

# Opacity and Paradox

Andrew Bacon

In 1963 Prior proved a theorem that places surprising constraints on the logic of intentional attitudes, like ‘thinks that’, ‘hopes that’, ‘says that’ and ‘fears that’. Paraphrasing it in English, and applying it to ‘thinks’, it states: If, at  $t$ , I thought that I didn’t think a truth at  $t$ , then there is both a truth and a falsehood I thought at  $t$ .

In this paper I explore a response to this paradox that exploits the opacity of attitude verbs, exemplified in this case by the operator ‘I thought at  $t$  that’, to block Prior’s derivation. According to this picture, both Leibniz’s law and existential generalization fail in opaque contexts. In particular, one cannot infer from the fact that I’m thinking at  $t$  that I’m not thinking a truth at  $t$ , that there is a particular proposition such that I am thinking *it* at  $t$ . Moreover, unlike some approaches to this paradox (see Bacon et al. [4]) the failure of existential generalization is not motivated by the idea that certain paradoxical propositions do not exist, for this view maintains that there *is* a proposition that I’m not thinking a truth at  $t$ . Several advantages of this approach over the non-existence approach are discussed, and models demonstrating the consistency of this theory are provided. Finally, the resulting considerations are applied to the liar paradox, and are used to provide a non-standard justification of a classical gap theory of truth. One of the main challenges for this sort of theory — to explain the point of assertion, if not to assert truths — can be met within this framework.

## 1 Prior’s paradox

In [30] Arthur Prior describes a remarkable theorem of propositionally quantified logic. Letting  $Q$  represent a unary propositional operator and  $X$  a sentential variable, it may be stated as follows:

**Prior’s Theorem**  $Q\forall X(QX \rightarrow \neg X) \rightarrow \exists X(X \wedge QX) \wedge \exists X(\neg X \wedge QX)$

The result is general, in that no special principles about the operator  $Q$  are assumed in its proof. Thus it has instances in which  $Q$  is replaced with negation or metaphysical necessity. But these instances are unsurprising. Prior is primarily concerned with instances in which  $Q$  is substituted for various intentional attitudinal verbs, such as ‘ $S$  fears that’, ‘ $S$  hopes that’, ‘ $S$  said that’, and so on. Reading  $Q$  as ‘Simon said at  $t$  that’, Prior would paraphrase his theorem as follows:

If Simon said at  $t$  that Simon didn't say anything true at  $t$ , then Simon has said a truth and a falsehood at  $t$ .

The puzzle is this: it seems evident that Simon could have said, at  $t$ , that he didn't say anything true at  $t$ , and not have said anything else. In which case he would not have said both a truth and a falsehood, he would have said at most one thing, contradicting the theorem.

Prior's proof follows reasoning that should by now be familiar from the liar, and related paradoxes. Suppose Simon had said, at  $t$ , that he didn't say anything true at  $t$ . Then either he didn't say anything true at  $t$ , or he did. Suppose the former. Then, since the proposition that he didn't say anything true at  $t$  is a proposition he said at  $t$ , it must be false. So it's not the case that he didn't say a truth at  $t$ . Contradiction. Suppose the latter: he did say a truth at  $t$ . Then the proposition that he didn't say a truth at  $t$ , is itself false, and thus a falsehood he said at  $t$ . So he has said a truth and a falsehood at  $t$ .

The informal reasoning above belies the inevitability of Prior's result. In order to appreciate this, some points about the relation between Prior's theorem, and the subsequent paraphrase in English, are in order. Prior's formal proof will be stated in the next section. First, Prior's theorem is stated in propositionally quantified logic: a language containing the familiar truth functional connectives, propositional letters, the unary operator  $Q$ , but also variables that can occupy the position of sentences, and quantifiers that can bind them. The English paraphrase replaces quantification into sentence position with singular quantification over propositions. Since a first-order variable cannot grammatically occupy the position of a sentence, occurrences of bound sentential variables are replaced by first-order variables as arguments to a propositional truth predicate.

Second, because of the use of the propositional truth and falsity predicates, certain avenues for resisting Prior's theorem that are suggested from inspection of the informal argument, are really illusory. This includes, for example, co-opting any of the familiar non-disquotational accounts of sentential truth to play the role of propositional truth, allowing propositions to be neither true nor false, or both true and false, but without relinquishing classical logic. For Prior's paradox is not really stated in terms of propositional truth or falsity. Where, in the paraphrase, we would assert that the proposition that  $P$  is true, the target of the paraphrase would simply assert  $P$ . And where, in the informal argument, we seem to move freely between 'the proposition that  $P$  is true' and ' $P$ ', no such move is made in Prior's official argument.

Other familiar responses to the liar paradox are also of no use here. For example, after proving that a given language  $L$ , that can talk about its own sentences, does not contain a predicate that is true of all and only the true sentences in  $L$ , some will emphasize that the matter is not hopeless, since there might be another, more expressive language,  $L^+$ , that *can* express truth in  $L$ . This goes along with the general idea that everything is in principle expressible, but not all at once in the same language. Yet Prior's theorem purports to tell us that there is a particular proposition — that Simon didn't say anything true at  $t$  — that Simon cannot say uniquely at  $t$ . The theorem doesn't have qualifications

— that he can't say the proposition in this particular language or that — it says that he can't say it uniquely at *t at all*. Thus he can't say it with a sentence of English, of Russian, a more expressive extension of these languages, or with an elaborate round of charades, unless he also says something else along with it.

Of course, the paraphrase in terms of first-order quantification is necessary because we do not have the equivalent of pronouns for sentences, or quantifiers that can bind them in ordinary English. But quantification into other grammatical positions, aside from first-order quantification, is possible in English, including the positions of plural and prepositional phrases, suggesting this restriction is only accidental to English (Prior [31], Rayo and Yablo [33], Boolos [7]). And even if not, why think that quantification into sentence position is intelligible only if it can be translated into a language that already has it? As children we are evidently able to learn first-order quantification in our native tongue, without first translating it into an already understood language, so why should these other modes of quantification be any different (Williamson [41])?

Might one nonetheless try to resist Prior's theorem by denying the intelligibility of quantification into sentence position? Those who do so will find the first-order paraphrases on better footing, but might reject the reasoning as illicit, since it involves principles concerning propositional truth that are extremely contentious in the context of sentential truth. For instance, we have freely moved between 'the proposition that *P* is true' and '*P*', in order to be faithful to Prior's reasoning, and this looks similar to the move, that is illicit given classical logic, between 'the sentence '*P*' is true' and '*P*'. But unlike sentential truth, these inferences are unproblematic in the propositional setting (and only become problematic when combined with a structured theory of propositions, or some other theory of propositional granularity in which propositions behave enough like sentences to define things like substitution and diagonalization). Indeed, a minimal theory of first-order propositions and truth can be consistently formulated in which all of Prior's reasoning may be faithfully represented.<sup>1</sup> It is most naturally formulated in a type theory, à la Church, that contains the type of first-order predicates  $e \rightarrow t$ , expressions that combine with names to form sentences, and a converse type  $t \rightarrow e$  of expressions that combine with sentences to form names. Given two primitives, *true* of type  $e \rightarrow t$ , and *that* of type  $t \rightarrow e$ , a minimal theory of first-order propositions can be formalized by an axiom stating that *true* is a left-inverse of *that* (in the type theory,  $\lambda X.true\ that\ X = \lambda X.X$ ). The consistency of such a theory roughly amounts to there not being 'more things' of type *t* than of type *e*, as in a full model of type theory *that* can be interpreted by any injection in type  $t \rightarrow e$ , and *true* by any of its left-inverses.<sup>2</sup> Despite being exceedingly minimal — the theory does not

<sup>1</sup>See Bacon [3], section 2.1.

<sup>2</sup>I do not make the assumption, sometimes made, there are only two elements in the domain of type *t*, the truth values, or even that the domain of type *t* consists of sets of worlds. A suitably general class of models free of these assumptions is described in Benz Müller et al [6]. The matter is slightly more subtle in non-full Henkin models of type theory, as it turns out that having a left-inverse and being injective are not provably equivalent in standard axiomatizations of higher-order logic: see the appendix of Bacon [2].

prove that Julius Caesar and the proposition that grass is green are distinct, for example — I think the theory is suitable for almost all of the things one might want from a first-order theory of propositions. In particular, it licenses the intersubstitutivity of ‘the proposition that  $P$  is true’ with ‘ $P$ ’.<sup>3</sup>

## 2 Possible Responses to Prior’s Paradox

Having accepted the intelligibility of Prior’s language, here, finally, is Prior’s proof. It has two assumptions:

**Classical Propositional Logic** All classical tautologies, Modus Ponens.

**Universal Instantiation**  $\forall X A \rightarrow A[B/X]$ , provided  $B$  is substitutable for  $X$  in  $A$ .

Propositionally quantified logic has further axioms beyond Universal Instantiation, but it is the only principle appealed to in Prior’s proof. We also take  $\forall, \neg$  and  $\rightarrow$  as primitive, and define  $\exists X A$  as  $\neg \forall X \neg A$  and  $\wedge$  as  $\neg(A \rightarrow \neg B)$ .<sup>4</sup> Here is Prior’s proof, translated from Polish notation, and annotated so as to use only the assumptions above.

1.  $\forall X(QX \rightarrow \neg X) \rightarrow (Q\forall X(QX \rightarrow \neg X) \rightarrow \neg \forall X(QX \rightarrow \neg X))$  (UI)
2.  $Q\forall X(QX \rightarrow \neg X) \rightarrow (\forall X(QX \rightarrow \neg X) \rightarrow \neg \forall X(QX \rightarrow \neg X))$  (1, CL)
3.  $Q\forall X(QX \rightarrow \neg X) \rightarrow \neg \forall X(QX \rightarrow \neg X)$  (2, CL)
4.  $Q\forall X(QX \rightarrow \neg X) \rightarrow \exists X(QX \wedge X)$  (3, CL,  $\exists$ Def)
5.  $Q\forall X(QX \rightarrow \neg X) \rightarrow (Q\forall X(QX \rightarrow \neg X) \wedge \neg \forall X(QX \rightarrow \neg X))$  (3, CL)
6.  $(Q\forall X(QX \rightarrow \neg X) \wedge \neg \forall X(QX \rightarrow \neg X)) \rightarrow \exists X(QX \wedge \neg X)$  (CL,  $\exists$ Def,  $\wedge$ Def)
7.  $Q\forall X(QX \rightarrow \neg X) \rightarrow \exists X(QX \wedge \neg X)$  (5, 6, CL)
8.  $Q\forall X(QX \rightarrow \neg X) \rightarrow (\exists X(QX \wedge X) \wedge \exists X(QX \wedge \neg X))$  (4, 7, CL)

---

<sup>3</sup>In type theory one may define a binary connective,  $=$ , that takes two sentences,  $A$  and  $B$ , and produces another sentence informally corresponding to the identity of  $A$  and  $B$ . The principle of  $\beta$ -equivalence, that  $(\lambda p.M)N = M[N/p]$  provided  $N$  is substitutable for  $P$ , ensures the following identities:

1.  $\text{true that } P = (\lambda X.\text{true that } X)P$  by  $\beta$ .
2.  $(\lambda X.\text{true that } X)P = (\lambda X.X)P$  by the minimal theory.
3.  $(\lambda X.X)P = P$  by  $\beta$ .

