a *variant* of a stage. At STAGE 1 we know, for example, that no matter how we interpret signification, the sentence $2 + 2 = 4$ is true, and the sentence $(\exists x)(x + 1 = 0)$ is false. All purely arithmetic sentences get their appropriate interpretation at stage 0 no matter how signification is interpreted. So, we can add, for example $\ulcorner 2 + 2 = 4 \urcorner : 2 + 2 = 4$ and rule out $\ulcorner 2 + 2 = 4 \urcorner : q$ for any arithmetical statement q that is false. However, signification is not yet completely fixed for *other* sentences. The construction discused in 'Modal Models' [8] defines the inductive procedure by which we find, at each stage the class of *variants* of a model at STAGE n (which allow signification to vary wherever it is not yet fixed by what occurs at the previous stage), and then define significantion at stage $n + 1$ to be defined by what holds in *every* variant at STAGE n. Signification for arithmetic statements is *settled* at STAGE 1, signification for those statements expressing signification of arithmetic statements is settled at STAGE 2, and so on.[3]

The model construction ensures that if, at some stage of the development, A is *settled* (A is the case, and furthermore, it will remain the case in each of the other models we construct, no matter how we further refine signification), then at the *next* stage, $T\ulcorner A \urcorner$ is be settled too. The process will eventually come to a halt (a *fixed* point, at which no more is added to signification), and at this stage, many claims of the form $A \supset T\ulcorner A \urcorner$ are true. For sentences A that are *grounded*, $A \supset T\ulcorner A \urcorner$ holds.[4]

Now we can choose our sentences to do the job of (r) and (s). The sentence S is straightforward. We will let it be the sentence[5]

$$2 + 2 = 4$$

For R, on the other hand, we must do more work. Gödel has shown us that in any theory of arithmetic strong enough, predicates have *fixed points* in the following way. If ϕ is a predicate, then there is some sentence ϕ such that we can prove in an arithmetical theory[6] that $Q \leftrightarrow \phi\ulcorner Q \urcorner$. If you like, it is a sentence which "says of itself" that it is P. I write "says of itself" in scare quotes, as this does not feature signification, which is our *official* account of what things say. However, given Bradwardine's axiom, this untutored notion of signification becomes the literal truth. If it is a theorem that $Q \leftrightarrow \phi\ulcorner Q \urcorner$, then since we have $\ulcorner Q \urcorner : Q$ by Bradwardine's axiom we can conclude that $\ulcorner Q \urcorner : \phi\ulcorner Q \urcorner$, so our official theory of signification indeed tells us that $\ulcorner Q \urcorner$ *says of itself* that it is ϕ.

[3] We then define *limit* stages as settling whatever is settled at any stage up to that limit, and a fixed point is found as usual, at some countable stage [8].

[4] See "Modal Models" [8] for a precise definition of the sense of groundedness in play here. It is not the same as being settled. Groundedness as defined there is a syntactic property of sentences. Settledness is relative to a model and a stage.

[5] Actually, in the vocabulary I chose, there is *no* sentence $2 + 2 = 4$, literally speaking, since there are no terms 2 and 4. Literally speaking, the sentence can be $(1+1)+(1+1) = ((1+1)+1)+1$, taking '2' as shorthand for '$1 + 1$', '3' as shorthand for '$2 + 1$' which is shorthand for '$(1 + 1) + 1$,' '4' as shorthand for '$3 + 1$', which is shorthand for ... etc.

[6] Which theory? Robinson's Arithmetic will do. It suffices to represent recursive functions: in particular, the diagonal function.

So, to do the job of (r) in our intuitive reflections, we want something that says that $2+2=4$ and that it is true. What will do this? We apply Gödel's trick to find a sentence R such that in an arithmetic theory we can prove

$$\mathsf{R} \leftrightarrow (2+2=4 \ \& \ \mathsf{T}\ulcorner\mathsf{R}\urcorner)$$

using the general 'predicate' $(2+2=4 \ \& \ \mathsf{T}x)$ as the context for the fixed point. Gödel proves for us that there is a fixed point and a sentence with this property. In the naïve sense, R says that $2+2=4$, and that $\mathsf{T}\ulcorner\mathsf{R}\urcorner$. In our model all truths of arithmetic are true, so Gödel's reasoning applies, and so in our model, $\mathsf{R} \leftrightarrow (2+2=4 \ \& \ \mathsf{T}\ulcorner\mathsf{R}\urcorner)$ holds. And so, the construction, as signification is refined and made more precise, $\mathsf{R} \leftrightarrow (2+2=4 \ \& \ \mathsf{T}\ulcorner\mathsf{R}\urcorner)$ will continue to hold, as this process merely refines signification, and leaves the arithmetic part of the model untouched. As our reasoning above has shown, it follows that $\ulcorner\mathsf{R}\urcorner : (2+2=4 \ \& \ \mathsf{T}\ulcorner\mathsf{R}\urcorner)$ too.

