Quantificational logic and empty names

Andrew Bacon

Current draft: 26/10/2011

Abstract

In a number of recent articles Timothy Williamson has built a strong case for the claim that everything necessarily exists. His argument rests on a combination of the derivability of this claim from quantificational logic, with standard modal principles and a robust understanding of the Kripke semantics for quantified modal logic. In this paper I defend a contingentist, non-Meinongian metaphysics within a positive free logic. It is argued that although certain names and free variables do not actually denote anything they might have actually done so, allowing one to interpret the contingentist claims without quantifying over non-existent possibilia.

The classical theory of quantification is subject to a number of difficulties relating to contingent existence and to the treatment of both empty and non-empty names. In this paper I propose and defend a modal metaphysics, couched in a weakening of classical logic, that avoids these objections.

It is well known that classical quantification theory does not provide a straightforward treatment of empty names. According to the simplest way of translating between English and first order logic there are false sentences of English which translate to theorems of first order logic. For a given first order language, $\mathcal{L}$, classical quantification theory proves every instance of the schema

$$\exists x t = x \tag{1}$$

where t here can be substituted for any term in $\mathcal{L}$. By a flat-footed translation of the false sentence, 'there is something identical to Pegasus', we obtain an instance of the above schema.[1] This is the problem of empty names.

Perhaps it is not first order logic that is to blame here, but rather our treatment of empty names. However matters become worse once we extend our logic and language to cope with modal reasoning. To achieve this one expands the language with an operator symbol, $\Box$, intended to be read informally as 'it is necessary that.' To capture some basic aspects of modal reasoning it is standard to include the rule of necessitation, that allows one to infer that $\Box\phi$ is a theorem if ϕ is a theorem. By the rule of necessitation we can obtain from any instance of (1) a corresponding instance of

$$\Box\exists x t = x \tag{2}$$

By this reasoning it can be seen that the false sentence, 'necessarily there is something identical to Timothy Williamson', is translated to a theorem of the system resulting from closing classical quantification theory under the rule of necessitation.[2] Unlike the former case, however, 'Timothy Williamson' actually refers to somebody. Whatever is responsible for our difficulties here cannot be attributed to a problematic treatment of empty names. The occurrence of a name is not even essential for generating the problematic phenomenon. If one has the rule of generalisation – a rule that allows you to infer the theoremhood of $\forall x\phi$ from the theoremhood of ϕ – one can obtain

$$\forall y\Box\exists x y = x \tag{3}$$

by applying the rule of generalisation to an instance of (2). This sentence corresponds informally to the claim that everything exists necessarily.

If one wants to avoid (2) one is faced with a clear choice: either reject classical quantification theory, or drop the rule of necessitation. As a matter of

[1]The claim that this sentence is false is far from uncontroversial (see Kripke [11], van Inwagen [27], Salmon [24].) In many of these disputes my point can be made with a different example, in the case of [24] for instance, by substituting 'Vulcan', a name for a hypothesised intra-Mercurial planet, for 'Pegasus'

[2]For views in which sentences like this are treated as actually true see Williamson [28], and Linsky and Zalta [17].

sociological fact, starting with Kripke's 1963 paper [10], modal logicians have blamed the derivability of (2) on first order logic and not the rule of necessitation.[3] The standard alternatives to the quantified modal logic based on classical quantification theory are based on logics that fall under the umbrella term 'free logics.' Thus we have some excellent reasons to want to weaken classical logic to some kind of free logic.

However, the philosophical foundations of free logic are not as solid as one would like. On the one hand there is no consensus about how a semantics should be given for a first order language containing empty names. It is generally thought that the standard model theories for this logic (such as the Meinongian and supervaluational theories - see [16]) cannot provide a semantics, as a special case, without having to illegitimately quantify over objects which, by the free semanticists own lights, do not exist.

In this paper a philosophical basis for the free logical approach to terms is given. I provide some linguistic evidence for a 'positive' free logic in which some atomic sentences involving non-referring terms are true. According to this view, there are a number of properties which Pegasus can have which do not force him to exist – being a mythical horse-god, being worshipped by the ancient Greeks, being believed to have wings, and so on. I argue that being referred to by the name 'Pegasus' is one of these properties and provide a causal account of reference which vindicates this. The use of positive free logic allows us to develop, in a free metalogic, an essentially Tarskian semantics for a language containing empty names. This semantics relies on non-denoting terms having a non-trivial semantic profile.

In §1 and §2 I outline and argue for a positive free logic. In §3 I defend the view that for every meaningful non-denoting term there could have been something to which that term actually refers. It is then shown how, in virtue of having non-trivial referential properties, meaningful empty names can contribute to the truth conditions of atomic sentences involving them. In §4 a semantics is given for positive free logic which does not involve quantifying over non-existent objects. It is crucial to this account that, unlike previously proposed semantics for free logic, the logic of the metalanguage is also free.

1 A weaker logic

Classical quantificational theory, construed so as to include identity axioms, contains every instance of (1) as a theorem. It can easily be verified that a derivation of (1) makes use of the classical axiom of universal instantiation: $\forall x\phi \to \phi[t/x]$ where t is free for x in ϕ. This is in fact the only culprit if one

[3] It is sometimes claimed that Kripke's response to this issue is to relinquish the rule of necessitation (for an example of this claim see [8].) However this description of his solution is, I think, highly contentious. Indeed, Kripke asserts (p69), and it is easily verified, that the unrestricted rule of necessitation is a *derived* rule in his system. In particular one does not need to restrict the rule of necessitation to closed formulae (as claimed in [8]) since there are no open theorems. On the other hand Kripke's system is a free logic since it does not contain universal instantiation.

wants to keep classical propositional logic, the principle of self-identity, and the standard equivalence between the universal and existential quantifiers.

In 1963 two papers appeared, one by Saul Kripke and one by Karel Lambert, that suggested weakening the axiom of universal instantiation to $\forall y(\forall x\phi \to \phi[y/x])$.[4] Kripke was mainly interested with the issue of contingent existence in modal logic whereas Lambert was concerned with the treatment of non-denoting singular terms, however it is arguable that both these issues stem from the same underlying problem. At any rate, it is extremely natural to wonder what happens if the axiom of universal instantiation is replaced by its weakening in an axiomatisation of classical first order logic with identity. Taking a standard axiomatisation of classical logic the result is the following system, where t and s can be substituted for arbitrary terms

- A1 Any substitution instance of a propositional tautology.
- A2 $\forall y(\forall x\phi \to \phi[y/x])$
- A3 $\forall x(\phi \to \psi) \to (\forall x\phi \to \forall x\psi)$.
- I1 $t = t$
- I2 $t = s \to (\phi \to \phi[t/s])$.
- MP From ϕ and $(\phi \to \psi)$ infer ψ.
- GEN From $(\phi \to \psi)$ infer $(\phi \to \forall x\psi)$ if x is not free in ϕ.

The rule GEN is sometimes called the Ackermann-Hilbert rule of generalisation. In some presentations of classical logic GEN is replaced with the simpler rule GEN2 – from ϕ infer $\forall x\phi$ – and an extra axiom for dealing with vacuous quantification – $\phi \to \forall x\phi$ where x is not free in ϕ. Substituting universal instantiation for (A2) in these variations will also result in an equivalent variation of free logic. We define what it means for a set of sentences, Γ, to prove ϕ, written $\Gamma \vdash \phi$, in the normal way (see Bridges [3] p41-42.)[5] In the classical variant one can prove the formula $\exists xx = x$ from (I1) and and the principle of universal instantiation, whereas in the above system it is no longer a theorem. On the face of it this is a benefit of the proposed system, since the sentence $\exists xx = x$ corresponds to the claim that there is at least one thing, and this claim does not appear to have the status of a logical truth. However one could retain this classical aspect by extending the system with the principle $\forall x\phi \to \exists x\phi$, allowing us to prove $\exists xx = x$ without running the risk of reinstating theorems like (1).