<sup>4</sup>In order for the proof to go through on these assumptions, the choice of primitives matters. Since neither assumption involves existential quantification, for example, it cannot be taken as primitive, or we would not be able to prove sentences involving it (such as Prior’s theorem).

If we are to avoid Prior’s conclusion, then we must reject one of the two assumptions. That is, one must either weaken classical propositional logic, or reject the principle of Universal Instantiation. (Given the duality of the quantifiers and the first assumption, Universal Instantiation is equivalent to Existential Generalization:  $A[B/X] \rightarrow \exists X A$  provided  $B$  is substitutable for  $X$  in  $A$ . We will treat these two equivalent formulations interchangeably.)

Both options have precedents. Weakening classical logic is a common response to other paradoxes, including the semantic and set theoretic paradoxes.<sup>5</sup> And there is a tradition of rejecting Universal Instantiation going back to Russell’s ban on impredicativity: a quantified proposition cannot belong to the domain over which it quantifies.<sup>6</sup> Since propositional quantification is thus necessarily restricted, and since the logic of restricted quantification does not include Universal Instantiation, Universal Instantiation is not valid.

Of course, another option is to simply accept the logic of Prior’s argument and follow it where it leads. Indeed, this is where my own sympathies lie (Bacon [3]). A radical interpretation of Prior’s theorem along these lines is that when Simon attempts to utter the sentence ‘Simon didn’t say anything true at  $t'$ ’, a mysterious force prevents him from completing his utterance. But this seems like a far-fetched moral to draw, and luckily one can accept Prior’s logic without drawing it. Rather, it is perfectly possible for Simon to make the utterance of the *sentence* ‘Simon didn’t say anything true at  $t'$ ’, but in doing so he doesn’t succeed in saying that Simon didn’t say anything true at  $t$ . According to Prior’s own interpretation, he doesn’t succeed in saying anything at all. This interpretation is not that radical, for few people would wish to identify saying that  $P$  with making the sounds that constitute uttering the sentence ‘ $P$ ’. As a way of warming up to this idea, note that one might cough in a way that, by fluke, sounds exactly like how one would say that snow is white in Farsi. But few would regard this as a way of saying that snow is white. More realistically, one might utter the sentence ‘that dog looks funny’ while failing to point out a dog, or even anything, and thus fail to say something with a sentence that, in other circumstances, *could* be used to say something. But Prior’s theorem, as we have noted already, is quite general and there are many other instances for which the costs are higher. For instance, consider:

If Simon tried to say, at  $t$ , that Simon didn’t try to say anything true at  $t$ , then he tried to say a truth and a falsehood at  $t$ .

Suppose that at  $t$  Simon utters the sentence ‘I didn’t try to say anything true at  $t'$ ’ (with full understanding of English, the intentions to assert, etc), and does nothing else. If we apply Prior’s diagnosis here, we should say that when Simon utters the sentence ‘I didn’t try to say anything true at  $t'$ ’ he fails to say anything

---

<sup>5</sup>Priest, for example, recommends weakening classical propositional logic in response to Prior’s paradox in [28].

<sup>6</sup>Ramification is spelled out as a response to Prior’s paradox in Tucker and Thomason [39], and Kaplan [20] and Kripke [22] express sympathy to this line of response in the context of some related paradoxes. See Tucker [40] for a non-standard approach to the ban on impredicativity, that does not involve explicit ramification.

at all. But this response is of no avail: one must deny that Simon even *tried* to say something. For if we concede that he tried to say that he didn't try to say anything true at  $t$ , then we must accept the absurd conclusion that there were two things he tried to say at  $t$ .<sup>7</sup>

In fact, accepting Prior's theorem at face value commits us to one of the following (whether we follow Prior's diagnosis of the original paradox or not): (i) that Simon didn't even try to say that he didn't try to say anything true at  $t$ , despite strong appearances to the contrary, or (ii) that Simon tried to say two things, one false one true, again, despite strong appearances to the contrary. So the cost of accepting Prior's theorem is high, and a thorough examination of the logic that ensures it seems like the responsible course of action, even if we ultimately choose to accept the conclusion.

### 3 Non-existent Propositions

What should the Universal Instantiation denier say about Simon's utterance? According to the most plausible theories that accept Prior's theorem, uttering the sentence ' $P$ ' is not sufficient for having said that  $P$ . Indeed, we have argued that it cannot be sufficient for having even *tried* to say that  $P$ .

To distance themselves from the pathological results of the fully classical view, a denier of Universal Instantiation ought to endeavour to uphold what I will call the *naïve model of speech*. According to this view, there is an action one can perform freely to a sentence ' $P$ ', which I shall call uttering ' $P$ ', which does suffice for saying that  $P$ , and moreover when one utters nothing else beyond ' $P$ ', one says nothing apart from that  $P$ , in normal circumstances.<sup>8</sup> Of course, the naïve view will have to be qualified in various ways to take into account context sensitivity; but we will set that aside by restricting attention to the pared down language described in section 2, and by imposing the working assumption that its expressions are context insensitive.<sup>9</sup> Such a view can accept all of the following: that uttering ' $P$ ' typically suffices for one having tried to say that  $P$  (bracketing context sensitivity), and that typically, when one tries to say that  $P$  one says that  $P$ .

What does this amount to with respect to Prior's paradox? According to the naïve model, when Simon utters the sentence 'I didn't say anything true at  $t$ ' he succeeds in saying that he didn't say anything true at  $t$ . But we cannot

<sup>7</sup>Cases related to this are discussed in Bacon [3], and explained in the context of a different interpretation of Prior's result.

<sup>8</sup>Thus, according to our understanding of utterance, a cough that sounds like a sentence is Farsi does not count as an utterance of that sentence, but the noise Simon makes in our envisioned situation, does count as an utterance.

<sup>9</sup>Of course, many people have argued that attitude verbs like 'say that' are context sensitive, and that quantifiers, including the propositional quantifiers appearing in Prior's language, are context sensitive. Some have even suggested that the context sensitivity of these things are key to understanding the paradoxes, and so should not be bracketed (for instance, Parsons [27], Glanzberg [17], Burge [8], Simmons [38]). But we are exploring these views, we are exploring the idea that the paradoxes can be solved without postulating context sensitivity, so bracketing the context sensitivity of these expressions is indeed the proper methodology.

apply Existential Generalization (an equivalent of Universal Instantiation) to conclude that there is something that Simon has said:

1. Simon said, at  $t$ , that Simon didn't say anything true at  $t$ .
2. For no  $X$  did Simon say that  $X$ .

This, of course, blocks the argument. For instance, at one point of the argument we had supposed that Simon didn't say anything true at  $t$ , and attempted to show, contrary to the assumption, that this proposition is itself a truth that Simon said. Given 2 we should indeed accept this supposition: it is vacuously true, because Simon hasn't said anything. We can also maintain 1, that Simon said that Simon didn't say anything true at  $t$ ; we just can't conclude from this that Simon said a truth — this is an instance of Existential Generalization that is rejected.

One way of understanding 2 is as an instance of a *no-proposition* view. Like the fully classical logician, this theorist accepts that Simon uttered the sentence 'I didn't say anything true at  $t$ ' but that he didn't succeed in saying anything. But it is a different kind of no-proposition view, because despite this he did succeed in saying that he didn't say anything true at  $t$ .

There are a couple of precedents for this sort of view. One, mentioned already, stems from Russell's ban on vicious circles, according to which a quantified proposition cannot belong to the domain over which it quantifies.<sup>10</sup> A consequence of this ban is that all quantification into sentence position is restricted quantification, for the range of the quantifier in a quantificational claim like  $\forall X A$ , must be restricted in such a way that it does not range over  $\forall X A$  itself. Thus an instance of Universal Instantiation that instantiates  $X$  with  $\forall X A$  is not legitimate (the instance:  $\forall X A \rightarrow A[\forall X A/X]$ ). More generally, we know that the logic of restricted quantification does not include the principle of Universal instantiation, as, after all, the following inference concerning the restricted quantifier 'every  $F$ ' is clearly invalid:

Every  $F$  is  $G$

Therefore,  $a$  is  $G$ .

Theorists sympathetic to Russell's ban on impredicativity generally pursue ramified approaches to propositional quantification, in which there is a hierarchy of different quantifiers, each quantifier restricted to ranging over propositions involving quantifiers from lower in the hierarchy. Each quantifier is restricted, but each proposition is in the range of some quantifier or other.<sup>11</sup>

<sup>10</sup>There are some substantial assumptions that must be taken on in order to even formulate this sort of ban. For instance, one cannot accept a theory of propositional granularity, such as the possible worlds theory, in which a quantified proposition, like  $\forall x x = x$  and a non-quantified claim, like  $Fa \rightarrow Fa$ , are identified, for otherwise one would not be able to make the distinction between quantified and non-quantified propositions.

<sup>11</sup>Although my remark here is representative of what a ramified theorist would like to say, it is strictly speaking nonsense. For in order for it to express a truth, I must have quantified

Another precedent for rejecting Universal Instantiation arises in the treatment of empty names within free logic.<sup>12</sup> According to common sense, there is no such thing as Pegasus, the winged horse-god of Greek mythology. But we still want to maintain that Pegasus is a mythical winged horse, or that the ancient Greeks told stories about Pegasus. One can also look (unsuccessfully) for Pegasus, and believe (mistakenly) that Pegasus is a horse, or that he would have been a horse-god had the mythology been true. We seem to want to assert that there are no mythological horse-gods, but at the same time assert that Pegasus is a mythological horse-god, that there is nothing you are in fact looking for when you are looking for Pegasus, and so on. But these are counterexamples to Existential Generalization (and thus Universal Instantiation).

On this interpretation we do not conceive of the failures of Universal Instantiation arising because the quantifiers are restricted in some way. For according to a view that takes the above judgments at face value, there is not some more expansive quantifier that ranges more widely over things like Pegasus, Sherlock Holmes, Vulcan, and so on. The counterexamples to Universal Instantiation, from this perspective, involve the *unrestricted* quantifiers.<sup>13</sup>

This perspective on Universal Instantiation can be applied to Prior's paradox. For instance, verbs like *says*, *hopes*, *fears*, and so on, create the sorts of contexts from which one cannot existentially generalize. They are not *existence entailing*: to illustrate, *being a horse* is existence entailing, since we may infer from the (false) assumption that Pegasus is a horse that he exists, whereas *being a mythical horse* is not, for the analogous entailment seems false. It seems like Simon could hope that Pegasus was real, but we wouldn't want to infer that there is something that Simon hopes is real. Similarly, he could have said that Pegasus is real, without there being anything that he said is real. One might therefore postulate things that stand to the grammatical category of sentences as empty names stand to the category of singular terms, and have the same attitude towards them as the free logician has towards empty names. Thus, one can say that *P* without there being anything that one has said. The proposition that Simon didn't say a truth at *t* must be like Pegasus in the sense that it can have certain non-existence entailing properties, like *being said* or *being hoped* (or, in the latter case, *being said/hoped to be real*) without existing.

