Australasian Journal of Philosophy
Vol. 78, No. 4, pp. 494–502; December 2000

OBJECTS OF THOUGHT

Graham Priest

I. Characterisation

Benny fears the man next door. The man is a rather nasty character and a convicted criminal. Benny fears that he may break in and murder him one night. But in reality there is no such man next door: the house is, in fact, empty. Benny has simply overheard snatches of conversations between his neighbours about different subjects (the house next door and a murderer that has just been arrested), and, being a nervous individual, jumped to confused conclusions.

So whom does Benny fear? The man next door; something that does not exist. That is the obvious answer. In pedantic terms, the natural semantic parsing of '*x* fears *y*' is as a relationship between two objects, where the latter may or may not exist. And the same goes for the objects of other intentional states: 'imagines', 'worships', 'pities', etc.

Since this analysis of objects of thought was advocated, particularly by Meinong, many—starting with Russell—have felt that there is something philosophically rebarbative about the very notion of a non-existent object, that it is beset with insurmountable objections. I think that this is not the case.[1] Moreover, other attempts to give predicates such as 'fears' a parsing that does not invoke non-existent objects face well known objections.[2] None of this do I want to go into here, however.

[1] Meinong has been defended against standard objections by Parsons (1980), Zalta (1988) and particularly Routley (1980), esp. chs. 3 and 4.

What I think is the most difficult objection to the claim that there are non-existent objects, is one not discussed by any of the above, and is as follows. If one believes that there are non-existent objects, one does not have to believe that every term of the language denotes some object or other. However, given that a prime application of the theory of non-existent objects is to provide the wherewithal for an analysis of the objects of thought; and given that one would seem to be able to think of an object falling under any description whatsoever; it is natural to suppose that every term of language denotes. Now, given the resources of a semantically closed language and standard principles of identity, the assumption that every term denotes leads to triviality, which is clearly unacceptable. For a taste of the reason, consider the term 'one plus the denotation of this term'. Call this term t, and let t denote n; then t also denotes $n+1$. Hence, $n=n+1$, and so $0=1$. See Priest (1997).

The most plausible Meinongian solution to the problem, I think, is as follows. Just as paradoxical sentences may have more than one truth value, so paradoxical names, such as t, may have more than one denotation. After all, t denotes n and $n+1$, and these are distinct. In a logic of multiple denotation, the transitivity of identity fails (see Priest (1995)), and this breaks the triviality argument, as is demonstrated in Priest (1998).

[2] The most common suggestion is to reparse the relation 'x fears y' (xFy) as one between the agent and some surrogate object, especially some mental representation. Thus, 'Benny fears the man next door' is understood as 'Benny has a man-next-door-representation in his mental "fear box" '. Call this relation F'. This suggestion won't do as it stands. For example, it makes a nonsense of 'Benny fears the man next door, and he has just moved out'. This is $\exists x(bF'x \wedge Mx)$. To make sense of the

The major problems with Meinongianism lie not in there *being* non-existent object, but in giving an account of what *properties* they have. What properties, for example, does the man next door have? Well, for a start, he lives next door. But that can't be right; after all, nobody lives next door. And if you say 'nobody *existent* lives next door', it still remains the case that non-existent objects characterised in certain ways cannot, in general, have the properties they are characterised as having. Just consider the person such that he lives next door and pigs fly. If he satisfied this characterising condition, it would follow that pigs fly.[3]

In attempts to answer this question, Meinongians often distinguish between characterising and non-characterising properties (or some equivalent terminology, such as nuclear and non-nuclear properties). Descriptions do satisfy their defining conditions provided that they are characterising; not if they are not. But the problem with this is that no one that I am aware of has managed to explain the distinction between characterising and non-characterising properties in any principled fashion.

Maybe, then, one can just bite the bullet, and insist that non-existent objects have none of the properties they are characterised as having. After all, nothing in the analysis of 'fears' as a binary predicate requires them to have these. Let 'the man next door' and its ilk denote the null set, and be done with it. But this won't do either. Benny fears the man next door; he doesn't fear the King of France. Non-existent objects would seem to have the properties attributed to them in *some* sense. What sense?