[4]Note that this *is* a strict weakening: in classical logic $\forall y(\forall x\phi \to \phi[y/x])$ entails universal instantiation, $\forall x\phi \to \phi[t/x]$. However that argument is itself an instance of universal instantiation, and cannot be carried out in the weakened system.

[5]One thing to note about Bridge's definition of $\vdash$ is that the application of (GEN) from a non-empty premise set is restricted. See Bridges [3] p42. Some presentations prefer not to restrict (GEN) and have instead a restricted soundness proof that applies only to closed formulae (see, for example, Mendelson [18] p69-70.) These complications are inherited from classical logic and need not worry us here.

The system described above is essentially that given by Lambert in [14], although I have presented it in terms of axiom schemas instead of with axioms and a rule of substitution (for reasons that will become clear in a moment.) The system is not conservative over its identity free fragment (see Fine [6]) so it is essential to keep the identity axioms. With that caveat, the system is as simple as classical logic.

A model theory that characterises these axioms can be given. A Meinongian model[6] is simply a pair $\langle D, I\rangle$ where:

- D is a set and I is a function on the non-logical vocabulary.
- $I(P_i^n)$ is a set of n-tuples for each n-ary predicate letter.
- $I(c_i)$ is an object for each constant symbol c_i.

A variable assignment is a function whose domain is the set of variables. Truth in a model is defined in exactly the same way as it would be for a first order model with the exception of the clause for the universal quantifier. Here one says that a quantified formula, $\forall x\phi$, is true with respect to an assignment v iff it is true with respect to every x-variant of v *which assigns x a member of D* - a restriction that is redundant in the classical model theory where assignments are defined as having codomain D. We say that $\Gamma \models \phi$ iff any model and assignment which makes every member of Γ true makes ϕ true. The system described above is sound and complete with respect to this model theory in the sense that $\Gamma \models \phi$ if and only if $\Gamma \vdash \phi$.[7]

Note that if this had been a definition of a standard model for first order logic the only difference would have been that $I(P_i^n)$ would be a set of n-tuples *from the domain D*, that $I(c_i)$ would be a member of D, and so on and so forth. Standard models are thus a special case of Meinongian models, which in turn means that the set of logical truths of free logic is a subset of the logical truths over standard models. Free logic is just a weakening of standard quantificational logic, and Meinongian model theory is just a generalisation of standard model theory.

Notice also the similarity between our models and Kripke's variable domain semantics for quantified modal logic [10]. Relative to a world a Kripke model

[6] I have modified the standard definition of a Meinongian model (see [13]) in some inconsequential ways however I shall continue to use the standard terminology.

[7] I have not seen this result explicitly proved for Lambert's original system, only for slightly less natural systems either without any well formed open formulae or ad hoc restrictions on modus ponens ([19], [16].) It is these presentations of Lambert's system that could be responsible for the impression that free logic is not a natural system. At any rate, Lambert's original 1963 system is quite natural and is not subject to any of these restrictions. It's soundness and completeness is a consequence of an observation due to Church [4]. Proof sketch: Let $\mathcal{L}$ be a first order language and $\mathcal{L}^+$ be the language augmented with a single unary predicate P. One can translate every formula of $\mathcal{L}$ into $\mathcal{L}^+$ by the mapping $Ft_1 \ldots t_n^* \mapsto Ft_1 \ldots t_n$, $(\phi \wedge \psi)^* \mapsto (\phi^* \wedge \psi^*)$, $(\neg\phi)^* \mapsto \neg\phi^*$ and $(\forall x\phi)^* \mapsto \forall x(Px \to \phi^*)$. Church's observation was that Lambert's logic proves ϕ from Γ iff classical logic proves ϕ^* from Γ^*. One can then infer the soundness and completeness of free logic with respect to Meinongian models from the soundness and completeness of classical logic with respect to classical models by taking the extension of P in a classical model to be the domain of a Meinongian model.

always determines a model of extensional first order logic. However, the kind of model it determines is not a standard first order model but a Meinongian model of the type above. That is to say, the extension of a predicate relative to a world need not in general be a subset of the domain relative to that world, and the value of a term at a world need not necessarily belong to that world. Free logic, rather than standard quantificational theory, is the natural companion to quantified modal logic. Indeed I take this to be a powerful response to Williamson's argument [28] that the Barcan formula and its converse are unavoidable since there are no clean axiomatisations of first order logic and propositional modal logic which fail to prove them when combined.

As we shall see later, it will be convenient also to have function symbols, in addition to relation symbols, in our language. The semantic clause for unary function symbols is as follows:

- $I(f_i)$ is a unary function whose domain contains the union of (i) D and (ii) the range of $I(f_j)$ for every j.

The procedure for computing the values of terms containing function symbols is completely familiar from standard model theory. We may accordingly revise our treatment of variable assignments so that they range into the extended domain defined by (i) and (ii).

2 Positive and negative free logic

Within free logic there are three positions concerning how one treats atomic sentences containing non-denoting singular terms and three analogous positions on evaluating variables in quantified modal logic. Negative semantics treats all atomic sentences involving empty names as false, neutral semantics treats all atomic sentences involving empty names as neither true nor false, and positive semantics treats at least some atomic sentences involving empty names as true. (Note, however, that some negative and neutral semanticists allow for true identity statements.) In the literature on modal logic the distinction between negative and positive semantics corresponds roughly to the distinction between serious and non-serious actualism (see [20], [1], [25].) Prior's modal logic **Q** could be thought of as corresponding to the neutral position. The slogan behind negative and neutral semantics is that all properties are existence entailing. For a one-place predicate, F, we can state what this amounts to with the formula $\Box\forall x\Box(Fx \to \exists y x = y)$. Similar formulations can be made for many place predicates.

The Meinongian models of the previous section can be restricted to be in line with negative free logic by stipulating that the extension of a unary predicate is always a subset of the domain; and by analogous stipulations for many place predicates. This restriction is not logically idle. A negative semanticist would endorse instances of the schema

$$\phi \to \exists x t = x \text{ whenever } \phi \text{ is atomic and contains } t \text{ free.} \tag{4}$$

which cannot be derived from Lambert's system. Systems for negative free logic are usually less well behaved. Among other things, they cannot be closed under the rule of substitution which guarantees that one will always get a theorem by substituting atomic formulae for complex formulae within a theorem. They also form a weak basis for a quantified modal logic. For example, Stalnaker shows that the combination a negative free logic with a propositional normal modal logic will fail to be complete with respect to the class of variable domain Kripke models (see [25].)

More importantly, it is not clear that negative free logic suffices to capture the way we ordinarily speak. I will now consider some examples which seem to me to present problems for the general spirit of negative semantics. Some of the examples are explicit cases of true atomic sentences involving empty names, others are not. My objective at this point is not so much to settle the debate about whether atomic sentences have a special status or not, although I'm inclined to think not, but rather to present convincing cases that show empty names can make an interesting semantic contribution to a sentence. Perhaps someone calling themselves a negative semanticist could find a way to accommodate these examples. This is good enough, for then she would have the resources to accommodate what I say in the rest of the paper.