The former view is less naturally classified as a no-proposition view. For

---

over quantifiers unrestrictedly. But if I could do that, I could introduce an unrestricted quantifier:  $\text{everything}_U$  is *F* if and only if, for every quantifier,  $\text{everything}_i$ ,  $\text{everything}_i$  is *F*. Whether this is the sort of Wittgensteinian nonsense one can simply kick away is a contentious matter, as is the status of heuristically useful nonsense more generally, and I won't attempt to adjudicate here. But see Williamson [41].

<sup>12</sup>This motivation is explored in Bacon et al [4], alongside the ramificationist response.

<sup>13</sup>Free logicians take failures of Existential Generalization with varying degrees of seriousness. For instance, some explain the failures by saying that atomic sentences involving empty names are by default false, and that empty names do not in general make an important semantic contribution to a sentence in which they occur. In the cases where they appear to, such as in intentional verbs, special mechanisms are invoked (see, for instance, Sainsbury [34]). I have defended a view that takes failures of Existential Generalization seriously, indeed, which takes them as primitive and not to be explained by any general property of the name 'Pegasus', but in terms of what properties Pegasus in fact has (see Bacon [1]).



each level of the quantifier hierarchy,  $i$ , there is a paradoxical sentence Simon can utter for which he can truthfully report, of his paradoxical utterance, both ‘Simon hasn’t said anything <sub>$i$</sub> ’ and ‘there is something <sub>$i+1$</sub>  that Simon said’.<sup>14</sup> And both views should be distinguished, as noted above, from the classical no-proposition view which maintains that, in addition to not saying anything at  $t$ , Simon hasn’t said that he didn’t say anything true at  $t$ .

Do these views avoid the trappings of the classical no-proposition view? A common problem for adherents of the classical no-proposition theory is that they have trouble stating their own view. For instance, someone like Prior might wish to communicate their view about Simon’s utterance at  $t$  by making their own utterance of the following: ‘Simon didn’t say anything true at  $t$ ’, perhaps by clarifying ‘because he didn’t say anything at all’. But of course, this is an utterance of the very sentence Simon failed to say anything with. Prior’s response is that, while one utterance of a sentence might fail to say anything, another can. Thus it is utterances, not sentences, that are the true couriers of successful assertion. But this response is insufficiently general, as variants can be formulated that do not fair so well:  $D =$  ‘no utterance of  $D$  can be used to say a truth and only truths’. A routine argument, mimicking Prior’s, shows that no utterance of  $D$  can be used to say truths and only truths. But the theorist might want to communicate this result as well, by uttering the sentence ‘no utterance of  $D$  can be used to say a truth and only truths’. In virtue of being an utterance of the problematic sentence itself, they will not succeed in saying anything.<sup>15</sup>

In this regard the non-classical no-proposition views do better. For they can maintain that, even though an utterance of ‘no utterance of  $D$  can be used to say a truth and only truths’ will not result in you saying anything, one might still end up saying, with such an utterance, that no utterance of  $D$  can be used to say a truth and only truths. What good is this, if one hasn’t said anything? Well, the purpose of assertion is presumably to pass on beliefs and knowledge to others which can inform their actions, and so forth. But intentional verbs like belief and knowledge are paradigm instances of the sorts of words that create non-existence entailing contexts: for example, Alice may believe that Pegasus is real, even though there isn’t anything that she believes is real.<sup>16</sup> Thus in saying that  $P$  one may pass on knowledge that  $P$ , or a belief that  $P$ , even if there is no proposition you’ve said and which your intended audience thereby knows or believes. One might worry that the intentionality will eventually peter out. For instance, if the beliefs and knowledge acquired are to be efficacious, they had better inform our actions, but whether one has done something seems to be an existence entailing context, as it appears to be objective and not tangled up with our attitudes. Call a belief existent if, for some  $P$  it is a belief that  $P$ , and non-

<sup>14</sup>The slightly awkward metalinguistic formulation here is to avoid the Wittgensteinian nonsense alluded to in footnote 11.

<sup>15</sup>Variants of this paradox are discussed in Hazen [18], Zardini [42], Bacon [3].

<sup>16</sup>And while Alice cannot know that Pegasus is real, she can know that Pegasus *isn’t* real, from which one still can’t infer that there’s something Alice knows isn’t real (or else one could infer that there’s something that isn’t real).

existent otherwise. Non-existent beliefs, by impacting your credences, might raise the value of existing propositions which you are in a position to make true. So non-existent beliefs can be efficacious, even if we assume that *making true* is an existence entailing context. That said, the notion of making a proposition true is arguably also a propositional attitude, albeit a factive one. It is not simply a matter of certain things happening, but of them happening as a result of the agents intentions, and is thus like knowledge, belief and saying in the relevant respects. For instance, given the naïve model of saying, Simon appears to be in control of what he says and doesn't say, as he is certainly in control of his utterances. When Simon utters the sentence 'nothing I've said at  $t$  is true' he doesn't say anything true. Since he was responsible for this, we should maintain that Simon has made it the case that nothing he said at  $t$  is true. Thus making true is not existence entailing, given the assumption, currently being explored, that there does not exist a  $P$  identical to the proposition that nothing Simon said at  $t$  is true.<sup>17</sup>

In order to uphold the naïve model of speech in full generality, we need not only to accommodate the case in which Simon utters 'nothing I said at  $t$  is false'. For he could have said any other sentence at  $t$ . For instance, if he uttered 'snow is white' then he would have said, instead, that snow is white. Indeed, for any sentence  $A$ , 'it is possible that Simon said that  $A$ ' should be true given the naïve model of saying, since it's possible that Simon utter ' $A$ '. Augmenting the language in section 2 with an operator  $\Diamond$ , representing possibility, we may formulate this as a schema:

### Possible Saying $\Diamond QA$

Note, however, that the naïve model of saying motivates something stronger: for given that it's possible that Simon utter 'snow is white' and nothing else, it should be possible that Simon say that snow is white without saying anything else. If we add a propositional identity connective,  $=$ , we might try to formulate this as follows:<sup>18</sup>

$$\Diamond(QA \wedge \forall X(QX \rightarrow X = A))$$

But this seems insufficiently strong without Universal Instantiation. For instance, this principle is consistent with the hypothesis that Simon can't say at  $t$  that snow is white, without also saying that nothing he said at  $t$  is true.  $\forall X(QX \rightarrow X = A)$  entails that any *existing* proposition which Simon has said at  $t$  must be identical to the proposition that snow is white. But this is satisfied when Simon says both that snow is white and that nothing Simon said at  $t$  is true, since the latter is not an existing proposition. In [4] a primitive notion of saying uniquely,  $Q!$  is introduced to deal with this: it is not defined in terms of  $Q$ , the quantifiers and an identity connective, but is taken as basic, and characterized by a model theoretic stipulation (more on this below), and moreover subject to the laws:

<sup>17</sup>Similarly, Simon can make it true that he said that Pegasus is real, even if there is nothing such that he made it true that he said that *it* is real.

<sup>18</sup>This is the principle called K2 in Bacon et al. [4].

**Subsumption**  $\Box(Q!A \rightarrow QA)$

**Uniqueness**  $\Box(Q!A \wedge Q!B \rightarrow A = B)$

We may thus articulate the desired component of the naïve model as follows:<sup>19</sup>

**Possible Saying!**  $\Diamond Q!A$

It is important to this project to know that Possible Saying!, Subsumption and Uniqueness are indeed consistent with a free version of propositionally quantified logic that does not contain Universal Instantiation. This is undertaken in Bacon et al. [4], where various models of this, and related theories, are constructed.

## 4 Opacity

In this section I will be exploring another view that responds to Prior’s paradox by rejecting Universal Instantiation. But it is a non-standard version of this response: it is not motivated by any general logical constraint, such as Russell’s ban on impredicative quantification. Nor does it follow from considerations of non-existence, as the counterexamples involving names like Pegasus. In fact, our quantifications will in general be interpreted as ranging unrestrictedly, and we will accept a schema,  $\exists X(X = A)$ , to the effect that all sentences express existing propositions: there are no sentences that stand to the propositional quantifiers as empty names stand to the first-order quantifiers.

The key observation is that the problematic instances of Prior’s theorem all involve intentional attitudes — attitudes like saying, fearing, hoping and so forth — that are commonly supposed to create *opaque* contexts. We have already noted that intentional attitudes create contexts which are not existence entailing. I will argue that opacity also creates contexts in which existential generalization is not permissible, albeit for fundamentally different reasons than in the cases involving non-existence, and restricted quantification.

Suppose that Simon knows that Hesperus is visible in the evening and Phosphorus is visible in the morning, but does not realize that Hesperus and Phosphorus are the same. He is looking at the sky in the evening and forms the belief that Hesperus is bright:

1. Simon believes that Hesperus is bright.

Because he does not realize they are the same, he form no such belief about Phosphorus

2. Simon does not believe that Phosphorus is bright.

But, of course, they are the very same planet:

3. Hesperus is Phosphorus.

This seems to be a counterexample to Leibniz’s law.

---

<sup>19</sup>This is called  $K2^+$  in Bacon et al. [4].

**L()**  $a = b \rightarrow A \rightarrow A[a/b]$

There are many different replies to this puzzle. Most assimilate it to some sort of equivocation. According to one version of this idea, words like ‘believes’ are context sensitive, and the attitude being ascribed by the word ‘believes’ are different in 1 and 2.<sup>20</sup> According to another, the equivocation is due to the names ‘Hesperus’ and ‘Phosphorus’: in embedded contexts names refer not to their customary referents, but to their customary senses.<sup>21</sup>

In [5], Jeff Russell and I have investigated the suggestion that we take the judgments 1-3 at face value, as counterexamples to Leibniz’s law. But this is entirely consistent with various quantified variants of Leibniz’s law, including:

**L(xy)**  $\forall x \forall y (x = y \rightarrow A \rightarrow A[x/y])$

(Sometimes principles like  $L()$  and  $L(xy)$  are restricted to instances involving direct predications; in this case, for instance,  $\forall xy (x = y \rightarrow Fx \rightarrow Fy)$ . Following [5] I shall treat these principles interchangeably in the presence of a device of  $\lambda$ -abstraction that allows one to turn a term  $t$  occurring in an arbitrary context  $A$ , into a direct predication  $(\lambda x A)t$ .<sup>22</sup> In our discussion of first-order logic, where  $\lambda$  is not present, the unrestricted principles are stronger, and will be our focus.)