I want to suggest a solution to this riddle. The suggestion is that the man next door neither lives next door, nor has any of the properties Benny fears him to have. There is, however, a way that Benny fears the world to be. And the man next door does have the property of living next door in the way he fears the world to be. Similarly, if Benny worships something, x, with property, P, x has P in the way he believes the world to be. And if Benny imagines something, x, with the property P, then x has the property P in the way he imagines the world to be. Generally, to say that some object, x, has property P *qua* object of thought, is to say that $a\varphi Px$, where a is the thinker, and φ is some appropriate psychological state (e.g., 'fears the world to be such that', 'imagines the world to be such that', 'believes the world to be such that').[4] Fictional objects, note, are just a special case of this. *Qua* fictional objects, they possess those properties that they have in the way the author imagines the world it to be.

2 *continued*
first conjunct, the quantifiers have to range over representations, but then the second conjunct is nonsense: the representation has not just moved out. One can save this view by invoking the relation 'x is a representation of y' (xRy). The sentence then becomes $\exists x\exists y(bF'x \wedge xRy \wedge My)$. But this is just the start of the problems. For example, how is one to understand 'Benny and Penny fear something (the same thing)'? $\exists x(bF'x \wedge pF'x)$ won't do: there is no guarantee that Benny and Penny have exactly the same mental representation of the object in question. We can try $\exists x\exists y\exists z(bF'x \wedge pF'y \wedge xRz \wedge yRz)$. This may work if they both fear the same existent object (z). But the object in question may not exist; and if this is so, non-existent objects are still being invoked.

3 Formally, the description is $\iota x(Lx \wedge p)$. Call this τ. If it satisfied its defining condition, we would have $L\tau \wedge p$.

4 Similar ideas have been expressed in Griffin (1998) and Nolan (1998).

II. A Formal Construction

Let me make this more precise by giving a formal semantics. Because the above analysis employs intentional propositional operators, it is necessary to have an analysis of these. I will adopt a standard possible-world semantics. Essentially, $a\varphi A$ is true at world w iff A is true at every world compatible with the way a φs the world to be. (The way Benny fears the world to be does not succeed in characterising a unique world—it is indeterminate, for example, on the matter of whether the man next door is left handed or right handed.)[5]

Take a first order language; for simplicity, with no function symbols. For the moment, we assume that the language does not contain identity or descriptions either; I will deal with these in the next section. It may still contain names for non-existent objects. Predicates, P, are of two kinds. The first are intentional predicates, those which concern the mental states of some agent, like, 'fears', 'loves', etc. The others are the familiar extensional predicates, like 'kicks', 'holds', etc.[6] To keep things simple, I will take all predicates to be one-place. In particular, I will assume that the agent we are talking about is fixed, and so may be ignored.[7] Where necessary, I will indicate extensional predicates with a suffix e, and intentional predicates with a suffix i, thus: P_e, P_i. One of the extensional predicates is distinguished as the existence predicate. I will write this as E. There is also a collection of operators, Ψ, expressing propositional states. Again, since the agent in question is fixed, we may take these to be simply sentential operators.

I will specify an interpretation for the language which is a semantics for a modal extension of the paraconsistent logic *LP*.[8] To a large extent, this is arbitrary. If the reader thinks that connectives ought to have other semantics, they are—to an extent—free to use those instead. The qualification is due to the fact that *LP* has one important property, namely, that *every* set of sentences (even the set of all sentences) is true in *some* world. This is important, since, after all, people being what they are, they could fear, imagine, etc., the world to be any way at all. If another logic is used in the place of *LP* it should have at least this property.[9]

An interpretation for the language is a structure $\langle D, W, @, \{R_\varphi; \varphi \in \Psi\}\rangle$. D is a non-empty set (of objects). W is a non-empty set (of worlds). $@ \in W$ is the actual world. For each $\varphi \in \Psi$, R_φ is a binary relation on W. (Intuitively $xR_\varphi y$ iff y is a world satisfying all the things that at world x the agent φs to be the case.) For each constant, c, $\delta(c) \in D$. ($\delta(c)$ is its denotation.) For each n-adic predicate, P, $\delta(P)$ maps each $w \in W$ to a pair, $\langle P^{w+}, P^{w-}\rangle$, where $P^{w+} \cup P^{w-} = D$. (P^{w+} and P^{w-} are P's extension and antiextension at world w.) An interpretation must satisfy one further condition to ensure that non-existent objects cannot satisfy extensional predicates. Specifically:

5 I do not employ standard possible-world semantics because I think that they are correct. Indeed, I think that they are not. For example, they make $a\varphi A$ true for every logical truth, A. A more sophisticated possible world semantics can avoid this; but I keep things simple here to highlight what is distinctive about the present approach to objects of thought.