Intensional transitive verbs. Suppose you are faced with a pile of Lego bricks, and detailed instructions on how to build a model castle out of them. You are then asked to draw the castle that would be constructed out of those bricks if the instructions had been followed correctly. If your drawing is good enough, it seems very natural to say that had the instructions been followed and the castle been made, the picture you actually drew is of that very castle. Thus we could say

> There could have been something which you actually drew but actually does not exist. (5)

Which is an instance of the formula $\Diamond\exists x(@Fx \wedge @\neg\exists y x = y)$ which the negative logician rejects.

This example worked because 'x drew y' is not existence entailing in its right argument; one can draw Zeus without there having been anything you've drawn. This property is common to other intensional transitive verbs, including verbs like 'imagine', 'look for', 'want', 'fear', (see [7] for a more comprehensive list.) With small modifications one could construct similar examples where, although there in fact are no model castles I'm imagining (looking for, wanting, fearing) there could have been a castle which I am actually imagining, (looking for, wanting, fearing.) The point extends to the case of non-denoting names as well.

> Botticelli drew Venus (6)

> The ancient Greeks worshipped Zeus (7)

The negative free logician will want to deny that the ancient Greeks worshipped Zeus, or that Botticelli painted Venus on the grounds that neither Zeus nor

Venus exist. However, on the face of it these claims are unobjectionable and perfectly compatible with the non-existence of Zeus and Venus.

Propositional attitudes. True propositional attitudes reports can involve empty terms. For example I might mistakenly believe that Sherlock Holmes is a detective living in 221B Baker Street, or I might doubt that Pegasus has wings. Although these are not strictly atomic sentences once doxastic attitudes have been included in the language, the problem for negative semantics is that it demonstrates that empty names have non-trivial semantic properties. For example, if I were under the impression that Conan Doyle's stories were accounts of real people then I might believe that Holmes is a detective without believing that Watson is a detective. To account for the difference in truth value between these belief reports there must be some semantic difference between the names 'Holmes' and 'Watson'; something hard to accommodate in the spirit of negative semantics.

It is possible to construct examples using only bound variables instead of names. Consider the following scenario.[8] An hour before the stabbing the prime suspect, Jones, obtained a knife handle and a blade. The chief investigator comes to believe that Jones stabbed the victim with a knife constructed from the handle and blade. Unfortunately the chief investigator is mistaken: the knife handle and blade were destroyed before they could be put together - there never was nor will be a knife made of that handle and that blade. Nonetheless, it seems extremely natural to say that had the blade and handle been put together there would have been a knife and it would have been the very knife the chief investigator in fact believed to be the murder weapon. Thus we can say

> There could have been a knife which the chief investigator actually believed to be the murder weapon but which doesn't actually exist. (8)

As before, this translates to an instance of the formula $\Diamond\exists x(@Fx \wedge @\neg\exists y x = y)$ which the negative logician cannot accommodate.

Counterfactual and fictional properties. Had Conan Doyle's stories been true, Sherlock Holmes would have existed and would have been a detective. However it is not the case that, had Doyle's stories been true, Watson would have been a detective. Thus Holmes has a counterfactual property that Watson doesn't. Similarly, according to Doyle's fiction, Holmes is a detective and Watson isn't . There are thus a large class of operators – 'according to the Sherlock Holmes fiction', 'according to Greek mythology', and so on – which are sensitive to semantic differences between empty names like 'Sherlock Holmes', 'Pegasus' and 'Vulcan.'

> According to the Sherlock Holmes fiction, Watson is a doctor (9)

> Had Holmes existed he would have been a private detective (10)

> Had we been in a Newtonian universe and had Mercury's orbit been as it actually is, there would have been an intra-Mercurial planet: Vulcan. (11)

[8] This is based on an example from Williamson [29].

As in the previous cases it is possible to formulate quantificational variants of these problems.

Predicate modifiers. It is also worth noting that some predicate modifiers like 'mythological', 'imaginary', 'possible' or 'fictional' can combine to produce a predicate which is not existence entailing.

Pegasus is a mythological horse-god. (12)

Sherlock Holmes is a fictional detective. (13)

This concludes my case against negative semantics. Before moving on, let me address one line of response which purports to concede the truth of the examples presented in this section while retaining the thesis that all properties are existence entailing. A negative free semanticist claims that since Zeus does not exist Zeus has no properties. Thus, in particular, Zeus does not have the property of not existing, or the property of having no properties. This is coherent provided one distinguishes carefully between asserting a complex sentence involving a name, '$\phi(a)$', and asserting a sentence ascribing a the property of being an x such that $\phi(x)$. At a purely technical level it seems like the negative free semanticist already has the resources to accept the examples presented in this section and keep in verbal agreement with the slogan that all properties are existence entailing. For example, a negative free semanticist could also hold the following

Botticelli painted Venus, although Venus does not have the property of being painted by Botticelli (14)

Jones believes that Sherlock Holmes is a detective, but Sherlock Holmes does not have the property of being believed by Jones to be a detective (15)

According to Greek mythology Pegasus is a winged horse, although Pegasus does not have the property of being a winged horse according to Greek mythology (16)

I have no objection to someone using property talk in this way if they wish to. At this point, however, I have been granted enough resources to write the rest of this paper. By replacing the locutions about properties like those appearing in the second conjuncts in the above examples, with locutions not involving properties like those appearing in the first conjuncts it appears that I can appease the kind of negative free logician described here whilst carrying out the aims of this paper.

Accounting for these intuitions will, I think, ultimately require dropping any form of negativaty or neutrality requirement on the semantics. However, once one has bought into a positive free logic of some sort, there are a number of further claims one could make which go beyond the intuitions presented, but which one might have theoretical reasons to adopt. These following examples are not to be taken as part of the argument for rejecting negative semantics but as optional theses which can be upheld once a positive free logic is adopted.

Logical relations. In order to keep the simplest logic of identity one typically has $t = t$ as an axiom schema where t can be any term. In a modal logic this would yield also the schema $\Box t = t$. This would involve holding the following claims

Pegasus is identical to Pegasus (17)

Necessarily Hesperus is identical to Hesperus (18)

It should be noted that some negative semanticists accept these sentences.

Essential properties. One might think there are certain kinds of properties such that if one has them at all then one has them necessarily. For example, if someone is a human, then they're necessarily human – $\Box\forall x\Box(Hx \to \Box Hx)$. According to the negative semanticist, however, there is a way I could have failed to be human – by failing to exist. A positive free logician needn't make this concession; humans are human at every world, whether they exist there or not.

Let me highlight a feature of this proposal. Presumably the line of reasoning would extend to natural kind terms which aren't actually instantiated. For example, perhaps being a unicorn is had necessarily if had at all - $\Box\forall x\Box(Ux \to \Box Ux)$. Now, if one assumes that there could have been a unicorn, $\Diamond\exists x Ux$ and the principles of $\mathsf{S5}$ (although KTB would suffice) one can prove that there could have been actual unicorns: $\Diamond\exists x @Ux$. Although I think that this is no cause for alarm – there aren't any actual unicorns, it's only that there might have been – I will not need this thesis in what follows so I shall put this issue to one side for the time being.

3 The meaning of an empty name

There is a well known objection to the Meinongian model theory outlined in section §1. If true sentences like 'there is nothing identical to Pegasus' show anything at all, they show that the classical principle $\exists x t = x$ does not merely fail to be logically true, but has actual false instances. However, the counterexamples to $\exists x t = x$ in the model theory involve assigning a referent to t that is not in the domain of the model. Since we must assign some object to t, we are committed in the metalanguage to there being something t refers to. The quantifiers of the metalanguage in this case range wider than those of the object language.