$L(xy)$  has been thought to be plausible, even by those who reject  $L()$  on broadly Fregean grounds. For while *names* may be associated with interesting senses, which can have interestingly different cognitive import, bound variables refer directly to the things they denote. Let  $Bx$  represent the open sentence ‘Simon believes  $x$  to be bright’.<sup>23</sup> Then we should accept:

$$a = b \wedge Ba \wedge \neg Bb$$

But given Existential Generalization we would be able to infer:

$$\exists x \exists y (x = y \wedge Bx \wedge \neg By)$$

But this directly contradicts an instance of  $L(xy)$ . So Existential Generalization cannot be valid, and thus neither can Universal Instantiation. Indeed, this is evident when one observes that  $L()$  can be derived from  $L(xy)$  by two applications of Universal Instantiation (instatiating  $x$  with  $a$  and  $y$  with  $b$ ).

It has long been suggested that opaque contexts generate failures of Existential Generalization (Quine [32], Kaplan [19]). For instance, many people believe that a single person, who has since been dubbed ‘Jack the Ripper’, terrorized

<sup>20</sup>See, e.g., Crimmins and Perry [11].

<sup>21</sup>Frege [15].

<sup>22</sup>In doing so, I am thus setting aside approaches that treat restrictions of  $L()$  involving direct predications as special. Such approaches typically deny the equivalence between  $(\lambda x A)t$  and  $A[t/x]$  (see Kripke [21] and Salmon [35]).

<sup>23</sup>Some philosophers place special significance on simple predications, and will therefore object to this sort of terminological shorthand. They will think that ‘Hesperus is believed by Simon to be bright’ to be different from ‘Simon believes that Hesperus is bright’. By simply replacing the former formulations with the latter in the following discussion, we may satisfy those who object to these manipulations, but at the cost of readability.

Whitechapel with a series of murders in the 1888. Thus we believe that Jack the Ripper committed the Whitechapel murders. But we don't know who committed the murders, so there is no person such that we believe that *they* committed the Whitechapel murders.

It is worth emphasizing how these failures of Existential Generalization differ from the failures due to non-existence. For in the first case, there is something, namely the planet Venus, that is identical to both Hesperus and Phosphorus. Hesperus and Phosphorus *exist*. So, too, does Jack the Ripper, assuming, as commonly supposed, that there was a single individual who committed the murders in 1888. So the failures of Existential Generalization are fundamentally different from the failures instanced in inference from 'Pegasus is a mythical winged horse' to 'something is a mythical winged horse', since in that case there is no such thing as Pegasus. That is, we have a counterexample to a principle, weaker than Existential Generalization, that is validated in the free logics designed to deal with empty-names (see, e.g., Nolt [26]):

**Unrestricted Export**  $\exists x a = x \wedge A[a/x] \rightarrow \exists x A$

Unrestricted Export is derivable given some pretty uncontroversial quantificational logic, and Leibniz's law,  $L()$ .<sup>24</sup> But since Leibniz's law is part of what is at stake in these cases, this argument has little suasive force.

Let's summarize this with an important definition.

**Opacity** A context,  $A$ , is *transparent*, relative to a language, iff every instance of the schema  $a = b \rightarrow A \rightarrow A[b/a]$  in that language is true. It is *opaque* otherwise.

There are philosophers who will be in verbal agreement with much of the foregoing. They will countenance failures of  $L()$  (but not  $L(xy)$ ), and will subsequently agree that there are opaque predicates and predicates that do not license export. But they do not take opacity 'seriously' in my sense, for they will maintain that attitude verbs don't create genuine contexts, whose meanings can be computed compositionally, any more than quotation marks do. What, on the surface, looks like an operator expression — 'Simon believes that ...' — in reality functions like a predicate of sentences, 'Simon believes "..."', and what looks like a context is merely a typographical constituent of a quotation name for a sentence. Provided such philosophers can make sense of my remarks in their preferred framework, it will not matter much, but I will not make systematic attempts to accommodate them.

Putting all this together, what might a first-order logic of opacity look like? We will assume a first-order language with non-logical predicates and operators that may be opaque. Let's start with the quantificational fragment:

**Quantified Instantiation**  $\forall x(\forall y A \rightarrow A[x/y])$

**Normality**  $\forall x(A \rightarrow B) \rightarrow \forall x A \rightarrow \forall x B$

---

<sup>24</sup>See Bacon and Russell [5].

**Quantifier Exchange**  $\forall x \forall y A \rightarrow \forall y \forall x A$

**Vacuous Quantification**  $A \rightarrow \forall x A$  when  $x$  does not occur free in  $A$

Notice that in place of the principle of Universal Instantiation,  $\forall x A \rightarrow A[t/x]$ , we have its universal closure, Quantified Instantiation.<sup>25</sup> This corresponds to the idea, implicit in our discussion of  $L(xy)$ , that bound variables, unlike names, are not susceptible to identity confusions. Thus one ought to be able to instantiate bound variables.<sup>26</sup>

Besides Quantified Instantiation, one should also think that instantiation of arbitrary terms is legitimate provided it is not in an intentional context. The distinction between intentional and non-intentional contexts is not a distinction that can be made prior to the interpretation of the non-logical constants. But we may assume purely logical contexts are transparent, independently of how the non-logical constants are interpreted, and so we should be able to instantiate into purely logical contexts. More generally, we should be able to instantiate even in non-logical contexts, provided the variable is not within the scope of a non-logical operation:<sup>27</sup>

**Logical Instantiation**  $\forall y A \rightarrow A[t/y]$  provided  $y$  is not in the scope of a non-logical operation.

To this we must add principles governing identity:

**L(xy)**  $\forall xy(x = y \rightarrow A \rightarrow A[y/x])$

<sup>25</sup>Note, also, that our axiomatization includes Quantifier Exchange — a principle that is usually derivable with the help of Universal Instantiation. It is provable, even in free logic, from  $L()$ . But without either  $L()$  or Universal Instantiation it is not, and must be put in by hand: see Fine [14].

<sup>26</sup>One might have thought that the quantified version is stronger, but one quickly sees that to derive Universal Instantiation from Quantified Instantiation, you have to already have Universal Instantiation in order to instantiate  $y$  with  $t$ .

<sup>27</sup>This is where I depart from Bacon and Russell [5]. They restrict instantiation to purely logical contexts, but run into trouble making seemingly unproblematic instantiations in non-logical contexts. For instance, one cannot straightforwardly infer  $P \vee Q = Q \vee P$  from  $\forall XY(X \vee Y = Y \vee X)$ . Instantiating  $X$  with  $P$  yields  $\forall Y(P \vee Y = Y \vee P)$ , but the result is no longer logical, and we thus cannot instantiate  $Y$  with  $Q$ . The notion of scope is a little tricky in higher-order logic with  $\lambda$ -terms, as we do not want to count  $X$  as in the scope of a non-logical operation in the context  $\wedge P X$  (even though  $P$  is non-logical), but we do want to count it as in the scope of the operator  $B$  in  $(\lambda Y Y) B X$ , as it is  $\beta$ -equivalent to  $BX$ . One must instead define the scope for things in  $\beta$ -normal form, which is the smallest set of terms that contain (i) the variables and constants, (ii)  $QM_1 \dots M_n$  (provided this is well-typed) whenever it contains  $M_1 \dots M_n$  and  $Q$  is a variable or constant, and (iii)  $\lambda x M$  whenever it contains  $M$ . The free variables in the scope of  $P$  in a  $\beta$ -normal form term  $M$ ,  $S(P, M)$ , may then be defined inductively:

1.  $S(P, Q) = \emptyset$  when  $Q$  is a variable or constant,
2.  $S(P, PM_1 \dots M_n) = \bigcup_i FV(M_i)$  when  $M_i$  are in  $\beta$ -normal form
3.  $S(P, QM_1 \dots M_n) = \bigcup_i S(P, M_i)$  when  $M_i$  are in  $\beta$ -normal form,  $Q$  a variable or constant,  $P \neq Q$
4.  $S(P, \lambda x M) = S(P, M) \setminus \{x\}$  when  $M$  is in  $\beta$ -normal form

here  $FV(M)$  denotes the free variables of  $M$ .

**Reflexivity**  $\forall x x = x$

**Existence**  $\exists x t = x$

From Logical Instantiation and identity we can derive the reflexivity schema,  $a = a$ . And from  $L(xy)$  and Logical Instantiation we can derive every instance of Leibniz's law in which the substituted terms do not occur in non-logical contexts. This allows one to derive the symmetry and transitivity schemas in the usual way:  $a = b \rightarrow b = a$  and  $a = b \rightarrow b = c \rightarrow a = c$ .<sup>28</sup> Existence corresponds to the idea that failures of Universal Instantiation are not due to non-existence. I will not adjudicate on the more general issue of Existence when there are empty names in the language. For now we may simply treat it as a stipulation that the language does not contain any empty names.

It is worth pausing, for a minute, to consider whether the existence of opacity *forces* one to reject Existential Generalization, and Unrestricted Export. Part of our argument for these failures rested on the assumption of  $L(xy)$ , which someone taking opacity at face value may wish to deny. A more theoretical argument for  $L(xy)$  is given in [5] from a higher-order generalization, where  $X$  is a variable taking predicate position:

**L(xyX)**  $\forall X \forall x \forall y (x = y \rightarrow Xx \rightarrow Xy)$ .

$L(xyX)$  can be justified as follows: if it were false, one could find  $x$  and  $y$  that are identical, but are distinguished by a property. In which case, there would be a relation that *did* guarantee that  $x$  and  $y$  shared the same properties, namely the relation of *sharing the same properties* (defined:  $\lambda xy \forall X (Xx \leftrightarrow Xy)$ ). This relation satisfies the logical role that identity is customarily thought to satisfy, given Universal Instantiation. For if  $a$  and  $b$  stand in this relation, then  $Ba \leftrightarrow Bb$ , instantiating  $X$  with  $B$ .<sup>29</sup> So this relation satisfies  $L()$ , whereas identity proper does not, according to this view, making conspicuous the charge that one is simply talking about the wrong relation. If there is a relation that satisfies the logical role of identity, then it is hard to see any other relation having a better claim to being the notion of identity appropriate to logic and metaphysics. Of course, some metaphysicians have postulated notions of 'loose identity' by which we sometimes count and individuate objects in natural language, whilst distinguishing it from the proper identity of logic and metaphysics (see, e.g., Chisholm [10] and Lewis [24]). But to object to the logicians use of Leibniz's law on these grounds seems like an overreach.<sup>30</sup>

<sup>28</sup> $L(xy)$  allows us to derive the quantified versions of these, but without Universal Instantiation we can derive the instances.

