

THE OXFORD HANDBOOK OF

METAPHYSICS

Edited by

MICHAEL J. LOUX

AND

DEAN W. ZIMMERMAN

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi

Kuala Lumpur Madrid Melbourne Mexico City Nairobi

New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece

Guatemala Hungary Italy Japan Poland Portugal Singapore

South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© the several contributors 2003

The moral rights of the authors have been asserted
Database right Oxford University Press (maker)

First published 2003

First published in paperback 2005

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Data available

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India

Printed in Great Britain

on acid-free paper by

Antony Rowe Ltd, Chippenham, Wiltshire

ISBN 0-19-825024-X 978-0-19-825024-1

ISBN 0-19-928422-9 (Pbk.) 978-0-19-928422-1 (Pbk.)

1 3 5 7 9 10 8 6 4 2

CHAPTER 5

EXISTENCE, ONTOLOGICAL COMMITMENT, AND FICTIONAL ENTITIES

PETER VAN INWAGEN

MEINONG has famously (or notoriously) said, ‘There are objects of which it is true that there are no such objects.’¹ What could have led him to make such an extraordinary statement? He was, or so he saw matters, driven to say that there were objects of which it was true that there were no such objects by data for which only the truth of this extraordinary statement could account. These data were of two sorts: linguistic and psychological.² The linguistic data

¹ ‘...es gibt Gegenstände, von denen gilt, dass es dergleichen Gegenstände nicht gibt’ (Meinong 1969: 490). I will take it to be uncontroversial that for Meinong everything, without exception, is an ‘object’ (*Gegenstand*). I am aware that Meinong distinguished objects from ‘objectives’ (*Objective*). If, for example, the thought crosses my mind that golf is a popular sport, golf is the *Gegenstand* of my thought, and the popularity of golf is its *Objective*. But objectives are objects: if I believe that the popularity of golf is regrettable, the object of my belief is the popularity of golf. Since, at least in Meinong’s sense of the word, everything is an object—since ‘object’, in Meinong’s usage, is the most general count noun—I will take it to be uncontroversial that ‘Every object is *F*’ is equivalent to ‘Everything is *F*’ and that ‘Some object is *F*’ is equivalent to ‘Something is *F*’.

² The psychological data pertain to the phenomenon of intentionality. I will not discuss these data. One of the unstated assumptions of this chapter (unstated outside this note) is that all human

consisted of sentences like the following and what seemed to be obvious facts about them:

The Cheshire Cat spoke to Alice
The round square is an impossible object
Pegasus was the winged horse captured by Bellerophon.

The obvious facts were these: first, each of these sentences is or expresses a truth; secondly, the result of writing ‘There is no such thing as’ and then the subject of any of these sentences is, or expresses, a truth. (I so use ‘subject’ that the subject of ‘the Taj Mahal is white’ is ‘the Taj Mahal’ and not the Taj Mahal. I use ‘there is no such thing as’ to mean ‘there is no such thing as, and there never was or will be any such thing as’.) Thus, for example, it is true that the Cheshire Cat spoke to Alice, and it is also true that there is no such thing as the Cheshire Cat. We have, therefore, the following general truth:

There are true subject–predicate sentences (i.e. subject–predicate sentences that express truths when uttered in appropriate contexts) such that the result of writing ‘there is no such thing as’ and following this phrase with the subject of any of these sentences is true.

These are the linguistic data. Reflection on these data suggests the following question. The proposition expressed by the offset sentence, the proposition that summarizes the linguistic data, is a *semantical* generalization, a proposition that asserts that there are linguistic items of a certain description (‘sentence’) that possess a certain semantical property (truth); How can we express this same generalization in the ‘material mode’? How can we state it as a thesis not about the semantical properties of linguistic items but about the things those linguistic items purport to refer to? Well, strictly speaking, we *can’t* do this: ‘Rome is populous’ and ‘“Rome is populous” is true’ are not, strictly speaking, two ways of expressing the same proposition. Perhaps we should instead ask this: how can we express in a single sentence the general fact that is expressed collectively by the ‘whole’ infinite class of sentences of which the sentences

The Cheshire Cat spoke to Alice and there is no such thing as the Cheshire Cat
The round square is an impossible object and there is no such thing as the round square
Pegasus was the winged horse captured by Bellerophon and there is no such thing as Pegasus

psychological phenomena can be adequately described and accounted for without any appeal to ‘objects of which it is true that there are no such objects’.

are three representatives? (This 'single sentence' would not be a semantical sentence, for sentences of the type illustrated by our three examples are not semantical sentences; they do not ascribe semantical properties like truth or reference to linguistic items.) The sentence 'There are objects of which it is true that there are no such objects' represents an attempt at an answer to this question, but Meinong obviously recognizes that there is something unsatisfactory about this attempt, since he does not baldly say that there are objects of which it is true that there are no such objects; rather, he says, 'Those who were fond of a paradoxical mode of expression could very well say, "There are objects of which it is true that there are no such objects".' Um . . . yes—but suppose one was *not* one of those who were fond of a paradoxical mode of expression; what non-paradoxical mode of expression would one use in its place?

One obvious suggestion is: 'There are objects that do not exist'. But Meinong would object to this suggestion on grounds that are related to a peculiarity of his metaphysical terminology, for he holds that things that are not in space and time—the ideal figures the geometer studies, for example—do not 'exist' (*existieren*), but rather 'subsist' (*bestehen*), another thing entirely, or almost entirely, for subsistence is, like existence, a species of being. And this terminological red herring (in my view it is a terminological red herring) confuses matters. We had better leave the word 'exists' alone for the moment. But if we do not allow ourselves the use of the word 'exists', our question is unanswered: What shall we use in place of 'There are objects of which it is true that there are no such objects'? Perhaps we should turn to the question, What, exactly, *is* wrong with this sentence? What grounds did I have for calling it an 'extraordinary' sentence; why did Meinong suggest that this sentence was paradoxical? The answer to this question seems to *me* to be simple enough: there could not possibly be objects of which it was true that there were no such objects: if there *were* an object of which it was true that there was no such object (as it), that object would *be*; and if it *were* (if I may so phrase my point), it would not be true of it that there was no such object as it. This point is inescapable—*unless*, of course, 'there are' has (and 'es gibt' has and 'il y a' has) more than one sense. For suppose 'there are' has two senses; let the phrase itself represent one of these two senses, and let the same phrase in bold-face represent the other: there will be no contradiction in saying that there are objects of which it is true that there are no such objects. Or, at any rate, no contradiction that can be displayed by the simple argument I have just set out. (This simple argument can be phrased very neatly in the formal quantifier–variable idiom, in what Quine has called the canonical notation of quantification: 'There are objects of which it is true that there are no such objects' is equivalent to ' $\exists x$ there is no such object as x '; 'there is no such object as x ' is equivalent to 'Nothing is x ' or ' $\sim \exists y y = x$ '; 'There are objects of which it is true that there are no such objects' is therefore equivalent to ' $\exists x \sim \exists y y = x$ '; and this formula is in its turn equivalent to ' $\exists x \sim x = x$ '—that is, 'Something is not identical with itself'. The force of this argument, of course, depends on the assumption that