From this we could infer that the Meinongian models, while perhaps sufficient for characterising the valid inferences, are not suitable for doing semantics. An argument for this might run as follows. Since Pegasus doesn't exist, the sentence 'Pegasus does not exist' must be true according to the intended model. But if the intended model is of the Meinongian sort described, then there is some object, the object the model assigns to 'Pegasus', which does not belong to the domain of the model. Thus the model in question can't be the intended model since it fails to accurately capture the actual range of our quantifiers.

Alternatives to the Meinongian models have been proposed. Supervaluational semantics have been developed, however these have similar commitments to the Meinongian model theory (see the articles by Bencivenga and van Fraassen in [13].) More recently Antonelli, [2], has provided a model theory for positive free logic which does better in many of the respects mentioned here. While I am optimistic that one could find a model within Antonelli's framework which extensionally characterised truth in some formalised fragment of English, in my view his account does not give an adequate account of semantics at the subsentential level.[9]

Unfortunately it is hard to see how a better semantics could be given. The fundamental problem is in specifying the truth conditions of atomic sentences. Bencivenga puts the problem as follows: 'an atomic sentence is true if it corresponds to a fact [...] but an atomic sentence containing an empty name cannot correspond to any fact; hence it is false' [13]. I shall interpret 'fact' here to mean a true proposition. Since an atomic sentence containing an empty name cannot correspond to any fact all such sentences must fail to be true. There seems to be two main responses to this thought. One is to treat atomic sentences involving empty names as false or truth valueless. The other is to deny that the putative examples really involve non-denoting names by allowing them to refer to possibilia, abstract objects, or some other special entity. The former approach is problematic for the reasons discussed in the previous section, and the latter approach has a highly inflated ontology.

The semantics I shall develop here is based on a rejection of Bencivenga's premise that no sentence containing an empty name can correspond to a fact. I agree with some instances of this premise. On the one hand it seems obvious that 'John is taking the train to London' cannot correspond to a true proposition if 'John' does not refer to anything. On the other hand it is a fact, or so I claim, that the ancient Greeks worshipped Zeus, and this very fact can provide the truth conditions for the English sentence 'the ancient Greeks worshipped Zeus' even though this sentence contains a non-denoting singular term. The thought that an atomic sentence containing an empty name cannot correspond to a fact, or true proposition, is plausible only in a restricted class of cases where the predicate is existence entailing. The motivation for the fully general premise seems to rely on (a) an inflationary notion of fact in which Zeus must be a constituent of the fact that Zeus was worshipped by the ancient Greeks, and (b) that *being a constituent of the fact that Zeus was worshipped by the ancient Greeks* is existence entailing, which in this case amounts to saying that if Zeus is a constituent of the aforementioned fact, then Zeus must exist. As far as I can see, neither (a) nor (b) is sacrosanct.

[9]In particular, the treatment of empty terms is language dependent in a philosophically objectionable way. Essentially, the role of a non-denoting term is uniquely determined by a function from predicates of the language to the set $\{+, -\}$, whose value depends on whether the predicate truly applies to the term. If one is willing to consider models built out of the syntax one can always construct from the set of true sentences a term model. This shares with Antonelli's approach an adequate account of truth, but would fail at the subsentential level since a term would refer to an equivalence class of terms. The approach in question is, of course, unobjectionable taken as a formal characterisation of logical consequence.

Perhaps a more compelling argument can be given for the conclusion that a positive semantics will fail at the subsentential level. The argument might go as follows: the truth value of a sentence of the form 'a is F' is determined partly by the referential properties of the name 'a'. If 'a' is an empty name it has no interesting referential properties, so any pair of sentences of the form 'a is F' and 'a is G' must have the same truth status, whether that be truth, falsity or neither truth nor falsity.

In this section I will be rebutting this argument by showing that empty names do in fact have interesting referential properties. This will involve two theses. Firstly, that 'refers' does not fall under the category of existence entailing verbs, and is closer in kind to intentional verbs such as 'imagines', 'draws' and 'thinks about' which were discussed in section §2. Secondly, I will argue that for meaningful empty names, 'a', there could have been something which 'a' actually refers to - a claim which can be formalised as $\Diamond\exists x(@`a\text{' refers to } x)$.[10] If 'Pegasus' actually refers to Pegasus the second thesis entails that Pegasus might have existed. This claim allows us to explain how sentences involving the name 'Pegasus' can have interesting truth conditions by appealing to the modal properties of Pegasus.

One might think this project is somehow incoherent. On the one hand we wish to assert the following

$$\text{`Pegasus' refers to Pegasus.} \tag{19}$$

However, since 'Pegasus' presumably doesn't refer to anything other than Pegasus, and Pegasus doesn't exist, we also have to say the seemingly incompatible

$$\text{`Pegasus' doesn't refer to anything.} \tag{20}$$

The supposed incoherence is easily resolved once one has noted that the incompatibility of (19) and (20) relies essentially on the classical principle of existential generalisation, and that (19) and (20) are consistent in a positive free logic. Just as there are no mythological horse-gods, there are no referents of 'Pegasus'. But that is not to say that Pegasus isn't a mythological horse-god, or that 'Pegasus' does not refer to Pegasus.

Specifying the application conditions for predicates is similar. We wish to say things like the following

'mythological horse-god' does not apply to anything since there are no mythological horse-gods. (21)

'is a mythological horse-god' applies to Pegasus. (22)

As before, these sentences are only inconsistent in classical logic. So long as we abide by free logic when reasoning about reference and application we should

[10] I qualify the claim to *meaningful* names since it is widely thought that there are at least some referring expressions which deserve a genuinely different treatment. For example, failed demonstratives, or perhaps names introduced by failed demonstratives (see Evans's [5] for more discussion of these cases.)

not run in to trouble. If there is an incompatibility between (19) and (20) (or between (21) and (22)) it is not a logical incompatibility. An argument for an incompatibility between (19) and (20) would have to establish that 'refers' is among a special class of existence entailing verbs. That is, that 'refers' is closer in kind to verbs like 'walks', 'jumps' and 'drinks' than verbs like 'imagines', 'draws' and 'fears'.

3.1 A suppositional model of baptism

In order to evaluate the possibility that 'referring to' is existence entailing we need to look more closely at the nature of reference.

One response is to adopt a deflationary attitude to questions about reference. The reference relation, according to a deflationist, is completely characterised by sentences like "Pegasus' refers to Pegasus.' Saying that Pegasus stands in reference relations does not require anything particularly substantial of Pegasus, and is quite plausibly not existence entailing either. The role of reference in giving the meaning of a name like 'Pegasus' could perhaps be eliminated in favour of a conceptual role semantics of some sort. However, if we are to give an inflationary referential semantics we must answer what I will call the 'metasemantic question': what features of the way we use the word 'Pegasus' ensure that we refer to Pegasus by it and not, say, to Sherlock Holmes.

To answer the metasemantic question I will adopt a roughly causal account of reference according to which a name's referential properties can be traced via its use back to an initial baptism. For our purposes a baptism could be achieved by a guesture, a definite description or just a network of initial uses of the name which pick out some object. The crucial aspect of this account is that baptisms may take place within the scope of a supposition of some kind. As a simple example consider counterfactual suppositions. The counterfactual model of baptism is a natural way of accounting for names introduced in the course of a mathematical proof. One of Frege's examples of a name from 'On Sense and Reference' is introduced in the course of trying to prove that the lines connecting the corners of a triangle to the midpoints of the opposite sides always intersect at the same point. To prove this you would begin by supposing there was a triangle. Then, you might say something like 'let a, b and c be the lines connecting the vertices of the triangle with the midpoints of the opposite sides' before going on to prove that a, b and c all intersect the same point. 'a', 'b' and 'c' do not refer to anything, but they can be shown to have interesting properties in the scope of the supposition. Indeed, philosophy itself is littered with names that were introduced through thought experiments in this way. In none of these cases, however, do the baptising descriptions actually denote. A phenomenon similar to baptism occurs with respect to anaphoric pronouns. In such cases a pronoun is introduced into the conversation with its reference fixed by an earlier use of a description, or some other noun phrase. This can even happen when the description was introduced in the consequent of a counterfactual.