<sup>29</sup>This ensures that  $L()$  holds for direct predications. One can show that  $L()$  holds for an arbitrary context  $A$ , in higher-order logic, by using the device of  $\lambda$  abstraction, which can turn an arbitrary context into a predication. Specifically, one must appeal to the principle  $(\lambda x A)t = A[t/x]$ , which states that substituting a (substitutable) term  $t$  into an arbitrary context,  $A$  (i.e. a formula with one free variable) is equivalent to a direct predication. Philosophers who have rejected this principle tend to insist that it was only the instances of  $L()$  involving direct predications that we should have cared about in the first place. My remarks to follow about the logical role of identity, then, should be acceptable to these philosophers as well.

<sup>30</sup>Caie et al. [9], for instance, accept the existence of a strict identity relation that satisfies

With  $L(xyX)$  so justified we may infer  $L(xy)$  by universally instantiating  $X$  with  $A$  in the logical context  $\lambda X \forall x \forall y (x = y \rightarrow Xx \rightarrow Xy)$ .<sup>31</sup> The critical observation here is that, while Universal Instantiation is not in general valid, one can make such instantiations in logical contexts. For while words like ‘belief’, ‘hope’, ‘says’ and so forth create opaque contexts, words like ‘not’, ‘and’ and ‘all’ appear not to. A similar justification may be made of another quantified version of Leibniz’s law, namely:<sup>32</sup>

$$\mathbf{L(X)} \quad \forall X (a = b \rightarrow Xa \rightarrow Xb)$$

since one can also infer it from  $L(xyX)$  by instantiation into a logical context.

Caie et al. [9] follow [5] in rejecting the orthodoxy regarding Leibniz’s law, but develop a classical account of opacity that keeps Universal Instantiation. Consequently  $L(xyX)$ ,  $L(xy)$  and  $L(X)$  are all rejected along with  $L()$ .<sup>33</sup> But in addition to the worries above, these sorts of views are ontologically wild in ways that the non-classical versions are not. For instance, in order to accommodate 1-3, there exists a property that Hesperus has which Phosphorus doesn’t (namely, being believed by Simon to be bright). Call the relation of sharing the same properties, strict identity, and its negation strict distinctness: thus Hesperus and Phosphorus are strictly distinct. Just as we talk about there being multiple planets satisfying some criteria when there exist a distinct  $x$  and  $y$  that are planets with the criteria, we say there are strictly multiple planets when there are strictly distinct planets,  $x$  and  $y$ , satisfying the criteria. We will follow similar conventions for words like ‘lots’, ‘several’, ‘more’, ‘most’ and so on. Since Hesperus and Phosphorus are strictly distinct, there are at least two planets, in the strict sense, colocated with Venus. Of course, there’s nothing special about the name ‘Venus’ either, so we should expect there to be strictly more than the strict two: there should be a ‘strict lot’ of them, one for each possible

---

$L()$ , but respond to our argument on the grounds that basic judgments about English seem to directly support the idea that identity does not satisfy  $L()$ . But if there is indeed a relation that satisfies the logical role that identity is supposed to, then it seems to me the primary matter of substance has been settled. What remains is a question concerning which role takes priority, and whose occupant deserves the name ‘identity’: the logical role of satisfying  $L()$ , and another role more closely connected to judgments about identity in English. But this is less substantial than it might seem. Consider, for instance, the dynamic semanticist who takes the fact that ‘ $A$  and  $B$ ’ and ‘ $B$  and  $A$ ’ are not always equivalent in English to mean that the principle of commutativity of conjunction is not valid. They might concede that there is a connective satisfying the classical laws of conjunction, but that this connective is not really *conjunction*. I think classical logicians can acknowledge this, but go about their usual business unabated, so long as they are explicit that it is the classical connective that is the target of their logical theorizing.

<sup>31</sup> Again, the assumption made at the outset that arbitrary contexts are equivalent to direct predications.

<sup>32</sup> There is a complication in deriving this second principle: after instantiating  $x$  with  $a$ , the context is no longer logical since it contains  $a$ . Thus second instantiation of  $y$  for  $b$  is not in a purely logical context. This can be circumvented as in [5] by a slightly more complicated argument. However, in the present context, it is straightforwardly derivable given Logical Instantiation, since although the context is no longer logical, it is a context in which  $y$  doesn’t appear in a non-logical context.

<sup>33</sup> Clearly, any of  $L(xyX)$ ,  $L(xy)$  and  $L(X)$  can be used to derive  $L()$  given Universal Instantiation.



identity confusion.<sup>34</sup> So there is a sense in which the view inflates the ontology, even if all these strictly distinct planets are all identical in the loose sense. But there is also an internal worry. Clearly Hesperus is a planet between Mercury and Earth, similarly for Phosphorus. Indeed, given their identity, they must share many properties — all the *transparent* properties, including the property of being colocated with Venus, being visible in the evening and being visible in the morning, orbiting the Sun, and presumably any properties definable in the language of physics. But why is it, then, that Simon has the belief that Hesperus is bright, but not Phosphorus? After all, there are two strictly distinct planets colocated with one another, sharing all the same physical properties: it's not like one can be invisible in the evening while the other isn't — they occupy the same region of space-time, have the same reflective properties, and so on. What is it about Hesperus, and not the physically indistinguishable Phosphorus, that allows Simon to attach his belief to it, and not a strictly different planet?<sup>35</sup>

The view raises other metaphysical concerns. For instance, many philosophers are attracted to physicalism, according to which every property can be defined from physical properties. The strongest version of this thesis is that every property is strictly identical to a logical combination of physical properties. But all physical and logical properties are plausibly transparent, as, plausibly, are things you can create by combining logical and physical operations. In which case belief must be strictly identical to a transparent property, and thus transparent itself (since strict identity obeys Leibniz's law).

The non-classical opacity theorist, by contrast, avoids the excesses of its classical cousin. It draws no distinction between identity and strict identity: indeed, according to it there are no relations that satisfy Leibniz's law. They can accept that belief is strictly identical to a transparent property, without inferring that belief is transparent, since *being transparent* is itself an opaque property.<sup>36</sup> And they may deny that there are two strictly distinct colocated planets.

---

<sup>34</sup>Presumably one for every identity confusion formulable in some possible language. Alternatively, one might think that there are only as many strictly distinct planets as actual identity confusions, but this either means that the strict number of planets depends on the practices language speakers in surprising ways, or else means there are surprising coincidences between these practices and the strict number of planets.

<sup>35</sup>Might one attempt to turn this into an argument against opacity *tout court*? How can Simon believe that Hesperus is bright, but not Phosphorus, while both Hesperus and Phosphorus are plainly visible to him and bright (irrespective of whether they are strictly identical or not)? The initial reason for believing in opacity is that one can fail to believe that Phosphorus is bright, even when it is plainly visible to you, because believing that Phosphorus is bright is more demanding. You must also think about Phosphorus in the right way — presumably in a different way than is required to believe that Hesperus is bright. None of this is explained by the fact that there's a thing out there in the world, Phosphorus, which you have failed to believe to be bright. The thought that opacity is all in the head and not in the world is born out in the theory of Bacon and Russell: there simply isn't a plainly visible planet which Simon fails to believe to be bright on the basis of his observations. By contrast, Caie et al. maintain that there is. According to them, there are two (identical, but strictly distinct) planets, which are both plainly visible to Simon, but of which he believes one to be bright, and not the other.

<sup>36</sup>See the discussion in Bacon and Russell [5] at the end of section 4.

## 5 Prior's Paradox and Opacity

We have argued that opaque contexts create a distinctive cluster of counterexamples to Existential Generalization (and Universal Instantiation). We have focused on the case of first-order quantification, for the sake of familiarity, but the considerations generalize to other semantic types. For instance, it seems plausible that the property of being a lawyer, and the property of being an attorney are the very same, but Simon might believe that Susan is a lawyer, without believing that she is an attorney. Similar considerations compel us to reject Existential Generalization for quantification into predicate position.

Of special interest is the case of quantification into sentence position. A solution to Prior's paradox along these lines, would accept the following (with similar things being said about the variants of Prior's paradox stated in terms of hope, fear, and so forth):

1. Simon said, at  $t$ , that nothing he said at  $t$  is true.
2. For no  $P$  did Simon say that  $P$  at  $t$ .
3. For some  $P$ , the proposition that  $P$  just is the proposition that nothing Simon said at  $t$  is true.

It is 3, in particular, that distinguishes this solution from the solutions based on non-existence, and bans on impredicativity.

But despite this difference, many of the goodmaking features of the non-existence approach apply here too. For instance, unlike the classical no-proposition view, there is a point to uttering sentences that do not result in you saying anything. For one can say that Simon didn't say anything true at  $t$ , even if one hasn't thereby said anything, .

According to the non-existence view, the reason Simon hasn't said anything with his utterance is that the proposition that nothing Simon said at  $t$  is true simply doesn't exist. Thus no one can say anything by uttering the sentence 'nothing Simon said at  $t$  is true' (even though one will nonetheless say that nothing Simon said at  $t$  is true). This non-existence makes general theorizing difficult. For instance, the symmetry of disjunction is sometimes captured with a generalization:

$$\forall XY \ X \vee Y = Y \vee X$$

But according to the view under consideration, this is insufficiently general. It will not entail the instance in which  $X$  is instantiated with 'nothing Simon said at  $t$  is true' and  $Y$  with 'snow is white'.

By contrast, even though we have established that there is no  $P$  such that Simon said at  $t$  that  $P$ , we have not established that someone else couldn't say that nothing Simon said  $t$  is true and thereby have said something. Nor have we established that Simon couldn't say something with the same utterance at another time. By analogy with the first-order case, while there's no-one who I know to have committed the Whitechapel murders, there might be someone

else for which there is someone they know to have committed the Whitechapel murders (Jack the Ripper himself, for instance). Whereas, there can't be anything that that is believed (by *anybody*) to be a winged horse-god on the basis of believing that Pegasus is a winged horse-god.

Similarly, our generalization above is sufficient to prove all its instances, given the sentential analogue of Logical Instantiation. Since we can instantiate  $X$  with  $P$  into a purely logical context to get  $\forall Y P \vee Y = Y \vee P$ . And we may then instantiate again to get  $P \vee Q = Q \vee P$ , as  $Y$  was not in the scope of a non-logical operation.

Of course, not all theorizing is insulated from opacity. For instance, Bacon et al. [4] object that the universal generalization  $\forall X (KX \rightarrow X)$ , expressing the factivity of knowledge, is not general enough. An appeal to Logical Instantiation is of no avail here, since  $X$  is in the scope of an intentional operator,  $K$ . One must instead be content with a schema:  $KA \rightarrow A$ . But there is an extent to which schemas are unavoidable. For instance, in classical propositionally quantified logic, to ensure the that the quantified claim  $\forall X (KX \rightarrow X)$  really does imply all of its instances, one needs a different schema, namely Universal Instantiation,  $\forall X A \rightarrow A[B/X]$ . It cannot be substituted with its universally quantified variant,  $\forall Y (\forall X A \rightarrow A[Y/X])$ , what we earlier called Quantified Instantiation, because it doesn't imply its instances. See footnote 26.