only one sense can be given to '∃'. For suppose this symbol is ambiguous; suppose there are two senses it might have. If we allow the symbol itself to represent one of these two senses, and 'E' to represent the other, then we are forced to admit nothing more than that 'There are objects of which it is true that there are no such objects' is equivalent either to '∃x ~ Ey y = x' or to 'Ex ~ ∃y y = x'. And this thing, we are forced to admit, is not obviously self-contradictory. To deduce an absurdity like '∃x ~ x = x' or 'Ex ~ x = x' from either of these formulae, one would have to make use of some principle that governed the relations between the 'two' existential quantifiers, some principle along the lines of '∃αFα ⊢ EαFα' or its converse, and a Meinongian is unlikely to assent to the validity of any such principle.) But it is not evident that 'there is' can plausibly be regarded as having two senses. Whether this is so is a question to which we shall return. For the moment, it seems safe to say that a strong prima facie case can be made for the logical equivalence of 'There are objects of which it is true that there are no such objects' and 'Some objects are not identical with themselves'.

Meinong, so far as I know, was not aware of the strong prima facie case for the equivalence of 'There are objects of which it is true that there are no such objects' and 'Some objects are not identical with themselves', but, as we have seen, he was obviously aware that there was something logically unsatisfactory about the former sentence. Meinong and I agree, therefore, that the sentence 'There are objects of which it is true that there are no such objects' must, in the last analysis, be replaced with some other sentence. But what sentence? Chisholm has made some suggestions:

Meinong wrote 'There are objects of which it is true that there are no such objects'. But he was well aware that this statement of his doctrine of *Aussersein* was needlessly paradoxical. Other statements were: "The non-real" is not a "mere nothing"; and 'The object as such . . . stands beyond being and non-being.' Perhaps the clearest statement was provided by Meinong's follower, Ernst Mally: '*Sosein* is independent of *Sein*.' We could paraphrase Mally's statement by saying: 'An object may have a set of characteristics whether or not it exists and whether or not it has any other kind of being.'³ (Chisholm 1972: 15)

Let us follow Chisholm and use 'the doctrine of *Aussersein*' as a name for the thesis such that the words 'There are objects of which it is true that there are no such objects' is a needlessly paradoxical formulation of that thesis. Our problem is this: How is the doctrine of *Aussersein* to be formulated without paradox? (Is this the same problem as: 'How are we to express, in a single general sentence—not a semantical generalization—the fact that is expressed collectively by the infinite class of sentences of which our three sample sentences are representatives?' The answer is a qualified yes: if, as Meinong believed, 'Pegasus' denotes an object of which it is true both that it was a winged horse and that it has no kind of being, the answer is yes; otherwise, it is no. Cf. n. 6.) It seems to me that the alternative formulations

³ '*Aussersein*' may be translated as 'independence [*sc.* of objects] of being'. '*Sosein*' may be translated as 'being-thus' or 'predication' or 'having characteristics'. '*Sein*' means 'being' (the mass term, not the count noun).

Chisholm mentions fare no better than the original. Let us first consider Chisholm's paraphrase of Mally's suggestion.

What does 'an object may' mean when it is followed by a predicate? It is clear that it is not Chisholm's intention to use 'an F may G ' to express epistemic possibility—as in 'A disgruntled employee may be the murderer'. Chisholm's use of 'an F may G ' is rather illustrated by sentences like 'A quadratic equation may have only one solution' and 'A Bengal tiger may weigh over 600 pounds'. That is to say, 'an object may G ' (in this sense) means just exactly 'some objects G '. And an object that has no kind of being must be just our old friend 'an object of which it is true that there is no such object (as it)'. (For if there *is* such an object as x , then—surely?— x must have some kind of being.) Chisholm's paraphrase, therefore, is equivalent to 'Some objects are such that there are no such objects as they'—that is to say, 'There are objects of which it is true that there are no such objects'.

Let us turn to Mally's actual words, or to Chisholm's semi-translation of his actual words: '*Sosein* is independent of *Sein*'. Suppose one said, 'Mathematical ability is independent of sex.' This *could* mean that there was no 'lawlike' connection between mathematical ability and being male or being female, a thesis logically compatible with the proposition that every mathematically able person is a woman. But, surely, Mally does not mean the doctrine of *Aussersein* to be consistent with the following statement: 'Every object of predication *in fact* is (has being), but there's no nomic necessity in that; if the course of history had gone otherwise, it might well have turned out that there were objects of predication that were not (had no being).' No, Mally's words are certainly meant to be a way of saying that the class of objects of predication that have no being is *in fact* non-empty; that is to say, his words must mean or be equivalent to 'Some objects of predication are not (have no being)'. That is to say: 'There are objects [of predication] of which it is true that there are no such objects'. The reader who agrees with what I have said so far in this paragraph will almost certainly agree with me when I say, as I do, that 'The non-real is not a mere nothing' and 'The object as such stands beyond being and non-being' are also essentially equivalent to 'There are objects of which it is true that there are no such objects'. The problem of finding a non-paradoxical expression of the doctrine of *Aussersein* is therefore so far unsolved. The problem of expressing in full generality the (non-semantical) thesis of which

The round square is an impossible object and there is no such thing as the round square

and

Pegasus was the winged horse captured by Bellerophon and there is no such thing as Pegasus

are particular cases is therefore so far unsolved.

Many philosophers would be perfectly content to say that this problem is unsolved for the same reason that the problem of trisecting the angle by Euclidean means is unsolved: it has no solution. *I* should be perfectly content to say this, at any rate.

In my view, the only generalization that has these two sentences as particular cases is a semantical generalization, something like the semantical generalization set out in the first paragraph of this chapter as a summary of the 'linguistic data.'