> If Phil had come, he would have made a dessert. It would have been pie. (23)

The reference of the word 'it' is introduced in the scope of a counterfactual and it is clear that the truth of (23) does not commit one to the existence of any of Phil's pies.

These are examples of speech acts that introduce a term while in the context of a false supposition. My thesis is that this kind of speech act provides a model for baptisms for meaningful empty names quite generally. In some cases the suppositional context might be provided by a widely held false belief. The ancient Greeks may have introduced the name 'Zeus' to denote the ruler of the twelve Olympians and the god of sky and thunder.[11] However the name remained in use even after the supposition that there was any such person was lifted. So it is quite natural to apply something like the counterfactual model of baptism here as well. In other cases a name might be introduced in the scope of a pretense, rather than a supposition, which no-one seriously believes. This seems quite plausible in the case of fictional names like 'Sherlock Holmes.'

Suppose that the name a was introduced by a description, 'the F', under the supposition p. Suppose also that had p obtained, there would have been exactly one F. Then we can specify a's referential profile by the counterfactual: had p obtained there would have been exactly one F and the F would have been such that a actually refers to it. In other words, according to the closest p world there is exactly one F which a actually refers to. This is the basic idea behind the suppositional model of baptism. The metasemantic question is thus answered as follows: 'Pegasus' refers to Pegasus and not Sherlock Holmes in virtue of the counterfactual properties Pegasus has which Sherlock Holmes doesn't. Furthermore, which counterfactual properties must be satisfied to be referred to by 'Pegasus' is determined by the initial baptismal use of 'Pegasus'. For example, it is Pegasus, and not Sherlock Holmes, who would have been a horse-god had the beliefs of the ancient Greeks been true. As noted in section §2, counterfactual properties like these are not in general existence entailing. Thus according to our analysis of reference one cannot infer that 'Pegasus' refers to something from the fact that 'Pegasus' refers to Pegasus.

It should be noted that I am implicitly in agreement with the Stalnakerian analysis of counterfactuals in which there always is a closest p-world to the actual world. One cannot therefore assume that there will be a fact of the matter concerning which the closest p-world is. I shall say more on this in the next section.

3.2 Could there have been something 'Pegasus' actually refers to?

Given what has been said so far, one might be tempted to characterise the current view as one in which empty names refer to possibilia. Say that something is a *mere* possibile if it doesn't exist, but might have done. We can state what it means for x to be a mere possibile with the formula $\neg\exists y x = y \wedge \Diamond\exists y x = y$. What

[11] It should be noted that the causal history of a name like 'Zeus' probably cannot be traced back to any single baptismal speech act, and is more likely linked to a network of uses.

is distinctive about empty names is that if they refer to possibilia at all they refer to mere possibilia. However, it is possible to prove that necessarily, there are no mere possibilia. One can prove in positive free logic that $\forall x \exists y x = y$. Given the rule of necessitation one can then obtain the claim that necessarily everything is identical to something, $\Box \forall x \exists y x = y$, which is incompatible with the possibility that there is something not identical with anything but which is possibly identical to something $\Diamond \exists x(\neg \exists y x = y \wedge \Diamond \exists y x = y)$, i.e., this is incompatible with the possible existence of mere possibilia.

The claim that 'Pegasus' refers to a mere possibile is unequivocally false. 'Pegasus' does not refer to anything at all, and if it did it wouldn't refer to a mere possibile since there are no such things. Nonetheless the present view asserts that there could have been something 'Pegasus' actually refers to. If by 'refers' one simply means 'refers in English as it is actually spoken' this statement can be simplified to the claim that 'Pegasus' might have referred to something. Even with this improved formulation the claim is highly contentious. The primary source of contention originates from an addenda to Kripke's 'Naming and Necessity' lectures [12]. There it is argued that names like 'Pegasus' and 'Sherlock Holmes' couldn't have even *possibly* referred to anything since the fictional or mythical suppositions under which names like 'Pegasus' are introduced are radically incomplete. For example, Greek mythology leaves it open whether Pegasus has a white hind leg. As a result there are many possible worlds in which events play out according to Greek mythology, and between those worlds many possible winged horses playing the role that Pegasus does according to the Greek myth. Now given that 'Pegasus' refers rigidly, if it refers at all, which thing in which world does 'Pegasus' refer to? What could there possibly be in our use of the word 'Pegasus', the thought goes, that singles out exactly one of these objects over the rest?

If this argument is successful it runs the risk of overgenerating. Given that the name 'Princeton' refers to something, one might reasonably ask what aspect of our use of the name determines which precise geographical region it refers to. Presumably there are plenty of regions that could approximately have been designated by a baptism of Princeton. Yet in this case we have no temptation to conclude that 'Princeton' does not refer to anything at all; only that it is somewhat vague what it refers to. Indeed most names for macroscopic concrete objects are vague to some degree - there will almost always be multiple fusions of atoms in the vicinity of an named object between which the name indeterminately refers.

Perhaps the issue is not that there could have been several things among which 'Pegasus' indeterminately refers, but that the indeterminacy is quite radical. Let us suppose, for the sake of argument, that at any world any horse at that world could have filled the Pegasus role in some world. If possibly filling the Pegasus role is all it takes to be a candidate referent of 'Pegasus' then it follows that necessarily every horse is a candidate (actual) referent of 'Pegasus'.

But the account in question is not one in which possibly satisfying some description is all it takes to be a candidate referent of 'Pegasus.' A candidate referent of 'Pegasus' must be the subject of a baptism to which subsequent uses of the name 'Pegasus' can be traced back to. To determine when something is a candidate 'Pegasus' referent we go to the closest world in which the Greek myths are true and look for the object satisfying the relevant features determined by the baptism. I think there are two plausible implications of this account for this case. Firstly, if it's a determinate matter which the closest world in which the Greek myths are true is, then there is no more indeterminacy concerning which object at that world was baptised than in a ordinary case of a baptism which is not inside the scope of a false supposition. Secondly, if it is indeterminate which the closest world is, the indeterminacy is not as radical as it would be if one were permitted to look at any world whatsoever to find a Pegasus candidate. As was argued in section §2 Pegasus has a rich set of counterfactual properties, and it is these properties that provide us with the distinguishing features of Pegasus; features that allow a name to latch on to it with some level of determinacy.

A competing picture has it that, for example, Pegasus and Sherlock Holmes are featureless and interchangeable - Pegasus could well have played the role Sherlock Holmes does and vice versa. It is natural to treat empty names on this view as radically indeterminate. In contrast the picture I'm endorsing is one in which Pegasus and Sherlock Holmes have rich counterfactual properties which allow us to say something a bit more substantial about the truth conditions of sentences involving names like 'Pegasus' or 'Sherlock Holmes.' It's worth noting that there are examples in which the competing view is quite implausible. For example, Salmon [24] introduces the name 'Noman' for the person that would have been born if some specific sperm had fertilised some specific egg. Reference to the sperm and egg are both determinate, since both actually exist, so given some form of essentialism about origins it is natural to think that no two possible objects [sic.] could satisfy this description.[12] Williamson [29] describes an example, considered earlier, in which there is a knife handle and blade which haven't been put together. Once again it seems that no more than one knife could have been made by putting that handle and that blade together, allowing us to construct a determinate name such that, had the handle and blade been put together, it would have referred determinately to the resulting knife.