Another issue raised in [4] is that in some of the models of Prior's paradox considered there (motivated by the free logic approach to empty names, and not opacity), the domain of existing propositions weren't closed under certain logical operations. In some models negations, conjunctions and disjunctions of existing propositions didn't exist. In the most plausible models, propositions were closed under the finitary Boolean operations, but you would either have a proposition  $P$  that existed, without the proposition that Simon said that  $P$  existing, or you would have a context such that the proposition defined by  $A[P]$  existed for each  $P$ , but  $\forall X A$  didn't. These limitations were no accident, as one could show in that context that the following closure principles would lead to contradiction:

$$\mathbf{CL}\top \exists p(p = \top)$$

$$\mathbf{CL}\perp \exists p(p = \perp)$$

$$\mathbf{CL}\neg \forall p \exists q(q = \neg p)$$

$$\mathbf{CL}\Box \forall p \exists q(q = \Box p)$$

$$\mathbf{CL}Q \forall p \exists q(q = Qp)$$

$$\mathbf{CL}\wedge \forall p \forall q \exists r(r = (p \wedge q))$$

$$\mathbf{CL} = \forall p \forall q \exists r(r = (p = q))$$

$$\mathbf{CL}\forall \forall p \forall q \exists r(q = \phi \rightarrow (r = \forall p \phi))$$

That derivation made essential use, however, of Unrestricted Export:  $(\exists X X = B \wedge A[B/X]) \rightarrow \exists X A$ . The informal idea was to show, on the basis of these closure conditions, that Prior's proposition,  $\forall X(QX \rightarrow \neg X)$  existed. Unrestricted Export tells us we may universally instantiate and existentially generalize on propositions we know to exist — in effect, classical reasoning is legitimate provided the relevant propositions exist.<sup>37</sup>

But the logic of opacity, in propositionally quantified logic, should include analogues of all of the principles we discussed in section 4, including the schema Existence:

**Existence**  $\exists X(X = A)$

Thus each of the closure principles listed above is thus trivially secured. And as for the derivation in [4], Unrestricted Export must be relinquished. But inspection of the argument reveals that one must apply Unrestricted Export to variables under the scope of  $Q$ , a move we have rejected on independent grounds, from considerations of opacity.

So an opacity-inspired response to Prior's paradox is desirable. But is it consistent? That is to say, can we have a model of Possible Saying within the logic of opacity. For clarity, I will restate all the principles we wish to show to be jointly consistent, converting the first-order logic of opacity of section 4 to the context of propositionally quantified logic:

**Possible Saying!**  $\Diamond Q!A$

**Subsumption**  $\Box(Q!A \rightarrow QA)$

**Uniqueness**  $\Box(Q!A \wedge Q!B \rightarrow A = B)$

**Quantified Instantiation**  $\forall X(\forall Y A \rightarrow A[X/Y])$

**Normality**  $\forall X(A \rightarrow B) \rightarrow \forall X A \rightarrow \forall X B$

**Quantifier Exchange**  $\forall X \forall Y A \rightarrow \forall Y \forall X A$

**Vacuous Quantification**  $A \rightarrow \forall X A$  when  $X$  does not occur free in  $A$

**Logical Instantiation**  $\forall Y A \rightarrow A[B/Y]$  provided  $y$  is not in the scope of a non-logical operation (i.e.  $Q$  or  $Q!$ ).

**Reflexivity**  $\forall X X = X$

**L(xy)**  $\forall XY(X = Y \rightarrow A \rightarrow A[Y/X])$

**Existence**  $\exists X(X = A)$

---

<sup>37</sup>Actually [4] do not appeal explicitly to Unrestricted Export — rather they are implicitly assuming  $L()$  which allows them to derive it from Quantified Instantiation. Omitting the assumption of  $L()$  seems like an oversight.

For the sake of simplicity, we shall assume that propositions are individuated by necessary equivalence. Thus we may alternatively take  $\Box$  as primitive, and defined  $A = B$  by  $\Box(A \leftrightarrow B)$ , or take  $=$  as primitive, and define  $\Box A$  as  $A = \top$ . I will take the latter option, as it strikes me that identity is the more basic notion. Our language will thus be that of propositionally quantified logic, with a binary connective,  $=$ , and two unary connectives  $Q$  and  $Q!$ .

On this last point, one thing that should be emphasized is that our approach to opacity, despite capitulating much to the thought that attitudes are capable of drawing fine distinctions, does not take sides on propositional granularity. Indeed, everything we have said so far is consistent with *Booleanism*, the thesis that propositions are governed by the Boolean identity, as encoded by generalizations, statable in our propositionally quantified language with identity, like the commutativity of disjunction discussed earlier.

**Theorem 1.** *The listed principles are consistent, and indeed have a possible worlds model securing Booleanism.*

The full proof of this may be found in the appendix. But here is an informal explanation of the construction. Firstly, following [4], a model will consist of:

- A set of worlds,  $W$
- An accessibility relation,  $R$ , on  $W$ , for interpreting  $=$  (and consequently  $\Box$ ). It will be an equivalence relation.
- An mapping,  $|Q| : W \rightarrow P(P(W))$ , that provides the extension of  $Q$  at each world  $w$ .  $|Q|(w)$  is the set of propositions  $Q$ ed at  $w$ .
- A mapping  $D : W \rightarrow P(P(W))$  where  $D(w)$  represents the domain of the propositional quantifiers at  $w$ .

A sentence of the form  $A = B$  is true at a world  $w$  iff  $A$  and  $B$  are true at the same worlds accessible to  $w$ . Provided  $R$  is not the universal relation, an identity like this can be true even when  $A$  and  $B$  are true at different worlds.  $QA$  is true at a world  $w$  if the set of worlds where  $A$  is true is in the extension of  $Q$  at  $w$  (i.e. in  $|Q|(w)$ ).  $Q^+A$  is true at  $w$  if, moreover, that set is the unique member of  $|Q|(w)$ . The remaining clauses are the obvious ones, and may be found in the appendix.

In our specific model,  $W$  will consist of two copies of the natural numbers, where worlds in the first copy see only worlds in the first copy (along the accessibility relation  $R$ ), and similarly, worlds in the second copy see only worlds in the second copy.

The domain of quantification in a world in the first copy will consist of finite and cofinite sets of worlds in the first copy. Likewise the domain of a world in the second copy will consist of finite and cofinite sets of worlds from the second copy. (Cofiniteness is being computed relative to the copy in question, not relative the whole set of worlds.) To interpret  $Q$  we fix a bijection  $\sigma$  between the naturals,  $\mathbb{N}$ , and the finite and cofinite subsets of  $\mathbb{N}$ . The interpretation of

$Q$  at the world  $n$ , in either copy, is the same. It contains exactly one set of worlds — the union of the two copies of  $\sigma(n)$ . Note that this proposition is not in the domain of any world, either in the first or the second copy. But it is nonetheless necessarily equivalent to a proposition that is in the domain, in either case, for it coincides with a finite or cofinite set from the first (or second) copy respectively on the accessible worlds at that copy. Given the coincidence of identity with necessary equivalence in this model, this secures Existence.

Say that a set of worlds is symmetric if it is the same in both copies. Given the symmetry of the way the semantics is set up, a world in the first copy is indistinguishable from its sibling in the second copy, and vice versa. Thus one ought to expect a closed sentence could only be true at a world in the first copy if it was also true in its sibling in the second copy, and vice versa. Thus every closed sentence expresses a symmetric proposition. Moreover, if things work out nicely, every closed sentence should express a symmetric proposition that is finite or cofinite relative to either (and thus both) copies. In which case, for any closed sentence  $A$ , the proposition it expresses in the model is the unique proposition in the extension of  $Q$  at the two sibling worlds the bijection  $\sigma$  associates to that proposition. Thus,  $\Diamond Q!A$  will be true at any world in the first or second copy. The tricky part of this argument is establishing that every closed sentence expresses a finite or cofinite set, relative to both copies; the full details are in the appendix.

## 6 Truth and Meaning

The liar paradox poses familiar problems for many attempts at systematic semantic theorizing. Could these considerations on intentionality and opacity shed any light on it? In this section I will attempt to leverage our remarks on opacity and Prior’s paradox to motivate a novel theory of truth and a diagnosis of the liar paradox.

While much work on the liar has focused on the notion of truth, there is a large amount of work in semantics that is best thought of as characterizing the notion of a sentence meaning a proposition. According to this paradigm, lexical items more generally are assigned meanings, and the meanings of complex phrases are computed in terms of the meanings of their constituents, in accordance with the principle of compositionality. The meaning of a sentence is a proposition. The notion of sentential truth is then a derivative notion, to be understood as meaning a true proposition. We might formalize this by introducing a device,  $M$ , formalizing ‘means that’, that combines with the name of a sentence and a sentence to form a sentence, just as verbs like ‘says that’, ‘hopes that’ or ‘fears that’ combine with the name of a person and a sentence to form a sentence.

We can then introduce truth, in propositionally quantified first-order logic, via a definition:

**Truth**  $Tx := \exists Z (M x Z \wedge Z)$

**Falsehood**  $Fx := \exists Z (M x Z \wedge \neg Z)$

Suppose, moreover, that for each sentence of the language in question,  $A$ , there is a name  $\langle A \rangle$ , that on its intended interpretation denotes the sentence  $A$ . A *prima facie* compelling disquotational principle says:

**The Meaning Schema** The sentence ‘ $P$ ’ means that  $P$

which we may formalize:

$$M \langle A \rangle A$$

But now suppose that there is a name,  $c$ , identical to the representative name for the sentence,  $\forall X (M c X \rightarrow \neg X)$ , i.e.  $c = \langle \forall X (M c X \rightarrow \neg X) \rangle$ . An instance of the Meaning Schema is:

$$M \langle \forall X (M c X \rightarrow \neg X) \rangle \forall X (M c X \rightarrow \neg X)$$

Replacing  $\langle \forall X (M c X \rightarrow \neg X) \rangle$  for  $c$ , we get

$$M c \forall X (M c X \rightarrow \neg X)$$

Now interpret  $Q X$  in Prior’s theorem  $M c X$ . Applying Modus Ponens to Prior’s theorem we get:

$$\exists X (M c X \wedge X) \wedge \exists X (M c X \wedge \neg X)$$

In other words  $c$  means a truth and a falsehood. According to our definitions this means  $c$  is both true and false. Note that this is not strictly speaking inconsistent. But it contradicts the assumption, seemingly just as important for doing semantics compositionally, that a sentence means at most one thing. For if  $A$  meant several propositions, and  $B$  meant several propositions, how would we compute the meaning of  $A \vee B$ ? The obvious choice of letting it mean any proposition you can get by disjoining something  $A$  means with something  $B$  means doesn’t work, for otherwise  $A \vee \neg A$  can mean falsehoods. We need some way of coordinating the meanings of the two disjuncts, and we can’t do this with  $M$  as our only primitive.<sup>38</sup>

What about the more traditional liar sentence? Suppose we had a name,  $c$  such that  $c = \langle \neg Tc \rangle$ . Observe that expanding out the definition of  $\neg Tc$  you get  $\neg \exists X (M c X \wedge X)$ , and applying some logic this is equivalent to  $\forall X (M c X \rightarrow \neg X)$ . Thus there is a straightforward sense in which our paradox considered above is just another version of the liar, modulo our definition of truth.