And here, I think, the matter would stand if the linguistic data that supported the doctrine of *Aussersein* consisted only of sentences of the form 'Fx and there is no such thing as x'. But there are other linguistic data that support the doctrine of *Aussersein*, sentences not of this form that seem to be true and whose truth seems to imply that there are objects of which it is true that there are no such objects. (And, as we shall see, it is not easy to understand these sentences as supporting only some 'harmless' semantical thesis.) The most persuasive of these data, the only ones that are really hard for the anti-Meinongian to deal with, belong to what I shall call *fictional discourse*. (I will not attempt to defend this judgement.) By fictional discourse I mean not the sentences that are contained in works of fiction but rather sentences spoken or written *about* works of fiction—whether they issue from the pen of F. R. Leavis or from the mouth of the fellow sitting beside you on the plane who is providing you with an interminable defence of his conviction that Stephen King is the greatest living novelist. The sentences of fictional discourse that I mean to call attention to are those that have the following four features: (i) they are existential quantifications, or at least look as if they were; (ii) they have complex quantificational structures (e.g. $\exists \forall \exists$)—or look as if they did; (iii) the inferences from these sentences that standard quantifier logic endorses for sentences that have the quantificational structures these sentences appear to have are valid—or at least *appear* to be; (iv) they contain not only predicates such as you and I and our friends might satisfy (predicates like 'is fat', 'is thin', 'is bald', 'is the mother of') but also 'literary' predicates like 'is a character', 'first appears in chapter 6', 'provides comic relief', 'is partly modelled on', 'is described by means of the same narrative device the author earlier used in her more successful depiction of', and so on. Here is an example:

There is a fictional character who, for every novel, either appears in that novel or is a model for a character who does.

(This sentence would express a truth if, for example, Sancho Panza served as a model for at least one character in every novel but *Don Quixote* itself.) This sentence is (i) an apparent existential quantification; (ii) complex in its apparent quantificational structure; (iv) contains literary predicates: 'is a fictional character', 'appears in', and 'is a model for'. Moreover, (iii) it certainly appears that the inferences licensed by quantifier logic for sentences with the apparent quantificational structure of the above sentence are valid. It appears, for example, that we can validly deduce from the above sentence the sentence

If no character appears in every novel, then some character is modelled on another character.

And this inference is, or appears to be, endorsed by quantifier logic, for it seems that its premiss and conclusion can be correctly translated into the quantifier–variable idiom as follows:

$$\begin{aligned} & \exists x(x \text{ is a fictional character} \ \& \ \forall y(y \text{ is a novel} \rightarrow (x \text{ appears in } y \vee \exists z(z \text{ is a} \\ & \text{fictional character} \ \& \ z \text{ appears in } y \ \& \ x \text{ is a model for } z) \)) \)) \\ & \sim \exists x(x \text{ is a fictional character} \ \& \ \forall y(y \text{ is a novel} \rightarrow x \text{ appears in } y) \) \rightarrow \exists x \exists y \\ & y \text{ is a model for } x. \end{aligned}$$

And the second sentence is a formal consequence of the first. (And the thesis that these two translations are correct does not seem to be in any way implausible or far-fetched. They certainly *look* correct. And, really, what alternative is there? Surely these translations *are* correct? Surely the inference *is* valid?) Now note a second formal consequence of the first sentence: ‘ $\exists x$ x is a fictional character’—that is to say: ‘There are fictional characters’. It seems, therefore, that the logical relations among certain sentences of fictional discourse can be accounted for only on the assumption that there are fictional characters.⁴ (It is not to the point that the first of our sentences does not express a truth. The two sentences were chosen to provide

⁴ For an extremely interesting reply to this argument, see Walton (1990: 416–19). Walton’s reply—which is in aid of his thesis that there is nothing about works of fiction that threatens to force fictional entities upon us—is very complex and resists compression. At the centre of this very complex reply is the thesis that someone who utters (assertively; in a literary discussion; in ideal circumstances) the sentence ‘If no character appears in every novel, then some character is modelled on another character’ is using language in a very different way from someone who utters (assertively and so on) the sentence ‘If no one is a citizen of every country, then someone is carrying someone else’s passport’. The latter speaker is simply making an assertion about nations and their citizens and certain of the relations that hold among these things. The former speaker is *not* making an assertion about novels and the characters that occur in them and the relations that hold among these things. He is rather engaged in a certain game of pretence. It is a part of this game of pretence that the real universe is ‘divided into realms corresponding to the various novels’, and that each realm and its inhabitants were literally created by the author of the novel to which it corresponds. In uttering the sentence ‘If no character appears in every novel, then some character is modelled on another character’, he is pretending to describe this universe and the actions and motives of the creators of its several realms. He is not, in fact, saying something that has the logical structure that ‘If no character appears in every novel, then some character is modelled on another character’ has; he is, rather, *pretending* to say something that has that logical structure (which is why he uses that sentence). My main objection to this theory—to the theory of which I have given an incomplete and inadequate account: the reader is directed to the original—is that it simply does not seem to me to be *true* that the speaker who utters ‘If no character appears in every novel, then some character is modelled on another character’ (assertively and so on) is engaged in any sort of pretence. I would assimilate his case to the case of the speaker who says ‘Some novels are longer than others’—a case of simple description of how things stand in the world if ever there was one. I would ask: is it really plausible to suppose that the speaker who says ‘Some novels contain more chapters than others’ and the speaker who says ‘Some novels contain more characters than others’ are engaged in radically different kinds of speech act? Isn’t it much more plausible to suppose that each speaker is making the same *sort* of assertion and that their assertions differ only in *content*? Isn’t it much more plausible to suppose that each speaker is simply making an assertion about the relations that hold among novels, relations that are grounded in various features of the internal structures of these novels?

an example of a formal inference that was simple but nevertheless subtle enough that the utility of quantifier logic in demonstrating its validity was evident. But there are plenty of true sentences of fictional discourse whose obvious translations into canonical notation allow the immediate deduction of ‘ $\exists x$ x is a fictional character’. One example among thousands of possible examples would be: ‘In some novels, there are important characters who are not introduced by the author till more than halfway through the work.’)

Suppose, then, that there are fictional characters—objects of thought and reference like Tom Sawyer and Mr Pickwick. If this supposition is correct, how can we avoid the conclusion that there are objects of which it is true that there are no such objects? For is it not evident that Tom Sawyer and Mr Pickwick do not exist and never did exist? And if they do not exist (I continue to respect Meinong’s attempt to distinguish between two modes of being, existence and subsistence, ill-judged though I believe it to be), there are no such things as they, for, if there are such things as they, they are human beings, and human beings can participate in being only by existing.

We have reached this conclusion, that there are fictional characters, on the basis of certain linguistic data; primarily this datum: that the first of our sentences allows the formal deduction of the second. More generally, we have argued that if the obvious logical consequences of certain sentences of fictional discourse are accounted for in what seems the only possible way, there will be (true) sentences of fictional discourse from which ‘There are fictional characters’ can be validly deduced. But could these data perhaps be interpreted semantically (following our earlier model, following the way I proposed dealing with data like the apparent truth of ‘The Cheshire Cat spoke to Alice and there is no such thing as the Cheshire Cat’), and thus rendered ‘harmless’? Could they not be given some semantical interpretation that would have no consequences about fictional characters ‘themselves,’ but only such semantical consequences as ‘There are character-*names* that occur in works of fiction that can be used in sentences of fictional discourse that express truths’? I can say only that I see no way to do this. If there is indeed no way to do this, then the data of fictional discourse I have adduced constitute stronger support for Meinongianism than the linguistic data that Meinong and his followers appeal to.