6 Note, though, that intentional predicates are not intensional (in the technical sense). If I am thinking about x, and $x = y$, then I am thinking about y. I may just not be aware that x is y.

7 More generally, we would have a bunch of n-place predicates, and mark which argument places of each are intentional and which are not.

8 See Priest (1987), ch.6.

9 In particular, if one requires a language that has a detachable conditional operator, one could take the semantics to be a possible-world semantics for some relevant logic, such as B.

if $x \in P_e^{w+}$ then $x \in E^{w+}$

Given an interpretation, truth values are assigned to formulas in the standard *LP* fashion. Specifically, there is a binary relation, ρ_w, such that for every sentence, A, $A\rho_w 1$ or $A\rho_w 0$ or both. The truth conditions for the intentional operators are as follows:

$\varphi A\rho_w 1$ iff for all w' such that $wR_\varphi w'$, $A\rho_{w'} 1$

$\varphi A\rho_w 0$ iff for some w' such that $wR_\varphi w'$, $A\rho_{w'} 0$

For the quantifiers, I make the assumption that each object in the domain has a name. This is inessential, but allows us to talk officially in terms of truth, rather than the more complex satisfaction.

$\exists xA\rho_w 1$ iff for some c, $A(x/c)\rho_w 1$

$\exists xA\rho_w 0$ iff for all c, $A(x/c)\rho_w 0$

The truth/falsity conditions of $\forall$ are dual to those of $\exists$. Quantifiers, note, are not existentially loaded. So to express the fact that something exists satisfying A, one must write $\exists x(Ex \wedge A)$. Validity can be defined as truth preservation at @ in all interpretations.

The above semantics capture much of the semantic intuition described in the previous section. Objects may be either existent or non-existent (at a world). There are no constraints on existent objects. Non-existent objects satisfy no extensional predicates, though they may satisfy sentences containing such predicates, e.g., $\neg Ex$.[10] They may, however, satisfy intentional predicates, $P_i x$, and may also satisfy sentences of the form $\varphi A(x)$. If the language possessed modal operators, they would also be able to satisfy modalised sentences, and so have 'modal properties'. For example, if there were a possibility operator, $\Diamond$, where, as usual, $\Diamond A$ is true at a world if A is true at all worlds, or some specified subset of them, then it would be quite possible to model $\neg Ea \wedge \Diamond Ea$. That is, a is a (merely) possible object.

III. Descriptions

Adding descriptions to the machinery may be executed as follows. Let ε be an indefinite description operator with the usual syntax. (I shall not deal here with definite descriptions. These can be handled similarly, merely adding a uniqueness clause where required.) I will take descriptions to be rigid designators. It is certainly possible to modify the following

[10] Notice that the view described here does not presuppose any dubious dichotomy between positive and negative properties. We do not need, for example, to decide which of 'is transparent' and 'is opaque' is the genuine positive property, to be represented by a monadic predicate. Both are extensional properties, and may be represented by monadic predicates. Their incompatibility is recorded in the fact that every (physical) *existent* object has one or other of them, but not both. Non-existent objects have neither.

semantics in a way that does not assume this, but this certainly complicates things (requiring more machinery, such as λ-abstraction, to keep track of scope distinctions). And for present purposes, the extra complication would appear to be unnecessary. When descriptions characterise objects of thought, at least for a Meinongian, they would seem to function rigidly. When Benny fears that the man next door may murder him, it is *the man next door* (that person), who, in the world that realises Benny's fear, may murder him. Phenomenologically, it is clear that intentional states target *objects*, whether or not the object is characterised by a name or a description; and whether or not, as Husserl insisted, the targeted object exists.