3.3 Inconsistent fictions and counterpossibles

It is natural to think that names which are introduced in impossible or inconsistent suppositions pose a special problem for the kind of view defended here. Firstly it appears to be problematic because I have claimed that for every meaningful name there could have been something to which that name actually refers. Secondly, the account assumes an empty name's referential profile is determined by the object which would have satisfied a certain property under a certain supposition. However, on standard theories counterfactuals with impossible antecedents are all vacuously true.

We will begin by addressing the first problem. In 'Sylvan's Box' [21], Graham Priest presents a short story in which he describes how he stumbled across a box

[12] In non-possibilistic language, $\Box\forall x\Box\forall y\Box(Rxse \wedge Ryse \rightarrow x = y)$, where s and e denote the sperm and egg in question, and $Rxyz$ means had y fertilised z, x would have been conceived.

labelled 'Impossible Object' in Richard Sylvan's house shortly after his death. According to the story, when Priest opens the box he discovers both that the box contains a figurine and that it is completely empty. Unlike other putative examples from existing literature, the contradiction is central to the story and cannot be made consistent without significantly altering the content of the story.

In order to state the problem for my view, I'll introduce a name, 'Parabox', for the box Priest discovers in the story. Is there a problem for the view that Parabox might have existed, and is there a problem for the view that there might have been something to which 'Parabox' actually refers? I do not think so. Parabox isn't an impossible object in and of itself – it just has inconsistent properties according to the story. It would be unwise to conclude from this that Parabox couldn't have existed. By analogy, note that *Graham Priest* has inconsistent properties according to the story – according to the story he saw there there was an empty and non-empty box – but it clearly does not follow that Graham Priest could not have existed. There is no actual or possible contradiction involved in assuming that Parabox actually or possibly exists. One cannot, for example, infer that Parabox is empty and non-empty from the fact that Parabox is empty and non-empty according to Priest's fiction. Indeed, maybe Richard Sylvan really did keep an empty box labelled 'impossible object' as a joke, and this box inspired Priest to write his story about it. If this were the case, then Parabox would in fact exist, but it would be a boringly consistent empty box.

Let's turn to the second problem. We'll use a different example this time. Consider a (somewhat elliptical) proof by contradiction showing that there is no largest prime.

> If there were a largest prime p, $p! + 1$ would be prime. Therefore p is not the largest prime.

Here 'p' was introduced in the scope of an impossible supposition. In order to apply the suppositional model of baptism to this case we thus need to make sense of a counterfactual with an impossible antecedent.

According to one view, popular among the early possible world semantics for counterfactuals, counterpossibles – counterfactuals with impossible antecedents – are always vacuously true. On this view the referential profile of 'p' would be completely unconstrained: necessarily everything is a candidate actual referent for 'p'.[13]

I believe that some counterpossibles, such as the one above, are non-vacuous. While the view I have been developing does not require this commitment, we can greatly reduce the amount of indeterminacy that is present in a name like 'p' under this assumption. For example, I believe that the following counterpossible is determinately false: if there had been a largest prime, it would have been Saul Kripke. Although there are impossible worlds where Saul Kripke and other non-numbers are prime, presumably the *closest* impossible world where

[13]By the vacuity of counterpossibles it follows that necessarily, for any x, if there had been a largest prime, then 'p' would have actually referred to x.

there is a largest prime is a world where it's a number, and is not identical to Saul Kripke. Generalising from this example it would seem to follow that 'p' picks out an actual natural number at the closest world, so that 'p' refers to a number and is not an empty name. On the other hand it seems like the following counterpossible, for each possible substitution of 'n' for a numeral, is indeterminate: if there had been a largest prime it would have been n. So there is no number p is determinately distinct from - thus 'p' is an indeterminate name.

Are names introduced in inconsistent fictions radically indeterminate? As Priest himself points out, not everything is true in his fiction. For example, according to Priest's story Parabox is not Richard Sylvan – Sylvan plays a completely different role in the story. So it's determinate that 'Parabox' does not refer to Sylvan. If, as we imagined earlier, there was an actual box which Priest wrote his story about there would be a completely determinate referent for 'Parabox.' So it is not true in general that a name introduced in an inconsistent fiction is radically indeterminate.

3.4 Direct reference

Let me briefly say a few things about the thesis that proper names directly refer. Just as it is consistent in free logic to say that 'Pegasus' refers to Pegasus, even though Pegasus does not exist, it is also consistent to say that 'Pegasus' *directly* refers to Pegasus, who doesn't exist.

But consistency is a weak claim. For example, free logic does not rule out the logical possibility of the following scenario: that there could have been a foot, Foot-3, which doesn't actually exist, but is nonetheless actually a spatial part of my left leg and is not identical to my left or right foot. This scenario is clearly absurd since standing in parthood relations, especially spatial parthood relations, require existence. In other words, *being a part of something* is existence entailing.

In so far as I understand what is distinctive about the thesis that names directly refer, it is the metaphysical claim that the proposition expressed by a sentence containing a proper name contains the referent of that name as a constituent. Without this the direct referentialist cannot really distinguish herself from rival theories ('mediated' reference theories) that agree about what refers to what but for the insertion of the word 'directly.' The term 'refers directly' is too jargony to be grounding heavy philosophical distinctions on its own. And even if there is a non-technical use of 'refers directly', it is doubtful that doing conceptual analysis on it is going to help us with the issue of whether names directly refer.

However if this notion of constituenthood that can hold between objects and propositions is supposed to be metaphysically substantial one might expect it to behave much like spatial parthood. Since I exist, being a spatial part of me is existence entailing, which is why the example involving Foot-3 sounded so absurd. So if a direct reference theorist wishes to adopt this kind of free logic, she either needs some explanation of constituenthood between an object and an existing proposition which doesn't require existence, or some theory about how non-existing propositions with non-existent constituents can still play their role in semantics without the need to quantify over them.[14] For more discussion of these issues see [20], [26] and [30].

4 Semantics

Suppose that $\mathcal{L}$ is a regimentation of some first order fragment of English without semantic expressions into first order logic. What is required to describe the intended model of $\mathcal{L}$? To do this we must specify the referential profile of each term, the application conditions of each atomic predicate, and we need to describe how the truth values of complex formulae depend on the truth values of their parts.

For our purposes, at least, this will be a sufficient level of detail to describe the intended model in. I am not making the general claim, characteristic of Davidson, that giving a compositional theory of truth is all that is required of semantics. The characterisation of reference, extension and truth for the intended model is only instructive here because we are dealing with an extensional language.[15]

There are a number of metalanguages one could use. However, for concreteness I shall assume that the metalanguage extends $\mathcal{L}$ with some means to describe the syntax of $\mathcal{L}$, a function symbol, $\mathrm{Ref}(\cdot)$, on the terms of $\mathcal{L}$, a relation App('F', x) between predicates and objects, and some set theoretic machinery for constructing variable assignments. In cases where a is non-empty $\mathrm{Ref}(a)$ can be read intuitively as 'the referent of a' and the relation App('F', x) is to be read as "F' applies to x.' For simplicity we shall assume that $\mathcal{L}$ only contains monadic predicates. Given that our new vocabulary represents reference and application respectively the following two schemata should hold

- Ref('a') $= a$
- App('F', a) if and only if Fa.

where a and F can be replaced by any name and predicate belonging to the $\mathcal{L}$-fragment of English. Relative to a variable assignment v the satisfaction clause for atomic sentences is then:

- 'Fa' is satisfied by v if and only if App('F', Ref('a')).