How might opacity shed light on this paradox? I have suggested in the previous section that ‘says that’ is opaque. What things sentences mean supervenes in large part, on what propositions people have said. One could take this to suggest that ‘means’ is opaque. According to one theory, to say that  $P$  is just to utter a sentence that means that  $P$ . If ‘means’ was not opaque, then neither

---

<sup>38</sup>Cian Dorr is developing a view with this form, but the coordination problem is solved by having more complicated primitives.

would ‘says’ be. But this does not seem like a plausible account of saying to me. Merely uttering a sentence is not usually sufficient for saying anything. And it seems plausible that one can say things using *ad hoc* conventions, as when one is in a foreign country and doesn’t speak the language. Indeed, the fact that one can say things in the absence of a language suggests a very appealing theory of how sentences can acquire meanings in the first place: that one uses certain sentences when saying things to establish the more general convention that the sentence be used to say that thing.<sup>39</sup> On this view, *saying* is a primitive propositional attitude, and meaning is to be explained in terms of it. In this case too, we should expect meaning to be opaque, for it is explained in terms of an opaque attitude. This appears to be no accident: other reductive theories of meaning reduce meaning to other intentional attitudes. For example, Lewis, in his book *Convention* [23], attempts to reduce meaning to the desires and beliefs of the language speakers. Davidson’s account of radical interpretation involves the attitude of belief in a similarly central way [12].

Even if we do not make this reductive move, there are independent reasons for thinking that meaning is opaque. In Bacon and Russell [5] we are, among other things, concerned with upholding the Millian theory of names. According to this theory, the meanings of ‘Hesperus’ and ‘Phosphorus’ are the same: the planet Venus. We might formalize this  $\llbracket \langle a \rangle \rrbracket = \llbracket \langle b \rangle \rrbracket$ . But a principle of compositionality then appears to entail that the meaning of ‘Hesperus is believed to be bright’ is the same as the meaning of ‘Phosphorus is believed to be bright’, i.e. that  $\llbracket \langle Ba \rangle \rrbracket = \llbracket \langle Bb \rangle \rrbracket$ . For the principle of compositionality says:  $\llbracket \langle Ba \rangle \rrbracket = \llbracket \langle B \rangle \rrbracket \llbracket \langle a \rangle \rrbracket$ , and  $\llbracket \langle Bb \rangle \rrbracket = \llbracket \langle B \rangle \rrbracket \llbracket \langle b \rangle \rrbracket$ . Leibniz’s law allows us to derive the undesired conclusion. Undesired, because these sentences appeared to have different truth values in ancient times, and so, presumably different semantic values. We draw the following moral: opacity exists in the metalanguage — meaning is opaque.

Many consistent theories of truth have been proposed that avoid the liar paradox: the logical landscape is pretty well-explored at this point.<sup>40</sup> However, much work is still needed in the interpretation of these theories. In the following I want to motivate a theory of truth that falls under the umbrella of the so-called *classical gap* theories of truth, of which prominent proponents include Feferman [13] and Maudlin [25]. These theories say of the liar sentence,  $c = \langle \neg Tc \rangle$ , that it is neither true nor false. But many have argued that such theories are self-undermining. For in virtue of the liar being neither true nor false, it follows that the liar isn’t true. But this is exactly the liar sentence. Such theorists have to make seemingly absurd speeches like:

1.  $c$  is not true.
2. The sentence I just uttered, 1, is not true.

Scharp [36], for instance, calls this the ‘self-refutation problem’. Priest puts

---

<sup>39</sup>This account is obviously extremely crude as it stands, but see Schiffer [37] for more elaborate theory along these lines.

<sup>40</sup>For an early classification of the options in classical logic, see Friedman and Sheard [16].



it forcefully as follows: ‘truth is the aim of assertion. Once this connection is broken, the notion of assertion comes free from its mooring, and it is not clear why we should assert anything.’ [29].

Thus, we must do something to explain what we are doing when we assert, if the aim is no longer to assert truths. I think an opacity friendly theory of meaning can meet this demand.

Let us rehearse what we must say about the liar paradox, drawing on our solution to Prior’s paradox in the logic of opacity.

- $c$  means that nothing  $c$  means is true.
- For no  $X$  is it the case that  $c$  mean that  $X$ .

Given our definition of truth, as meaning a truth, it follows that  $c$  is not true: there is no proposition it means, and so no truth it means. Similarly, it is not false either, as there is no false proposition it means. Thus we have a version of the classical gap theory. In particular, we must make the seemingly absurd sequence of assertions listed in 1 and 2.

We began by suggesting that it is meaning, not truth, that is the more theoretically central notion. This opens up the option, not available to someone just theorizing in terms of truth, of giving a meaning-theoretic explanation of why we might assertively utter  $c$  — the sentence ‘ $c$  is not true’ — even though it is not itself true. The explanation is this: because by uttering that sentence, *one thereby says that  $c$  is not true*. And by saying that  $c$  is not true, one may pass on the knowledge that  $c$  is not true to ones audience, and they can act accordingly. Of course, these sayings and knowledge are opacity laden: there is no proposition one has thereby said, or that ones audience thereby knows. But, as we argued in section 3, in the context of the non-existence approach, this is no barrier to the knowledge playing the relevant role in informing your actions.

## 7 Conclusion

We have argued that by taking opacity seriously one may provide a philosophically appealing solution to Prior’s paradox. The resulting account of intentionality naturally extends to an opacity-laden theory of meaning that vindicates a classical gap theory of truth that is able to meet the challenges usually directed at that account.

## 8 Appendix

We shall work in the language of propositionally quantified logic. The wffs consist of an infinite stock of propositional variables,  $X_1, X_2, \dots$ , a wff  $\neg A$ ,  $QA$ ,  $Q!A$ ,  $A \wedge B$ ,  $A = B$  and  $\forall X A$ , whenever  $A$  and  $B$  are wffs, and  $X$  a propositional variable.

Following [4], a model of this language consists of:

- A set  $W$
- An equivalence relation  $R$  on  $W$ .
- A map  $|Q| : W \rightarrow P(P(W))$
- A map  $D : W \rightarrow P(P(W))$  such that if  $D(w) \subseteq P(P(R(w)))$

Here we follow the usual convention of writing  $R(w)$  for  $\{x \in W \mid Rwx\}$ . A variable assignment is a function  $g : Var \rightarrow P(W)$ , assigning each propositional variable to a set of worlds. Write  $g \sim_X h$  iff  $g$  and  $h$  agree on every variable except, possibly,  $X$ . We may interpret an arbitrary wff relative to an assignment as follows:

$$\llbracket X \rrbracket^g = g(X)$$

$$\llbracket A \wedge B \rrbracket^g = \llbracket A \rrbracket^g \cap \llbracket B \rrbracket^g$$

$$\llbracket \neg A \rrbracket^g = W \setminus \llbracket A \rrbracket^g$$

$$\llbracket \forall X A \rrbracket^g = \{w \mid w \in \llbracket A \rrbracket^h \text{ for every } h \text{ such that } h(X) \in D(w) \text{ and } g \sim_X h\}$$

$$\llbracket QA \rrbracket^g = \{w \mid \llbracket A \rrbracket^g \in |Q|(w)\}$$

$$\llbracket Q!A \rrbracket^g = \{w \mid \llbracket A \rrbracket^g \text{ is the only member of } |Q|(w)\}$$

$$\llbracket A = B \rrbracket^g = \{w \mid \llbracket A \rrbracket^g \cap R(w) = \llbracket B \rrbracket^g \cap R(w)\}$$

This model theory is sound for the logic of opacity, except for Existence and Logical Instantiation: for any  $w \in W$  of such a model, and any closed theorem of the logic of opacity without Existence and Logical Instantiation,  $A$ ,  $w \in \llbracket A \rrbracket$ .<sup>41</sup> In order to validate both Existence and Logical Instantiation a model must satisfy the further constraint:

For any wff,  $A$ ,  $w \in W$  and assignment  $g$  whose range contained in  $D(w)$ ,  $\llbracket A \rrbracket^g \cap R(w) \in D(w)$ .

I'll now spell out the model described in section 5 in a little more precision. Let  $\sigma$  be a bijection between  $\mathbb{N}$  and the finite/cofinite subsets of  $\mathbb{N}$ . We define a model as follows:

- $W = \mathbb{N} \times 2$
- $R = (\mathbb{N} \times \{0\})^2 \cup (\mathbb{N} \times \{1\})^2$
- $|Q|_{(n,i)} = \{\sigma(n) \times 2\}$
- $D((n,i)) = \mathcal{P}_{f/cof}(\mathbb{N} \times \{i\})$

Here are some useful definitions. In what follows  $X$  denotes a subset of  $W$

---

<sup>41</sup>One has to be a little more careful with open theorems, by restricting attention to assignments that assign variables to propositions in the domain of the relevant world.

**Definition 1.** A left set is a finite or cofinite subset of  $\mathbb{N} \times \{0\}$ , and a right set is a finite or cofinite subset of  $\mathbb{N} \times \{1\}$ . Denote the set of left sets  $L$ , and the set of sets  $R$ .

Note: a left (right) set must be finite or cofinite *relative to*  $\mathbb{N} \times \{0\}$  ( $\mathbb{N} \times \{1\}$ ). A cofinite subset of  $\mathbb{N} \times \{0\}$  will not be cofinite relative to  $W$ .

**Definition 2.** The set  $X \subseteq W$  is left finite or cofinite iff  $X \cap \mathbb{N} \times \{0\}$  is a left set (it is right finite or cofinite iff  $X \cap \mathbb{N} \times \{1\}$  is a right set).

**Definition 3.** The set  $X$  is symmetric iff  $(n, 0) \in X \Leftrightarrow (n, 1) \in X$  for every  $n \in \mathbb{N}$ .

Recall that our conjecture was that all closed sentences express sets that are both (i) symmetric and (ii) finite or cofinite (in  $W$ ). We will call the set of subsets of  $W$  satisfying both (i) and (ii),  $\mathcal{S}$

**Definition 4.** A left assignment is a function  $g$  from propositional variables to left sets (a right assignment maps variables to right sets). Call the set of all left and right assignments  $A_L$  and  $A_R$  respectively.