We seem, therefore, to have a strong argument for the doctrine of *Aussersein*. But, as we have seen, there is a strong argument against this doctrine, an argument that we have not seen how to deal with: The doctrine of *Aussersein* entails, or seems to entail, that something is not identical with itself—a *reductio ad absurdum* if ever there was one. But if we have not seen how to deal with this argument, we have at least mentioned in passing what will seem to many to be a promising way of dealing with it: the ‘way of the two quantifiers.’

Suppose, then, that we have two ‘existential’ quantifiers (but we must read nothing ontological into the label ‘existential quantifier’), ‘ \exists ’ and ‘ E ’. Let us propose the following two readings for these quantifiers (when we earlier touched on the

possibility of there being two existential quantifiers, we did not propose readings for them). Suppose we read 'E' as an existential quantifier whose range is restricted to those objects that participate in being, to the objects that are. And suppose we read '∃' as an existential quantifier whose range is absolutely unrestricted, whose range comprises *all* objects, even those that are not. It is easy enough to see that if we allow ourselves this distinction, and if we suppose that fictional characters fall within the range of the wider existential quantifier and do not fall within the range of the narrower, we may interpret our linguistic data in a way that entails no paradoxical consequence. Our data would support only this conclusion: '∃x x is a fictional character'. And nothing paradoxical follows from this conclusion. We cannot deduce from it either that something that lacks any sort of being (some fictional character, say) is not identical with itself, or that something that has being is not identical with itself. We can come no closer to this conclusion than what we supposed at the outset, that something that lacks being is not identical with anything that has being. And this miss is a good deal better than a mile; it is, in fact, not paradoxical at all.

These reflections on what Meinong must do if he is to state the doctrine of *Aussersein* without paradox are not very profound. I do not think it is controversial that the doctrine of *Aussersein* requires a kind of quantification that 'goes beyond being'. The important question is this: *can* there be a kind of quantification that goes beyond being? It is my contention that there cannot be, that the idea of quantifying beyond being simply does not make sense. I can hardly hope to demonstrate this to the satisfaction of the committed Meinongian, however: any argument I can present for this position must be an argument the Meinongian has already considered—or, at best, a technical refinement of an argument the Meinongian has already considered. I will do what I can, however: I will explain why the idea of quantifying beyond being does not make sense to *me*.

I begin by examining the idea of universal quantification, an idea expressed by a large variety of words and phrases, the most important of which are 'all', 'everything', and 'there is no'.⁵ More exactly, I begin by examining the idea of *unrestricted* universal quantification. (It is a commonplace of the philosophy of language that when one uses the idiom of universal quantification, one often, one perhaps usually, has some tacit restriction in mind. 'We've sold everything,' says the sales clerk after a particularly busy day behind the counter, and we who hear this assertion do not protest that the number 510, the Taj Mahal, and the counter—a concrete object right there in the shop—remain unsold.) We all, I believe understand the idea of universal quantification, and it does not require much philosophical instruction for us pass from an understanding of this idea to an understanding of the idea of unrestricted universal quantification. Now it seems to me that the idea of unrestricted universal quantification is a pellucid and wholly unambiguous idea. And it seems to me that

⁵ The argument that follows in the text is deeply influenced by Lewis (1990). But Lewis is not to be held responsible for the way I have formulated the argument.

everyone, everyone including the Meinongian and me, means the same thing by the phrase ‘unrestricted universal quantification’—although the Meinongian and I will certainly disagree about which unrestricted universal quantifications are *true*. Let us use the symbol ‘ \forall ’ to express absolutely unrestricted universal quantification (in other words, let us use this symbol in its usual sense). I say this:

$\forall x \sim x$ is a unicorn.

The Meinongian says this:

$\sim \forall x \sim x$ is a unicorn.

(In fact, the Meinongian says this is a necessary truth.) I say I don’t see how the Meinongian’s assertion could be true. The world being as it is, the Meinongian’s assertion seems to be false (if Kripke is right, necessarily false). If the Meinongian’s assertion were true—this is what I want to say—and if I were made free of all space and all time, I ought to be able to find, encounter, or observe a unicorn. But this I should *not* be able to do: no magic carpet or starship or time machine could take me to a place where there was a unicorn. The Meinongian will reply that the truth of ‘ $\sim \forall x \sim x$ is a unicorn’ does not entail the ‘findability’ of unicorns. Not everything [an absolutely unrestricted ‘everything’] is a non-unicorn, the Meinongian says—and yet unicorns are nowhere to be found. (More precisely: they *are* to be found in certain places, but I cannot visit these places because they do not exist.) Unicorns are nowhere to be found because they lack *being*. But when the Meinongian says this, I must protest that either he contradicts himself or I do not understand him. (He will no doubt respond to this protest as Chisholm once responded to a similar protest: ‘I accept the disjunction.’) In my view, on my understanding of being, each statement (after the first) in the following sequence is a consequence of—and is in fact equivalent to—the preceding statement in the sequence:

All unicorns lack being.

For every object that is a unicorn, it is true of it that there is no such object (as it).

Every unicorn is such that everything [an unrestricted ‘everything’] is not it.

$\forall x(x \text{ is a unicorn} \rightarrow \forall y \sim y = x)$.

$\forall x \sim x$ is a unicorn.

Thus, according to *my* understanding of ‘lacks being’, the Meinongian says both that all unicorns lack being and that it is false that all unicorns lack being (‘ $\sim \forall x \sim x$ is a unicorn’). It would therefore seem that—since the Meinongian obviously does not mean to embrace a straightforward formal contradiction—the Meinongian must mean something different by ‘has being’ and ‘lacks being’ from what I mean by these phrases. But what *does* he mean by them? I do not know. I say ‘*x* has being’ means ‘ $\sim \forall y \sim y = x$ ’; the Meinongian denies this. Apparently, he takes

'has being' to be a primitive, an undefinable term, whereas I think that 'has being' can be defined in terms of 'all' and 'not'. (And I take definability in terms of 'all' and 'not' to be important, because I am sure that the Meinongian means exactly what I do by 'all' and 'not'—and thus he understands what *I* mean by 'has being' and is therefore an authority on the question whether he and I mean the same.) And there the matter must rest. The Meinongian believes that 'has being' has a meaning that cannot be explained in terms of unrestricted universal quantification and negation. He therefore believes in two kinds of quantification where I believe in one. I have two quantifiers—'∀' and '∃' (that is '∼∀ ∼')—and he has four: the two I have and two others: 'A' and 'E'. These two quantifiers may be defined as follows:

$$AxFx =_{df} \forall x(x \text{ has being} \rightarrow Fx)$$

$$ExFx =_{df} \exists x(x \text{ has being} \& Fx).$$

Or so they may be defined for the benefit of someone who knows what the Meinongian means by 'has being'. But not for my benefit, for, as I have said, I do not know what the Meinongian means by 'has being'.