[14] There are weaker notions of constituenthood which seem to be more neutral about the metaphysics of propositions. For example, we can capture the thought a is a constituent of the proposition that p by saying that there is some unsaturated concept F such that the proposition that $p =$ the proposition that a is F. A possible way to formulate this would be: $\exists F(\Diamond\exists x\Diamond\exists y\Diamond(Fx \wedge \neg Fy) \wedge \Box(Fa \leftrightarrow p))$. This notion is not existence entailing. However, it is also a notion that anyone can make sense of without modifying their views about reference; this would seem to rob the direct reference theorist of a distinctive thesis.

[15] I believe it is possible to carry out the project of describing the intended model for intensional languages too, but that would take us too far afield for this paper.

- 'Fx' is satisfied by v if and only if App('F', v('x')).

The definitions of satisfaction for complex sentences carry on exactly as in Tarski's original definition of truth, i.e.

- '$(\phi \wedge \psi)$' is satisfied by v if and only if 'ϕ' is satisfied by v and 'ψ' is satisfied by v.
- '$\neg\phi$' is satisfied by v if and only if 'ϕ' is not satisfied by v.
- '$\forall x\phi$' is satisfied by v if and only if 'ϕ' is satisfied by u for every x-variant of v, u.

Before we move on let me address a technicality: the function symbol, Ref(). I stress that this is only *approximated* in English by the phrase 'the referent of ...'. The reason we can not properly reduce our function to a definite description involving the more natural relational word 'refers' is that on at least some accounts of definite descriptions a sentence like "mythological horse-god' applies to the referent of 'Pegasus" can only be true if something is the referent of 'Pegasus'.[16]

The current formulation of the metatheory is neutral on the issue taken up in the previous section - whether Pegasus, Sherlock Holmes, and so on, might have existed. If we assume the results of section §3 we can paraphrase away uses of the 'Ref' function symbol in favour of a more natural reference relation, which I will write Rxy. Thus, for example, we can restate the satisfaction clause for atomic sentences: 'Fa' is satisfied by v iff $\Diamond\exists x@(R\text{`}a\text{'}x \wedge App(\text{`}F\text{'}, x))$.

4.1 Model theory

In this section we return to the question of which inferences are valid. In §1 we gave a complete proof theoretic and model theoretic answer to this question, but noted that the model theory in question would not provide an intended model. In §4 it was shown that one could nevertheless describe an adequate semantics in a free metalanguage.

One might wonder whether this separate treatment of semantics and validity could lead to a conflict. Might a sentence be false according to the semantics of §4 and yet be true on every model from §1? In other words, could a false sentence be valid according to our account? Furthermore, given that we have admitted that there are some interpretations which are not among the Meinongian models, could there be an inference which preserves truth at every Meinongian model which is not a genuinely valid inference – i.e. an inference which preserves truth in all Meinongian interpretations but does not preserve truth in all interpretations.

There is a similar worry about the classical account of logical validity and the classical model theory. To capture the intended meaning of the universal

[16] There are free description theories in which this does not hold, see, for example [15], however I do not wish to take sides on this issue.

quantifier one must have an unrestricted domain. However no classical model has an unrestricted domain, since according to classical model theory domains are sets, and there is no set containing everything. One might wonder whether there are invalid inferences involving unrestricted quantifiers which just happen to be truth preserving relative to every restricted set sized interpretation of the quantifiers. Luckily for the standard model theory the answer is 'no' – in [9] Kreisel proved that the genuinely valid inferences coincide exactly with the inferences preserving truth in every classical model.

Let me run a version of Kreisel's argument for the correctness of Meinongian model theory with respect to logical consequence for a first order language $\mathcal{L}$. Firstly observe that possible interpretations of $\mathcal{L}$ fall into two categories:[17] (i) those interpretations - interpretations like the intended one - which leave some terms without a referent (for example, by assigning Pegasus to the constant 'c') or assign to some predicate a property such that there could have been something to which the property actually applies which doesn't actually exist (for example, by assigning F a property which applies to Pegasus) and (ii) those interpretations that assign every term a referent and every predicate an existence entailing property. The Meinongian models of section §1 are (isomorphic to) those models of type (ii) which are set sized.

Suppose that the inference from Γ to ϕ is valid so that every possible interpretation that makes every member of Γ true makes ϕ true. Since the Meinongian models are isomorphic to a special class of type (ii) interpretations it follows that Γ entails ϕ over the class of Meinongian models. However, by the completeness of Lambert's axiomatisation of free logic with respect to Meinongian models, it follows that Γ proves ϕ in free logic. But since the axioms of free logic are evidently logically valid, and rules of inference of free logic are valid inferences, it follows that the inference from Γ to ϕ is genuinely valid, closing the gap between validity and preservation of truth in every Meinongian model.

So, for example, while the intended interpretation is not a Meinongian model for the reasons we discussed above, there will be a 'proxy' Meinongian model to take its place. Take the set of true sentences of $\mathcal{L}$ (this set will contain translations of sentences like 'Pegasus is a mythical winged horse', 'Sherlock Holmes does not exist', et cetera). Since this set is evidently logically possible, it will be consistent in free logic, and thus there will be a Meinongian model that makes each of the sentences true - a model which agrees with the intended interpretation about every sentence.[18]

Can we hope to provide an intensionally correct model theory? In a couple of papers Rayo, Uzquiano and Williamson have developed such a 'model theory' for standard first and second order logic [23] [22]. In their framework there are no models, technically speaking, because the definition of validity is not carried out in terms of singular quantification over sets or models of any kind at all. The definition is instead formulated in second order logic. This allows

[17] Note that I am taking the notion of an 'interpretation' as antecedently understood, and not to be defined set theoretically in a language whose logic is not free.

[18] This model will be unlike the intended model at the *subsentential* level – for example a name like 'Pegasus' would denote an equivalence class of terms, instead of denoting Pegasus.

one to achieve the effect of talking about arbitrary domains: given some objects one talks about what would be true if our quantifiers ranged just over those objects. However, strictly speaking, talk of domains is misguided, and should be reconstrued in terms of second order quantification. Such an approach could be applied in this case too, I claim, provided one has a strong comprehension principle:[19]

$$\exists F \Box \forall x \Box (Fx \leftrightarrow \phi)) \tag{24}$$

This constraint rules out a plural interpretation of the second order quantifiers; the presence of either $\Box$ is sufficient to rule out that interpretation, and we shall need both for our purposes. However it is compatible with a Fregean, or 'conceptual', interpretation of the quantifiers (see e.g. Williamson's [31].) In particular one may have the following instance of comprehension

$$\exists F \Box \forall x \Box (Fx \leftrightarrow x \text{ is a unicorn})) \tag{25}$$

In this way one may quantify over possible interpretations for a predicate without committing oneself to the actual existence of things the predicate only possibly applies to. Similarly one may quantify over possible interpretations of names without quantifying over the referents of those names. In the simplest case where the language contains only one name, 'c', an interpretation would just be a two place relation which was functional (i.e. $\Box \forall x \Box \forall y \Box (Rzx \wedge Rzy \rightarrow x = y)$.) For example, the comprehension schema entails there is a relation which satisfies the following

$$\Box \forall y @ \exists R \Box \forall x \Box (R\text{`}c\text{'}x \leftrightarrow x = y)) \tag{26}$$

This guarantees, for example, that there will actually be a relation R such that according to R, 'c' refers to Pegasus, and that there will actually be another relation in which 'c' refers to Sherlock Holmes, and so on and so forth. The point is that the second order quantification used to quantify over these interpretations of c does not commit one to the existence of Pegasus or Sherlock Holmes. Although I think it would take us too far afield to spell out the details fully, I hope this demonstrates that it is not hopeless to expect an intensionally correct model theory.