Let us pause here: we have a formal theory (arithmetic, augmented by propositional quantification and signification), and sentences R and S exhibiting the behaviour that ought to give rise to the puzzle of truth-tellers. Is there any difference between a sentence S which merely *happens* to signify its own truth, and a sentence R which does so explicitly?

What does the model construction say? What happens to R and S in the model construction? Do they share the same fate, or do they differ? At this point I will not go through all of the technical details, but give the gist of the process. In the original model S holds, since it is a purely arithmetic sentence (it does not feature signification or anything defined in terms of signification, such as T or F). So, not only does S hold: it is settled. It would continue to hold however signification is refined. This means that at the next stage of the construction $\mathsf{T}\ulcorner\mathsf{S}\urcorner$ will be settled as true too.

R fares very differently. At the first STAGE of the construction T fails of every object whatsoever. So, $\mathsf{T}\ulcorner\mathsf{R}\urcorner$ does not hold, and in the model $(2+2=4 \ \& \ \mathsf{T}\ulcorner\mathsf{R}\urcorner)$ does not hold. Since, in the first model $\mathsf{R} \leftrightarrow (2+2=4 \ \& \ \mathsf{T}\ulcorner\mathsf{R}\urcorner)$ holds, we can conclude that R does not hold either.

However, R is *not* settled, for if we vary signification in the model a little bit to make $\mathsf{T}\ulcorner\mathsf{R}\urcorner$ hold, then R would hold too (the other conjunct $2+2=4$ is no problem—it is satisfied at every STAGE). So, R does not hold at the first stage, but it is not settled as false either. So, we do not vary signification with respect to $\ulcorner\mathsf{R}\urcorner$ at this stage, and $\mathsf{T}\ulcorner\mathsf{R}\urcorner$ is not made true at the next stage. But as with the first stage, so with the second. The situation here is no different. $\mathsf{T}\ulcorner\mathsf{R}\urcorner$ does not hold, so R does not hold. But it *would* hold if we refined signification for $\ulcorner\mathsf{R}\urcorner$ a little. But since R is not settled, we do not add $\mathsf{T}\ulcorner\mathsf{R}\urcorner$ yet. And so on. In the process of model refinement, S comes out as true, but R does not.

But look! From stage two of the process, not only do S and $\mathsf{T}\ulcorner\mathsf{S}\urcorner$ hold: since $\mathsf{T}\ulcorner\mathsf{S}\urcorner$ is now the case in our model from this point on, we have $\mathsf{S} \rightarrow \mathsf{T}\ulcorner\mathsf{S}\urcorner$, and since $\ulcorner\mathsf{S}\urcorner : \mathsf{S}$ and we have Bradwardine's axiom, we can conclude $\ulcorner\mathsf{S}\urcorner : \mathsf{T}\ulcorner\mathsf{S}\urcorner$. In our construction S is also a truth-teller. Yet S came out to be true and R did not. The model treats S and R very differently.

4 WHAT DOES THIS MEAN?

What can we say about signification and truth-telling with what we have learned? The first consequence is that you cannot tell simply say that if $x : Tx$ then x is paradoxical. Truth-telling need not be a sign of paradox, even for a proponent of Bradwardine's axiom. It is *true* that if x is true, then everything that x says is the case, and hence, that x is true, and so, if x is true, it must be the case that x is true. But *that* fact is unsurprising. It is always the case that if x is true, it must be the case that x is true. (In general, if p then p, no matter what p you choose!) It is another matter to say that if $x : Tx$ then if x is true you must *first* convince yourself that x is true, and that it is true that x is true, etc. Here is an analogy: if $y : p$ then since p implies $p \& p$, then $y : (p \& p)$ too. Now, to check whether y is true or not, do we need to check p and then $p \& p$, and then $(p \& p) \& (p \& p)$ … *ad infinitum*? No, we do not. Not all unfoldings of consequences such as these lead to the type of 'ungroundedness' found in sentences that explicitly say of themselves that they are true.