I therefore cannot accept the Meinongian doctrine of *Aussersein*. But what then of our strong argument *for* the doctrine of *Aussersein*?—the argument based on the data of fictional discourse? Since I do not understand the idea of objects of which it is true that there are no such objects, nothing can be (for me) an argument for the existence or reality or being (none of these is the right word, of course, but what would be the right word?) of objects of which it is true that there are no such objects. What is for the Meinongian an argument for the doctrine of *Aussersein* becomes for someone like me a *problem*: What are those of us who cannot understand objects that lack being to say about fictional discourse, which appears to be a vast repository of evidence for the [insert proper verb-stem here]-ing of such objects? We must understand fictional discourse in a way that does not presuppose the doctrine of *Aussersein*. We must adopt a non-Meinongian analysis of, or account of, or theory of, fictional discourse. And what might such a theory be?⁶ What are the available

⁶ One sort of non-Meinongian analysis of fictional discourse might make use of the idea of 'substitutional quantification'. (I take it that no one can properly say, 'I am a Meinongian because I have two sorts of existential quantifier that bind variables in nominal positions, the objectual or the referential—that's the narrow one—and the substitutional—that's the wide one.' A Meinongian, surely, is a philosopher who thinks there are two kinds of *objectual* quantifier, a wide objectual quantifier whose range comprises all objects and a narrow one whose range is restricted to the objects that are. Essentially the same point can be made in terms of reference. One does not qualify as a Meinongian in virtue of saying that the sentence 'Pegasus is a winged horse' is true despite the fact that 'Pegasus' does not refer to anything. A Meinongian must say that 'Pegasus is a winged horse' is true because 'Pegasus' refers to a winged horse, and is true despite the fact that the horse 'Pegasus' refers to does not exist.) I will not discuss analyses of fictional discourse based on substitutional quantification in this chapter. I refer the reader to van Inwagen (1981).

non-Meinongian theories of fictional discourse? This question is the subject of the remainder of this chapter.⁷

I begin with a brief exposition of a theory I have presented in various publications (van Inwagen 1997, 1983, 1985) When I have set out this theory, I will describe two other non-Meinongian theories of the ontology of fiction, those of Nicholas Wolterstorff and Amie Thomasson. Wolterstorff and Thomasson's theories are, in a sense I shall try to make clear, in substantial agreement with mine; they differ from mine in being much more specific than I care to be about the metaphysical nature of fictional characters.⁸

A non-Meinongian theory of fiction (that is, a theory of fiction that allows only one sort of existential quantifier) must answer the following question: How are we to deal with the fact (or is it a fact?) that when fictional discourse is translated into the quantifier–variable idiom, it can be seen to imply that fictional characters like Tom Sawyer and Mr Pickwick *are* or *have being*, that they *exist*? (In the remainder of this chapter, I will use 'exist' to mean the same as 'are' and 'have being', for I need no longer attend to Meinong's spurious distinction between existence and subsistence; I need no longer pretend to respect the idea that existence is one of two

⁷ For Meinongian theories of fiction, see Routley (1980, esp. ch. 7); Parsons (1975, 1980); Castañeda (1979).

Edward Zalta's 'object theory' (see n. 18) can be given a Meinongian interpretation, and 'encoding', a fundamental concept of object theory, depends to a large degree for its intuitive content on examples drawn from fiction.

I am aware of two other theories of fictional objects that *might* be described as Meinongian. According to Robert Howell (1979: 130), fictional objects are 'non-actual but well-individuated objects that exist in a variety of fictional worlds'. (Howell would repudiate the suggestion that his theory is appropriately described as Meinongian; he describes Meinongian theories as 'quasi-actualist'.) Howell does not explicitly say what a 'non-actual object' is, but it seems clear from what he says about them that they are objects that exist in non-actual worlds and do not exist in the actual world. In my view, there are no non-actual objects (in this sense)—despite the fact that there are non-actual worlds in which everything that exists in the actual world exists, and other things as well. (See Plantinga 1974, ch. vii and viii.) The theory presented in Charles Crittenden's *Unreality: The Metaphysics of Fictional Objects* (Crittenden 1991) is pretty clearly a Meinongian theory in *some* sense, but Crittenden, despite his title, adopts a resolutely anti-metaphysical attitude that leads him to avoid any attempt to give a systematic account of the nature of 'grammatical objects' (in which category he places fictional objects). What he says about the nature of grammatical objects is haphazard in the extreme, and I can't help thinking that he says things in some places that flatly contradict things he says in other places. But I am not sure of this, because I do not really understand the scattered remarks that are supposed to explain the notion of a grammatical object.

⁸ Kendall Walton's theory, the theory described in n. 4, is also a non-Meinongian theory of the ontology of fiction. But Walton's theory is wholly different from Wolterstorff's and Thomasson's and mine. Walton's theory denies the existence of fictional characters—and not in the subtle way (or the unintelligible way: take your pick) in which the Meinongian denies the existence, and even the being, of fictional characters. Walton denies the existence of fictional characters in the same straightforward sense as that in which the naturalist denies the existence of supernatural beings and the nominalist the existence of universals. The Meinongian says that 'Tom Sawyer' names something that lacks being; Wolterstorff, Thomasson, and I say that 'Tom Sawyer' names something that has being; Walton says that 'Tom Sawyer' names nothing at all.

modes of being.) I propose that we simply accept this implication. I propose that we adopt a theory according to which fictional characters exist. I propose, in fact, that the existence of fictional characters is just what our examination of fictional discourse has demonstrated. More exactly, I hold that our examination of fictional discourse has demonstrated that this follows from two assumptions: that what is said by those engaged in fictional discourse is (often) true, and that there is no way to rewrite or paraphrase the true sentences of fictional discourse so as not to allow the deduction of ‘ $\exists x$ x is a fictional character’ from the obvious and proper translations of these sentences into the ‘canonical notation of quantification’. The first of these assumptions seems obviously right: ‘In some novels there are important characters who are not introduced by the author till more than halfway through the work’ seems to be, without qualification, *true*. As to the second, it may be possible to understand sentences like ‘In some novels there are important characters who are not introduced by the author till more than halfway through the work’ in a way that allows their truth to be consistent with there being no fictional characters, but I have never been able to think of any way to do this and I have never seen any workable suggestion about how it might be done. Since, therefore, I think there are true sentences of fictional discourse (vast numbers of them, in fact) that entail ‘There are fictional characters’ (which I take to be equivalent to ‘Fictional characters exist’), and since I think one should accept the perceived logical implications of that which one believes,⁹ I conclude—tentatively, perhaps, but all philosophical conclusions should be tentative—that fictional characters exist.