Suppose that  $A$ s free variables lie in some finite set  $V$  of size  $n$ . Then  $\llbracket A \rrbracket : L^n \rightarrow \mathcal{P}(W)$  defines an  $n$ -ary function from left sets to subsets of  $W$  (and similarly defines a function from right sets).

**Definition 5.** An  $n$ -ary function  $f : L^n \rightarrow \mathcal{P}(W)$  is well behaved if and only if

- a.  $f$  only takes right finite or cofinite sets as values.
- b.  $\{f(\bar{x}) \cap \mathbb{N} \times \{1\} \mid \bar{x} \in L^n\}$  is finite. In other words, as we run through the set of possible values for  $\bar{x}$ , the right half of  $f(\bar{x})$  takes at most finitely many different values.

A function from  $f : R^n \rightarrow \mathcal{P}(W)$  is well behaved if the natural parallel conditions obtain.

Fix  $V$  ks be a finite set of variables, and  $A$  a formula whose variables all lie in  $V$ . I will write  $\llbracket A \rrbracket$ , without the assignment superscript, to denote the  $n$ -ary function on left sets, that takes  $P_1, \dots, P_n \in L$  to  $\llbracket A \rrbracket^{g[X_1 \mapsto P_1 \dots X_n \mapsto P_n]}$  (for any assignment  $g$ ).

**Theorem 2.** Let  $V$  be a finite set of variables, and  $A$  a formula whose variables all lie in  $V$ . Then  $\llbracket A \rrbracket$  defines an  $n$ -ary function on left sets that is well behaved

*Proof.* For  $X_i \in V$ ,  $\llbracket X_i \rrbracket$  is constant if we vary  $X_j$  for  $j \neq i$ . If we vary the value of  $X_i$  clearly condition (a) holds, and the only value  $\llbracket X_i \rrbracket \cap \mathbb{N} \times \{1\}$  takes on is  $\emptyset$ .

If  $\llbracket A \rrbracket$  and  $\llbracket B \rrbracket$  are well behaved, so are  $\llbracket \neg A \rrbracket$  and  $\llbracket A \wedge B \rrbracket$  (straightforward).

$\Box$  maps every intension to either  $\emptyset, \mathbb{N} \times \{0\}, \mathbb{N} \times \{1\}$  or  $W$  so  $\llbracket \Box A \rrbracket$  is clearly well behaved.

$Q$  and  $Q!$  behave the same way: they map every member of  $\mathcal{S}$  to a symmetric pair of the form  $\{(n, 0), (n, 1)\} \in \mathcal{S}$ , and maps every other set to  $\emptyset$ . Thus:

1.  $\llbracket QA \rrbracket$  takes only the emptyset and symmetric pairs as values, so it satisfies condition (a) for well-behavedness.
2. If  $\llbracket A \rrbracket \cap \mathbb{N} \times \{1\}$  takes only  $k$  many different values, then  $\llbracket A \rrbracket$  can take at most  $k$  different values in  $\mathcal{S}$ . So  $\llbracket QA \rrbracket$  can take on the emptyset and at most  $k$  different symmetric pairs as values, thus it satisfies condition (b).

Suppose  $\llbracket A \rrbracket$  is well-behaved. Note by assumption all  $A$ s variables lie in  $V$ , so if  $X_i \notin V$   $\llbracket \forall X_i A \rrbracket = \llbracket A \rrbracket$  which is well-behaved.

If  $X_i \in V$  then  $\llbracket \forall X_i A \rrbracket \cap \mathbb{N} \times \{1\}$  maps a tuple to a finite intersection of the different values of  $\llbracket A \rrbracket \cap \mathbb{N} \times \{1\}$  (we know  $\llbracket A \rrbracket \cap \mathbb{N} \times \{1\}$  takes on finitely many different values because it's well-behaved). So if the values of  $\llbracket A \rrbracket \cap \mathbb{N} \times \{1\}$  are all right finite/cofinite so are  $\llbracket \forall X_i A \rrbracket \cap \mathbb{N} \times \{1\}$ .

Note that if  $f(x_0, \dots, x_n)$  takes on finitely many different set values as we vary  $x_i$  the function  $h(x_1, \dots, x_n) = \bigcap_{x_0} f(x_0, \dots, x_n)$  can only take on finitely many different values. Thus  $\llbracket \forall X_i A \rrbracket$  satisfies condition (b) if  $\llbracket A \rrbracket$  does.  $\square$

**Theorem 3.** *The logic of opacity, Possible Saying!, Subsumption and Uniqueness are all true in this model at any world in the left half of  $W$ . (And by a symmetrical argument, also true at any world in the right half of  $W$ .)*

Here is a proof sketch.

*Proof.* Evidently every closed formula expresses a symmetric set (by the symmetry of the semantics). Closed sentences express constant functions from  $L^n \rightarrow \mathcal{P}(W)$ . By theorem 2, every closed sentence expresses (constantly) a right-finite/cofinite set.

Putting 1 and 2 together, closed sentences express sets in  $\mathcal{S}$ , and every member of  $\mathcal{S}$  is in the extension of  $Q$  and  $Q!$  at some world. This secures the validity of Possible Saying!

Also, since every closed formula expresses a left finite/cofinite set, every proposition expresses a proposition necessarily equivalent to a left set securing Existence and Logical Instantiation.  $\square$

## References

- [1] Andrew Bacon. Quantificational logic and empty names. *Philosophers' Imprint*, 13, 2013.
- [2] Andrew Bacon. The broadest necessity. *Journal of Philosophical Logic*, 47(5):733–783, 2018.
- [3] Andrew Bacon. Radical anti-disquotationalism. *Philosophical Perspectives*, 2019.
- [4] Andrew Bacon, John Hawthorne, and Gabriel Uzquiano. Higher-order free logic and the prior-kaplan paradox. *Canadian Journal of Philosophy*, 46(4-5):493–541, 2016.

- [5] Andrew Bacon and Jeffrey Sanford Russell. The logic of opacity. *Philosophical and Phenomenological Research*, forthcoming.
- [6] Christoph Benzmüller, Chad E. Brown, and Michael Kohlhase. Higher-order semantics and extensionality. *Journal of Symbolic Logic*, 69(4):1027–1088, 2004.
- [7] George Boolos. To be is to be a value of a variable (or to be some values of some variables). *Journal of Philosophy*, 81(8):430–449, 1984.
- [8] Tyler Burge. Semantical paradox. *Journal of Philosophy*, 76(4):169–198, 1979.
- [9] Michael Caie, Jeremy Goodman, and Harvey Lederman. Classical opacity. *Philosophical and Phenomenological Research*, forthcoming.
- [10] Roderick M. Chisholm. The loose and popular and the strict and philosophical senses of identity. In Norman S. Care and Robert H. Grimm, editors, *Perception and Personal Identity*, pages 82–106. Cleveland: Press of Case Western Reserve University, 1969.
- [11] Mark Crimmins and John Perry. The prince and the phone booth: Reporting puzzling beliefs. *Journal of Philosophy*, 86(12):685, 1989.
- [12] Donald Davidson. Radical interpretation. *Dialectica*, 27(1):314–328, 1973.
- [13] Solomon Feferman. Reflecting on incompleteness. *Journal of Symbolic Logic*, 56(1):1–49, 1991.
- [14] Kit Fine. The permutation principle in quantificational logic. *Journal of Philosophical Logic*, 12(1):33–37, 1983.
- [15] Gottlob Frege. On sense and reference. In Darragh Byrne and Max Kölbel, editors, *Arguing About Language*, pages 36–56. Routledge, 2010.
- [16] Harvey Friedman and Michael Sheard. An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33(1):1–21, 1987.
- [17] Michael Glanzberg. The liar in context. *Philosophical Studies*, 103(3):217–251, 2001.
- [18] Allen Hazen. Contra buridanum. *Canadian Journal of Philosophy*, 17(4):875–880, 1987.
- [19] David Kaplan. Quantifying in. *Synthese*, 19(1-2):178–214, 1968.
- [20] David Kaplan. A problem in possible worlds semantics. In Walter Sinnott-Armstrong and Diana Raffman and Nicholas Asher, editor, *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus*, pages 41–52. Cambridge University Press, 1995.

- [21] Saul A. Kripke. Russell’s notion of scope. *Mind*, 114(456):1005–1037, 2005.
- [22] Saul A. Kripke. A puzzle about time and thought. In Saul A. Kripke, editor, *Philosophical Troubles. Collected Papers Vol I*. Oxford University Press, 2011.
- [23] David Lewis. *Convention: A Philosophical Study*. Wiley-Blackwell, 1969.
- [24] David K. Lewis. Survival and identity. In Amelie Oksenberg Rorty, editor, *The Identities of Persons*, pages 17–40. University of California Press, 1976.
- [25] Tim Maudlin. *Truth and Paradox: Solving the Riddles*. Oxford University Press, 2004.
- [26] John Nolt. Free logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018.
- [27] Charles Parsons. The liar paradox. *Journal of Philosophical Logic*, 3(4):381–412, 1974.
- [28] Graham Priest. Intensional paradoxes. *Notre Dame Journal of Formal Logic*, 32(2):193–211, 1991.
- [29] Graham Priest. Truth and paradox. *Journal of Philosophy*, 102(9):483–486, 2005.
- [30] A. N. Prior. On a family of paradoxes. *Notre Dame Journal of Formal Logic*, 2(1):16–32, 1961.
- [31] A. N. Prior. *Objects of Thought*. Clarendon Press, 1971.
- [32] W. V. Quine. Quantifiers and propositional attitudes. *Journal of Philosophy*, 53(5):177–187, 1956.
- [33] Agustin Rayo and Stephen Yablo. Nominalism through de-nominalization. *Noûs*, 35(1):74–92, 2001.
- [34] Mark Sainsbury. *Reference Without Referents*. Clarendon Press, 2005.
- [35] Nathan Salmon. Lambda in sentences with designators. *Journal of Philosophy*, 107(9):445–468, 2010.
- [36] Kevin Scharp. *Replacing Truth*. Oxford University Press UK, 2013.
- [37] Stephen Schiffer. *Meaning*. Oxford, Clarendon Press, 1972.
- [38] Keith Simmons. *Universality and the Liar: An Essay on Truth and the Diagonal Argument*. Cambridge University Press, 1993.
- [39] Dustin Tucker and Richmond H. Thomason. Paradoxes of intensionality. *Review of Symbolic Logic*, 4(3):394–411, 2011.

- [40] Dustin Tucker. Paradoxes and restricted quantification: A nonhierarchical approach. *Thought: A Journal of Philosophy*, 7(3):190–199, 2018.
- [41] Timothy Williamson. Everything. *Philosophical Perspectives*, 17(1):415–465, 2003.
- [42] Elia Zardini. Truth and what is said. *Philosophical Perspectives*, 22(1):545–574, 2008.