5 Logical revisionism

Let me conclude this paper by looking at one of the more radical claims it makes: that classical logic needs to be revised. My argument for this revision, however, is based on premisses that I believe are not considered radical or controversial. The argument is simple: classical logic is not sound. Or in other words, some of its theorems are false. In particular according to classical logic every instance of the schema

$$\exists x t = x \tag{27}$$

[19] I state just the monadic version here. One can generalise this to an n-adic version in the following way: $\exists R \Box \forall x_1 \Box \ldots \Box \forall x_n \Box (Rx_1 \ldots x_n \leftrightarrow \phi))$.

is a theorem. However, I take it that there are clearly false instances of (27) on the interpretation of first order logic in which terms correspond to proper names in English.

The standard response to this well known objection is to deny that the terms of first order logic correspond to proper names in English. The alternative is to paraphrase a sentence like 'there is something identical to Pegasus' with a different logical formula, one that is not a theorem. Perhaps the paraphrase is one featuring a Russellian definite description, or one in which a predicate replaces the role of the name 'Pegasus.' Given that one wants to avoid not just the existence of Pegasus, but also the necessary existence of, say, Saul Kripke, the paraphrase would have to be made across the board – for proper names both empty and non-empty.

The view under consideration is one that accepts classical logic as the 'correct' logic of (a first order fragment of) English, but denies that proper names, such as 'Pegasus', should be translated as names in the predicate calculus. Since I take this to be the standard way of blocking the argument let me say something more about this.

There is a good sense, I think, in which such a view has already accepted some of the logical revisionism I am proposing. For to assimilate names to, say, descriptions is to revise the logic of names to that of the logic of descriptions according to which existential generalisation and universal instantiation fails.

To this the defendant of classical logic might respond by saying that, although we are rejecting the principle of existential generalisation and universal instantiation for 'real' names, classical logic still holds because proper names don't correspond to the constant terms found in first order logic. But then *what claim is the classical logician making* when she says that classical logic as opposed to free logic is correct? Is she making a claim about an uninterpreted calculus? Or is she merely making a claim about classical logic's soundness with respect to a particular abstract semantics? The former claim has little content unless we have a standard of correctness to judge it by, and the latter is a mathematical theorem which is presumably not the subject of disagreement.

To see whether a logic is correct we must look to see whether it governs correct reasoning. However, if no expression of the same semantic category as constants in first order logic appears in ordinary reasoning, it becomes hard to evaluate whether the classical logic of constants is correct or not.

While I am sympathetic with the revisionary theorist who gives constants an alternative semantics, or proposes a different first order translation for proper names, there are still many things to be said in favour of my form of revisionism. For one thing, such theories will typically not be able to account for true atomic sentences involving empty names. More worryingly, many of the issues of this paper arise also for free variables. For example $\exists xy = x$ is also a classical theorem, and a quantified modal logic with necessitation and universal generalisation can prove in a few lines that $\Box\forall y\Box\exists xy = x$ - that necessarily everything exists necessarily. This contradicts the platitude that if my parents had never met, I would not have existed. There does not appear to be any way of paraphrasing away variables that is analogous to the way of paraphrasing away constants.

The revision to classical reasoning I am proposing is quite minimal, and is one, I think, that most people would be happy with. It is strong enough to allow semantic reasoning without commitment to a Meinongian ontology and is the natural logic for modal reasoning about contingent existence.

References

[1] R.M. Adams. Actualism and thisness. *Synthese*, 49(1):3–41, 1981.

[2] G.A. Antonelli. Proto-Semantics for Positive Free Logic. *Journal of philosophical logic*, 29(3):277–294, 2000.

[3] J. Bridge. *Beginning model theory: the completeness theorem and some consequences*. Oxford Univ Pr, 1977.

[4] A. Church. Review: Karel Lambert, Existential Import Revisited. *Journal of Symbolic Logic*, 30(1):103–104, 1965.

[5] G. Evans. *The varieties of reference*. Oxford University Press, 1982.

[6] K. Fine. The permutation principle in quantificational logic. *Journal of Philosophical Logic*, 12(1):33–37, 1983.

[7] Graeme Forbes. Intensional transitive verbs. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2010 edition, 2010.

[8] J.W. Garson. Applications of free logic to quantified intensional logic. *Philosophical Applications of Free Logic*, pages 111–144, 1991.

[9] G. Kreisel. Informal rigour and completeness proofs. *Problems in the Philosophy of Mathematics*, pages 138–157, 1967.

[10] S. Kripke. Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16(1963):83–94, 1963.

[11] S.A. Kripke. *Reference and Existence: The John Locke Lectures for 1973.* 1973.

[12] S.A. Kripke. *Naming and necessity.* Wiley-Blackwell, 1981.

[13] K. Lambert. The nature of free logic. *Philosophical Applications of Free Logic*, pages 3–16.

[14] K. Lambert. Existential import revisited. *Notre Dame Journal of Formal Logic*, 4(4):288–292, 1963.

[15] K. Lambert. A theory of definite descriptions. *Philosophical applications of free logic*, page 17, 1991.

[16] H. Leblanc and R.H. Thomason. Completeness theorems for some presupposition-free logics. *Fundamenta Mathematicae*, 62:125–164, 1968.

[17] B. Linsky and E.N. Zalta. In defense of the simplest quantified modal logic. *Philosophical Perspectives*, 8:431–458, 1994.

[18] E. Mendelson. *Introduction to mathematical logic*. Chapman & Hall/CRC, 1997.

[19] J. Nolt. Free Logic. *Stanford Encyclopedia of Philosophy*, 2009.

[20] A. Plantinga. On existentialism. *Philosophical Studies*, 44(1):1–20, 1983.

[21] G. Priest. Sylvan's Box: A Short Story and Ten Morals. *Notre Dame Journal of Formal Logic*, 38(4):573–582, 1997.

[22] A. Rayo and G. Uzquiano. Toward a theory of second-order consequence. *Notre Dame Journal of Formal Logic*, 40(3):315–325, 1999.

[23] A. Rayo and T. Williamson. A completeness theorem for unrestricted first-order languages. *Liars and Heaps: New Essays on Paradox*, pages 331–56, 2004.

[24] N. Salmon. Nonexistence. *Noûs*, 32(3):277–319, 1998.

[25] R. Stalnaker. The interaction of modality with quantification and identity. *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus*, 1995.

[26] J.E. Tomberlin and P. Van Inwagen. *Alvin Plantinga*. Number 5. Dordrecht: D. Reidel, 1985.

[27] P. Van Inwagen. Creatures of fiction. *American Philosophical Quarterly*, 14(4):299–308, 1977.

[28] T. Williamson. Bare possibilia. *Erkenntnis*, 48(2):257–273, 1998.

[29] T. Williamson. Existence and contingency. In *Proceedings of the Aristotelian Society*, volume 100, pages 117–139. Blackwell Publishing, 2000.

[30] T. Williamson. Necessary existents. *Royal Institute of Philosophy Supplements*, 51:233–251, 2002.

[31] T. Williamson. Everything. *Philosophical Perspectives*, 17:415–465, 2003.