The preceding paragraph and the preceding passages to which it alludes illustrate a certain style of reasoning concerning matters of existence. This importance of this style of reasoning in ontological disputes, trivial though it may seem, was not appreciated by philosophers till Quine’s very persuasive writings on ontological method forced them to attend to it. The reasons I have given for thinking that fictional characters exist are, in fact, an application of what is sometimes called ‘Quine’s criterion of ontological commitment.’¹⁰ Having said this, I must immediately record my conviction that there is an important sense in which there is no such thing as Quine’s criterion of ontological commitment. That is, there is no proposition, no *thesis*, that can be called ‘Quine’s criterion of ontological commitment’—and this despite the fact that several acute and able philosophers (see e.g. Church 1958; Cartwright 1954) have attempted to formulate, or to examine possible alternative

⁹ Unless, of course, these perceived logical implications are so incredible as to lead one to withdraw one’s assent from the proposition that has been seen to imply them. This reservation does not seem to me to apply in the present case. ‘Fictional characters exist’ does not seem to me to be so incredible that it should lead me to withdraw my assent from ‘In some novels there are important characters who are not introduced by the author until more than halfway through the work’.

¹⁰ See his classic essay ‘On What There Is’, in Quine (1961: 1-19) and *Word and Object* (Quine 1960, ch. v11). Quine soon came to prefer ‘ontic commitment’ to ‘ontological commitment’, but few philosophers have followed his example; we seem to be stuck with the more cumbersome phrase.

formulations of, 'Quine's criterion of ontological commitment.' In so far as there is anything that deserves the name 'Quine's criterion of ontological commitment,' it is a strategy or technique, not a thesis. This matter is important enough to warrant a brief digression on ontological commitment.

Strategies and techniques can be applied in various contexts. Let us concentrate on the context supplied by a debate, an ontological debate, a debate between two philosophers about what there is. Argle, let us say, contends that there are only concrete material objects. Bargle points out that Argle has asserted that there are a great many holes in this piece of cheese, and calls Argle's attention to the fact that a hole does not seem to be describable as a 'concrete material object.' I trust you know how this story goes.¹¹ It is, as its authors intended it be, a paradigm of the application of Quine's strategy. It has, however, a special feature. One of the characters in the dialogue (Bargle) is, as we might say, forcing the application of the strategy; but the other character (Argle) *cooperates*; Argle does not dispute the legitimacy of the questions Bargle puts to him. But some philosophers might not be so cooperative as Argle. Consider, for example, the late Ernest Gellner. In a review essay on Quine's contributions to philosophy Gellner gave a very nice description of Quine's ontological strategy, and, having paused briefly to identify himself as a nominalist, went on to say:

The dreadful thing is, I haven't even tried to be a serious, card-carrying nominalist. I have never tried to eliminate 'quantification' over abstract objects from my discourse. I shamelessly 'quantify over' abstractions *and* deny their existence! I do not try to put what I say into canonical notation, and do not care what the notation looks like if someone else does it for me, and do not feel in the very least bound by whatever ontic commitments such a translation may disclose. (Gellner 1979: 203)

In an ontological debate with someone like Gellner one would have to apply different strategies from those that are appropriate in a debate with someone like the admirable Argle. But I shall not further consider philosophers like Gellner. I have a lot to say to them, but I will not say it in this chapter. Here I will simply assume that Gellner's confession comes down to this: I don't mind contradicting myself—I don't mind both saying things that imply that there are abstractions (for to quantify over abstractions is *inter alia* to say things that imply that there are abstractions) and saying that there are no abstractions—if figuring out how to avoid contradicting myself would require intellectual effort.

Those philosophers who, like Argle, admit the legitimacy of Quine's strategy in ontological debate will, I think, mostly be willing to accept the following thesis: The history of ontological debates in which all parties admit the legitimacy of Quine's strategy shows that it is harder to avoid tacitly asserting the existence of things like numbers, sets, properties, propositions, and unrealized possibilities than one might

¹¹ I allude, of course, to David and Stephanie Lewis's classic paper 'Holes' (Lewis and Lewis 1970).

have thought it would be. If, for example, you think there are no numbers, you will find it difficult to recast all you want to say in the quantifier–variable idiom (and to do so in sufficient ‘depth’ that all the inferences you regard as valid will be valid according to the rules of first-order logic) without finding that the sentence

$\exists x$ x is a number

is a formal consequence of ‘all you want to say’. It may be possible in the end for you to do this—for you to ‘avoid ontological commitment to numbers’—but you will not find it a trivial undertaking.

What I have said about numbers I say about fictional characters: If you think there are no fictional characters, you will find it difficult to recast all you want to say in the quantifier–variable idiom (and to do so in sufficient ‘depth’ that all the inferences you regard as valid will be valid according to the rules of first-order logic) without finding that the sentence

$\exists x$ x is a fictional character

is a formal consequence of ‘all you want to say’. It may be possible in the end for you to do this—for you to ‘avoid ontological commitment to fictional characters’—but you will not find it a trivial undertaking. (I am inclined to think you will find it an impossible undertaking.)

It seems, therefore, that much of what we say in fictional discourse is true and that the truths of fictional discourse carry ontological commitment to fictional characters. That is to say, it seems that fictional characters exist. And, since the names that occur in works of fiction, names like ‘Mr Pickwick’ and ‘Tom Sawyer’ (when they occur not in works of fiction, but in discourse *about* works of fiction, in what I am calling fictional discourse), denote fictional characters if fictional characters are there to be denoted, Mr Pickwick and Tom Sawyer are among the things that are—an assertion that we anti-Meinongians regard as equivalent to the assertion that Mr Pickwick and Tom Sawyer are among the things that exist. (It should be noted that, at least in certain circumstances, ordinary speakers are perfectly willing to apply the word ‘exist’ to fictional characters. Consider: ‘To hear some people talk, you would think that all Dickens’s working-class characters were comic grotesques; although such characters certainly exist, there are fewer of them than is commonly supposed’; ‘Sarah just ignores those characters that don’t fit her theory of fiction. She persists in writing as if Anna Karenina, Tristram Shandy, and Mrs Dalloway simply didn’t exist.’)