Denotation of ε-terms and truth/falsity, are now defined by a joint recursion. In particular, the clause for denotation is as follows. If $A(x)$ has at most x free then:

If $\{a: A(a)\rho_{@}1\} \neq \phi$ then $\delta(\varepsilon x A(x))$ is $\delta(a)$ for some a in this set.
Otherwise, $\delta(\varepsilon x A(x))$ is z for some z such that $z \in E^{@-}$.

As is evident, for this to be a good definition, there must be non-existent objects at @. Hence, we need to assume that every interpretation satisfies the condition:

$$E^{@-} \neq \phi$$

which we do henceforth. The denotation conditions suffice to ensure that if $\exists x Px$ holds at @, then so does $P\varepsilon x Px$. (The converse is obvious.) This fact does not hold for worlds other than @, though, since ε-terms are rigid. Note that even ε-terms denoting non-existent objects may satisfy their defining conditions. For example, since $\exists x \neg Ex$ is true at @, so is $\neg E\varepsilon x \neg Ex$.

Note also that we assume nothing about how the denotation of an ε-term is selected from the appropriate set. In particular, it is not selected, as is often the case in semantics for descriptions, by a choice-function on subsets of the domain, which gives extensionality (if $\{a : A(a)\rho_{@}1\} = \{a : B(a)\rho_{@}1\} \neq \phi$ then $\delta(\varepsilon x A(x)) = \delta(\varepsilon B(x))$). The selection is purely non-deterministic.[11] This has the consequence that even relabelling bound variables may change denotation. Thus, $\varepsilon x Px$ and $\varepsilon y Py$ may well denote different things. This is actually a strength of the account. Suppose that Benny fears a man next door, and so does Penny. The objects of both fears are non-existent. Do they fear the same man or not? That depends. They may do, but the objects of their fears may be entirely different. For example, the object of Benny's fears may be tall and bearded; whilst the object of Penny's fears may be short and clean-shaven. If they do fear the same thing, the situation may be represented by the sentence $B_i\varepsilon x Mx \wedge P_i\varepsilon x Mx$. If they do not, then it may be represented by the sentence $B_i\varepsilon x Mx \wedge P_i\varepsilon y My$. As is easy to check, the first entails that $\exists x(B_i x \wedge P_i x)$; but given the non-deterministic nature of denotation, the second does not.

Talking of the same man raises the subject of identity. One might insist that identity is a regular extensional predicate, so that its extension at a world, w, is $\{\langle x, x\rangle : x \in D \cap E^{w+}\}$. This semantics verifies substitutivity of identicals, but not the law of identity, $a = a$,

11 Technically, this can be handled by employing a map from terms to choice functions, as in Priest (1979).

which holds only for existent *a*s. This means that we cannot express some identities in the most natural way. For example, we cannot express the fact that Benny and Penny fear the same thing by $\exists x \exists y(B_i x \wedge P_i y \wedge x = y)$; for this will not be true if the object of the fear does not exist (though of course, the same thing can be expressed in other ways, e.g., by $\exists x(B_i x \wedge P_i x)$. Alternatively, though identity is hardly an intentional predicate in the standard understanding of that word, we can take it to be a special case (a logical predicate *sui generis*), and allow non-existent objects to satisfy it. In particular, we can take its extension at every world to be the usual $\{\langle x, x\rangle : x \in D\}$. (Its anti-extension may be arbitrary.) This is the simplest treatment of identity, verifies standard principles concerning identity, and is, I think, preferable.

IV. Consequences of the Account

The above account, it seems to me, has a number of virtues. First, it preserves the natural parsing of intentional predicates. Next, it accommodates the fact that there are intentional truths about non-existent objects, e.g., that the Greeks worshipped Poseidon. For example, it is easy to construct a model of $G_i p \wedge \neg Ep$. It does not allow for non-existent objects to have extensional properties. (Thus, it is not true that Holmes lived in Baker St. $B_e h$ is false on every interpretation.) That seems right: Baker Street never really had a Sherlock Holmes living in it. But the acceptability of 'fictional truths' of this kind is accommodated in another way. (In the world as Conan Doyle imagined it, Holmes lived in Baker St. It is easy enough to construct a model for: $\varphi B_e h \wedge \neg Eh$.) It also accommodates the truth of other things that we would naturally want to say about non-existent objects. For example, let *a* be any actual detective. Then it is true that Holmes is more famous than *a*. That is, more people have heard of Holmes than *a*—where, here, 'hear of' is an intentional predicate. Whilst this sentence cannot be formalised in the language we have used for illustration, due to its lack of expressive power, it is an easy matter to extend the language, and the corresponding semantics, in such a way as to allow it to be true. Details are left as an exercise.