What can we say about the difference between benign and benighted regresses? This must be a matter for further reflection: my contribution to the discussion from this work is that—at least in the cases I have been discussing—ungroundedness can sometimes be a syntactic matter. More precisely, you cannot tell that something is grounded or not *just* by being told some of the things that this thing *says*. If I tell you that $z : 2+2=4$ and that $z : T\ulcorner z \urcorner$, I cannot conclude that z suffers from a benighted regress, for we have seen two objects—(r) and (s)—both satisfying this condition, one of which is is true, and the other not. What matters, at least as far as the models we have been discussing is concerned, is not simply whether or not something says of itself that it is true, but the *way* it says of itself that it is true. We can verify to our satisfaction that (s)

$$2+2=4 \tag{s}$$

is true by verifying that 2 and 2 add to 4. Even though (s) says that (s) is true, we do not need to independently verify that *before* concluding that (s) is true. But the same cannot be said for (r).

$$2+2=4 \text{ and (r) is true} \tag{r}$$

Here, the syntax *demands* that we consider the matter of the truth of (r) in the process of evaluating it.

If this analysis is right, then it seems that we cannot read off whether or not a sentence is paradoxical simply in terms of what it says. What *can* we say about benighted regress? If we are to say that some sentences are properly 'ungrounded', in terms of what is groundedness defined? Is it purely syntactical?

(What then of declarative objects without syntax?) This leads to one question for further research.

QUESTION 1: What is a good theory of groundedness consistent with Read's account of Bradwardine's theory? How does it relate to signification? Is it the case that all insolubilia concerning signification are ungrounded? Is ungroundedness itself a safe notion, or does adding the concept of groundedness make our theory inconsistent?[7]

A coherent account of groundedness and signification must also address questions of epistemology and our access to truth. Can you know that something is true without knowing everything it signifies? It seems that this *must* be the case, for we can know that p without knowing all of the consequences of p. But how does knowledge interact with signification, and with groundedness?

QUESTION 2: If $x : p$ do you need to know that p in order to know that x is true? In general, what do you need to know about x's signification in order to know that x is true?

These are difficult matters: someone could write the axioms of Peano Arithmetic on a blackboard and I could concur that *that* (the long statement on the blackboard) is true. Now, for all I know, Goldbach's conjecture (that every even number is the sum of two primes) could be is a consequence of these axioms. Also, for all I know, the *negation* of Goldbach's conjecture could be a consequence of these axioms. Or, for all I know, the *unprovability* of the conjecture could be a consequence. But it seems to me that I neither know that Goldbach's conjecture holds, nor that it fails, nor that it is unprovable in Peano Arithmetic. Yet, if any of these are genuine consequences then that consequence is *signified* by the statement of the board. Surely I don't need to know this, in order to know that the axioms hold.

To conclude: not all truth-tellers are ungrounded, but some are. Sorting out which are ungrounded will give us greater insight not only into groundedness and signification, but also epistemology and other intricacies of Bradwardine's account of signification and truth.

REFERENCES

[1] DOROTHY GROVER. "Propositional quantifiers". *Journal of Philosophical Logic*, 1(2):111–136, 1972.

[2] DOROTHY GROVER. *A Prosentential Theory of Truth*. Princeton University Press, 1992.

[3] ANIL GUPTA AND NUEL BELNAP. *The Revision Theory of Truth*. MIT Press, 1993.

[4] SAUL KRIPKE. "Outline of a Theory of Truth". *The Journal of Philosophy*, 72(19):690–716, 1975.

[5] STEPHEN READ. "The Liar Paradox from John Buridan back to Thomas Bradwardine". *Vivarium*, 40(2):189–218, 2002.

[7] Consider "this sentence is neither true nor grounded" — is *that* paradoxical? It seems to be false, and ungrounded. But then, it is neither true nor grounded . . .

[6] STEPHEN READ. "Symmetry and Paradox". *History and Philosophy of Logic*, 27:307–318, 2006.

[7] GREG RESTALL. "Models for Liars in Bradwardine's Theory of Truth". To appear in *Unity, Truth and the Liar — the modern relevance of medieval solutions to semantic paradoxes*, edited by S. Rahman, T. Tulenheimo and E. Genot, Springer Verlag.

[8] GREG RESTALL. "Modal models for Bradwardine's theory of truth". *Review of Symbolic Logic*, 1(2):225–240, 2008.

[9] PETER W. WOODRUFF. "Paradox, Truth and Logic. Part I: Paradox and Truth". *Journal of Philosophical Logic*, 13(2):213–232, 1984.