There is an obvious objection to any theory of fiction that implies that fictional characters exist. It might be stated as follows. There are characters in some novels who are witches—for example, in John Updike’s *The Witches of Eastwick*. If the characters of this novel exist, therefore, it follows that witches exist—and, as we all know, witches don’t exist. For an adequate reply to this objection I must refer you elsewhere. (See van Inwagen 1977, 1983, 1985.) The essence of the reply is that we

must distinguish between those properties that fictional characters *have* and those that they *hold*.¹² Fictional characters *have* only (a) ‘logical’ or ‘high-category’ properties like existence and self-identity, (b) properties expressed by what I have called ‘literary’ predicates—being a character in a novel, being introduced in chapter 6, being a comic villainess, having been created by Mark Twain, being modelled on Sancho Panza . . .¹³ Properties that strictly entail the property ‘being human’—being a resident of Hannibal, Missouri, being an orphan who has a mysterious benefactor, being a witch—they do not have but *hold*. (Of course, if a fictional character holds the property *F*, then it *has* the literary property ‘holding the property *F*’.) It is therefore not true in, as they say, the strict and philosophical sense that any fictional characters are witches—or that any of them are human, female, or a widow who lives in Eastwick, Rhode Island. What we should say in, as they say, the philosophy room is this: some of them *hold* the properties expressed by these predicates.

But what about our firm conviction—everyone’s firm conviction—that Tom Sawyer and Mr Pickwick and Sherlock Holmes do not exist? Let us consider two cases in which someone might use the sentence ‘Sherlock Holmes does not exist’. Consider, first, a frustrated detective who says in exasperation, ‘It would take Sherlock Holmes to solve this case, and unfortunately Sherlock Holmes doesn’t exist.’ Consider, next, an amused London cop who is responding to a flustered tourist who can’t find 221B Baker Street (‘You know, Officer—where Sherlock Holmes lived’). ‘Lord bless you, sir, Sherlock Holmes doesn’t exist and never did. He’s just a chap in a story made up by someone called Conan Doyle.’ It seems to me that the first use of ‘Sherlock Holmes does not exist’ expresses the proposition

No one has all the properties the fictional character Sherlock Holmes holds (nor has anyone very many of the most salient and striking of these properties).

The second use of ‘Sherlock Holmes does not exist’ expresses—I would argue—something like the following proposition.

Your use of the name ‘Sherlock Holmes’ rests on a mistake. If you trace back the use of this name to its origin, you’ll find that it first occurs in a work of fiction, and that it was not introduced into our discourse by an ‘initial baptism’. That is, its origin lies in the fact that Conan Doyle wrote a story in which one of the characters held the property ‘being named “Sherlock Holmes”’, and we

¹² Holding, like having, is a two-place relation. In ‘Creatures of Fiction’ (van Inwagen 1977) I employed instead of this two-place relation the three-place relation ‘ascription’, a relation that holds among a character, a property, and a ‘place’ in a work of fiction. This is a technically more satisfactory device, since it allows us to represent the fact that one and the same character may be, say, unmarried in one ‘place’ (chapter 4, for example), and married in another ‘place’, such as the second half of chapter 6.

¹³ Or, rather, these are the only properties they have other than those that may be prescribed by a specific theory of the nature of fictional characters. Compare: ‘Numbers have only logical properties like self-identity and arithmetical properties like being prime or being the successor of 6’. There is no doubt a sense in which this is true, but we must recognize that a specific theory about the nature of numbers may ascribe further properties to them—like being an abstract object or being a set.

customarily refer to fictional characters by their fictional names. (That is to say: if x is a name, and if a fictional character holds the property of being named x , we customarily use x as a name of that character.) You have mistaken this story for a history or have mistaken discourse about a fictional character for discourse about an historical figure—or both.

The difference between these two examples is this: In the first example both the speaker and the audience know that Holmes is fictional and the speaker is making a comment that presupposes this knowledge in the audience; in the second, only the speaker knows that Holmes is fictional, and is, in effect, informing the audience of this fact. The lesson I mean to convey by these examples is that the non-existence of Holmes is not an ontological datum; the ontological datum is rather that we can use the *sentence* ‘Sherlock Holmes does not exist’ to say something true. (Or something false. I can imagine cases in which it was used to say something false.¹⁴) Different theories of the ontology of fiction will account for this datum in different ways. According to one ontology of fiction, the reason we can use this sentence to say something true is that ‘Sherlock Holmes’ does not denote anything. According to another, the reason is that ‘Sherlock Holmes’ denotes something non-existent. I prefer a third account, the rather more complicated account I have briefly outlined. These ontologies should be compared and evaluated not simply by seeing how well they explain our reactions to special and isolated sentences like ‘Sherlock Holmes does not exist’; they should be compared and evaluated by seeing how well they explain our reactions to the whole range of sentences we use to talk about fiction—and our ability to integrate these explanations with an acceptable philosophy of the quantifier and an acceptable general ontology.

Here, then, is a non-Meinongian theory of fictional characters, non-Meinongian in that it rejects the thesis that fictional characters lack being, and hence does not allow the deduction of the Meinongian conclusion that it is false that everything [unrestricted] has being. What other non-Meinongian theories are possible? I will attempt to categorize the possible non-Meinongian theories of fictional characters. I will begin by isolating the two central assumptions of the theory I have just set out. They are:

1. Fictional characters exist or have being. (And this despite the fact that a sentence formed by prefixing the name of a fictional character to ‘does not exist’ can often be used to express a truth.)

¹⁴ It is a hundred years in the future. Sally is being examined on her Ph.D. thesis, ‘The Detective in British Popular Fiction before the First World War’. A pompous (and ill-informed) examiner speaks as follows: ‘This thesis appears most impressive. But it is concerned largely with the appropriation by the popular imagination of a fictional detective called Sherlock Holmes. I know the popular fiction of the period well, and I’m sorry to have to tell you that Sherlock Holmes does not exist. Conan Doyle never created any such character. The author simply made him and his supposed popularity with the public up. Apparently she believed that no one on this committee would know the period well enough to expose her fraud.’