For descriptions, $\varepsilon x A(x)$ satisfies $A(x)$ iff something satisfies $A(x)$.[12] If nothing satisfies $A(x)$, then $\varepsilon x A(x)$ still denotes; and though it may not satisfy $A(x)$, it can still do so in the way that things are imagined/believed, etc., to be. Thus, a Golden Mountain, $\varepsilon x(G_e x \wedge M_e x)$ has the property of being gold and a mountain in the world as it is told to be in the legend. If 'g' is $\varepsilon x(G_e x \wedge M_e x)$, it is easy enough to construct an interpretation where $\varphi(G_e g \wedge M_e g)$ is true.

Finally, the semantics allow for 'inconsistent and incomplete' non-existent objects. Inconsistency: one may think of a thing that is both round and not round. Let *a* be '$\varepsilon x(Rx \wedge \neg Rx)$'. It is easy to construct an interpretation which renders $\varphi(Rx \wedge \neg Rx)$ true. Incompleteness: Holmes is either right handed or left handed (in the way Conan Doyle imagined the world to be), without it being the case that he was either left handed (in that way) or right handed (in that way). It is easy enough to construct a model where $\varphi(L_e h \vee R_e h)$ is

12 Provided that x is not itself in the scope of an ε in A. The completely non-deterministic nature of the denotation of ε-terms destroys substitutivity features inside their scope. These can be preserved with suitable fine-tuning.

true, but neither $\varphi L_e h$ nor $\varphi R_e h$ is. (The denotation of h is in the extension of either $L_e h$ or $R_e h$ at every world, w, such that $@R_\varphi w$; but in some it is in the extension of L_e but not R_e, and vice versa.)

V. Objections to the Account

I will end by discussing two objections to the account. The first is easier to explain. The account accommodates naturally 'truths of fiction', and also truths about fictional objects provided that these concern intentional states (like being famous). But there would appear to be truths about fictional objects which concern extensional predicates. For example, Tolkein tells us that Bilbo Baggins, being a Hobbit, was short. Priest is 6′4″. It would therefore appear to be true that Priest is taller then Bilbo. But *being taller than* is an extensional predicate. We may accommodate the truth of this in the following way. There are numbers x and y such that x is Priest's height, in the *Hobbit,* Baggins' height is y, and $x > y$. It is easy to construct a model in which $\exists x \exists y(P_e x \wedge \varphi B_e y$ and $x > y)$ is true.[13]

Similar examples can be treated in the same way, though there may be a touch more artifice involved. For example, we might want to say that a fictional character was more angry (in some fictional context) than some actual character (in some real context). It is less natural to talk of some literal degree of anger than of some degree of height. But, it seems to me, it is quite possible, none the less (though such degrees might not form a linear order).

The second objection is harder to explain. Let us approach it gently. Consider a non-existent object, such as Holmes or Benny's man next door. What makes them different? It cannot be the fact that they have different extensional properties at this world: at this world they have no such properties. They may have different intentional properties at this world. For example, Benny fears the man next door, but not Holmes. But this cannot be what makes them different, since they must already be different things, or it would not be possible to fear the one but not the other. And in any case, different non-existent objects may have the same intentional properties at this world. For example, neither of them may be feared, imagined, or in any other way enter into anyone's intentional state.

They may have different properties at other possible worlds. For example, in the worlds which are compatible with the way Benny fears things to be, the man next door lives next door, but—at least in some of them—Holmes does not. But there are other worlds, not compatible with the way Benny fears the world to be, where Holmes lives next door, and the man next door does not.

So what makes the man next door the man next door, as opposed to Holmes? Rhetorical answer: Nothing; the account is incoherent. The problem has an analogue, of course, in standard possible world semantics. We know who Nixon is in this world, but what makes an object Nixon, and not Clinton, in some other world? The person in

[13] Actually, things are not quite so simple, since the *Hobbit* specifies no particular height for Baggins. Rather, there is a range of heights such that, in the *Hobbit*, Baggins has a height in that range, and each height in the range is less than Priest's height. But this extra complexity changes nothing essential here.

question may have no properties there in common with those possessed by Nixon in this world. They may have a different name, different job, different gender.

There are a number of possible and well known answers to this question. One is to appeal to haeccicities, individual essences.[14] Thus, Nixon has a certain essential property that makes him what he is. The property of being identical to Nixon will do (assuming there to be such a property). For Nixon, and only Nixon, has that property. We may say exactly the same about the man next door, or about Holmes.

Another possibility is to reject the question as meaningless, based on a misconceived view of the nature of possible worlds. What makes an object Nixon (or Holmes, or the man next door) in a possible world is simply that that's how we specify things. Kripke, for example, takes this line. I quote his own words:

> If we say 'If Nixon had bribed such and such a Senator, Nixon would have gotten Carswell through,' what is *given* in the very description of that situation is that it is a situation in which we are speaking of Nixon, and of Carswell, and of such and such a Senator. And there seems to be no less objection to *stipulating* that we are speaking of certain *people* than there can be objection to stipulation that we are speaking of certain *qualities*. Advocates of the other view take speaking of certain qualities as unobjectionable. They do not say, 'How do we know that this quality (in another possible world) is that of redness?' But they do find speaking of certain *people* objectionable. But I see no more reason to object in the one case than in the other. I think it really comes from the idea of possible worlds as existing out there, but very far off, viewable only through a telescope.[15]

There are certainly other possibilities for handling the situation.[16] It is not my purpose here to advocate any of these views. I simply point out that the problem in question is one for anyone who uses a possible-world semantics, whether they are dealing with existent or non-existent objects. And as far as I can see, any solution for existent objects will work equally well for non-existent objects.

It may be that this is not the case, for reasons of which I am unaware. Or it may be that there are other objections to the account I have suggested, which show it to be wrong. But in default of these, and given that the semantics seems to provide very neatly so much of what is required of an account of the properties of non-existent objects, I conclude that it is a very plausible such account.[17]

University of Queensland

Received: September 1999
Revised: June 2000

[14] See, e.g., Kaplan (1975).

[15] Kripke (1977), p.82.

[16] For example, we may give up the idea that objects exist in more than one world, and talk in terms of objects' counterparts, as does Lewis (1968).

[17] A draft of this paper was given at the 1999 meeting of the Australasian Association for Logic, held at Melbourne University. I am grateful to those present on the occasion for many helpful thoughts. I am also extremely grateful to Ed Zalta for his comments on an earlier draft of the paper.

REFERENCES

N. Griffin (1998), 'Problems in Item Theory', a paper read at the 1998 meeting of the Australasian Association for Logic.

D. Kaplan (1975), 'How to Russell a Frege-Church', *Journal of Philosophy* 72, 716–29.

S. Kripke (1977), 'Identity and Necessity', ch.2 of S.P. Schwartz (ed.), *Naming, Necessity and Natural Kinds*, Cornell University Press.

D. Lewis (1968), 'Counterpart Theory and Quantified Modal Logic', *Journal of Philosophy* 65, 113–26.

D. Nolan (1998), 'An Uneasy Marriage', a talk given at the 1998 meeting of the Australasian Association of Philosophy.

T. Parsons (1980), *Nonexistent Objects*, Yale University Press.

G. Priest (1979), 'Indefinite Descriptions', *Logique et Analyse* 85–6, 5–21.

G. Priest (1987), *In Contradiction*, Martinus Nijhoff.

G. Priest (1995), 'Multiple Denotation, Ambiguity, and the Strange Case of the Missing Amoeba', *Logique et Analyse,* 150–2, 361–73.

G. Priest (1997), 'On a Paradox of Hilbert and Bernays', *Journal of Philosophical Logic* 26, 45–56.

G. Priest (1998), 'The Trivial Object and the Non-Triviality of a Semantically Closed Theory with Descriptions', *Journal of Applied and Non-Classical Logic* 8, 171–83.

R. Routley (1980), *Exploring Meinong's Jungle and Beyond*, Philosophy Department, RSSS, Canberra.

E. Zalta (1988), *Intensional Logic and the Metaphysics of Intentionality*, MIT Press, 1988.