2. What appears to be the apparatus of predication in ‘fictional discourse’ is ambiguous. Sometimes it expresses actual predication, and sometimes an entirely different relation. When, for example, we say, ‘Tom Sawyer was created by Mark Twain,’ we are using the copula ‘was’ to assert that the property expressed by the predicate ‘created by Mark Twain’ actually belongs to the fictional character Tom Sawyer. When we¹⁵ say (and say truly), ‘Tom Sawyer was a boy who grew up along the banks of the Mississippi River in the 1840s,’ we use the copula ‘was’ to express that other, ‘entirely different,’ relation to which I have given the name ‘holding.’ Tom Sawyer, in other words, belongs to the extension of the property expressed by ‘created by Mark Twain.’ But if we look, or God looks, at all the members of the extension of the property expressed by ‘boy who grew up along the banks of the Mississippi River in the 1840s,’ we, or he, will not come upon Tom Sawyer. We, or he, will come upon no one but the inhabitants of man-shaped regions of space-time who are spatio-temporally related to you and me. There is a particular number of such filled, man-shaped regions (now for ever fixed), and the size of this number cannot be, and never could have been, affected by purely literary creation of the sort Mark Twain was engaged in. Tom Sawyer, therefore, does not *have* the property expressed by ‘boy who grew up along the banks of the Mississippi River in the 1840s.’ And yet he bears some intimate relation to it—a relation he does not bear to any of its ‘competitors’ (for example: the property expressed by ‘boy who grew up along the banks of the Rhine in the 1680s’). And this relation is such that when we say,

¹⁵ ‘We’ who are engaged in fictional discourse. If Mark Twain had been so artless as to include the sentence ‘Tom Sawyer was a boy who grew up along the banks of the Mississippi River in the 1840s’ in *The Adventures of Tom Sawyer*, he would not have been engaged in fictional discourse—discourse *about* fiction—and would not have expressed the proposition that we should express if we used this sentence as, say, a part of a summary of the plot of *Tom Sawyer*. If he had included this sentence in the book, it would there have been a sentence of fiction, not of fictional discourse, and would have expressed no proposition at all, for, when a writer of fiction writes a sentence (even a straightforward declarative sentence like this one), he makes no assertion. If someone had been looking over Mark Twain’s shoulder when he wrote the sentence ‘Tom Sawyer was a boy who grew up along the banks of the Mississippi River in the 1840,’ and had said, ‘That’s true’ or ‘That’s false,’ this person could only have misunderstood what Mark Twain was doing; this person must have thought that Mark Twain was writing not fiction but history. What then was Mark Twain doing (or what would he have been doing if this imaginary literary episode were actual) when he wrote ‘Tom Sawyer was a boy who grew up along the banks of the Mississippi River in the 1840s’? Well, this is a question for the philosopher of fiction. It is a question I need not answer in the present chapter, for its correct answer is irrelevant to my purposes. I will say just this: in this matter, I am a ‘Waltonian’. Despite my disagreement with Kendall Walton about the existence of characters (see n. 4), I am in general agreement with him about the nature of fiction. Mark Twain was engaged in crafting a literary object that he intended to be usable as a prop in a certain game of pretence it would authorize, an object a reader could use, in a special, technical sense, as a prop in a game in which the reader pretended to be reading a history—and to understand what a novel or story is to understand that, like a hobby-horse, its purpose is to be so used as a prop in a game of pretence. (As with most rules, there are exceptions. Authors do sometimes make assertions in works of fiction: in ‘asides to the reader,’ asides that are general observations on human life and not comments on the events in the fiction, by ‘omniscient’ third-person narrators—that is, third-person narrators who are not themselves inhabitants of the world of the story.)

'Tom Sawyer was a boy who grew up along the banks of the Mississippi River in the 1840s', we say something that is true *because* Tom Sawyer bears that relation to that property; and it is such that if anyone said, 'Tom Sawyer was a boy who grew up along the banks of the Rhine in the 1680s', what that person said would be false *because* Tom Sawyer failed to bear it to the property expressed by the predicate of that sentence. It is possible for a character to hold inconsistent properties—to have first met her only lover in London *and* to have first met her only lover in Buenos Aires, for example—but this is normally due to authorial inadvertence. A Meinongian object can be an impossible object, an object that literally *has* inconsistent properties (witness the round square), and Meinong would say that a fictional character who first met her only lover in London and who first met her only lover in Buenos Aires was an impossible object in just this sense. But for the non-Meinongian, for the philosopher who recognizes only one existential quantifier, this is not an option. The fact that an author can, by inexplicable accident or Borgesian design, compose a story one of whose characters first met her only lover in London and first met her only lover in Buenos Aires, is by itself enough to show that a non-Meinongian theory of fiction must distinguish having and holding (or must at any rate distinguish having from something that does the same work as holding). I am not able to define holding (unless to specify the role the concept plays in our talk about fiction is to define it: I mean I am unable to provide a Chisholm-style 'replacement definition' of holding, one whose definiens contains only unproblematic, perfectly clear terms), or even to find a good name for it. (My choice of the word 'hold' has *no* basis other than the familiar phrase in the wedding service, 'to have and to hold'.¹⁶) But 'holding' makes sense if fictional discourse makes sense. And fictional discourse, for the most part, makes sense.

I will regard any theory of the ontology of fictional characters that endorses both these 'central assumptions' as in fundamental agreement with mine. (I include, as part of the second assumption, the point made in note 15: 'typical' declarative sentences in works of fiction are not the vehicles of assertions made by the authors of those works.)

Three philosophers have presented theories of fiction I regard as in fundamental agreement with mine: Saul A. Kripke, in his as yet unpublished Locke Lectures (delivered in 1973), Nicholas Wolterstorff in his book *Worlds and Works of Art* (Wolterstorff 1980), and Amie Thomasson in her book *Fiction and Metaphysics* (Thomasson 1999).

Wolterstorff's and Thomasson's theories are, as I say, in fundamental agreement with mine: they endorse my two central assumptions. But Wolterstorff and Thomasson are more specific about the nature of fictional characters than I have

¹⁶ For a discussion of the inadequacies of the other word I have used for this relation, 'ascription', see van Inwagen (1977: 50–1).

been.¹⁷ (I will not discuss Kripke's Locke Lectures because they have not been published.) My own theory has nothing to say about what larger ontological categories (other than, perhaps, 'abstract object') fictional characters belong to, and I do nothing to explain 'holding' beyond giving examples and hoping for the best. Wolterstorff and Thomasson, however, say a great deal about these matters, and their theories stand in instructive opposition.¹⁸

I will first discuss Wolterstorff's theory. According to Wolterstorff, characters are *kinds*.¹⁹ They are kinds 'maximal within a work' (or, as with Sherlock Holmes and Huckleberry Finn, maximal within two or more works). I will explain this idea through a series of definitions. (They are rough and could be refined. I do not always use Wolterstorff's technical terms.) Suppose we know what it is for a proposition to be 'true in' a work of fiction.²⁰ We can agree, perhaps, that it is true in *War and Peace* that there are human beings, that some early nineteenth-century Russian aristocrats

¹⁷ Compare what is said about 'the property role' versus 'the nature of properties' in the discussion of the problem of universals in van Inwagen (1998). I have argued for the conclusion that 'the character role' is filled, and I have made only a few very abstract remarks about the nature of the things that fill it. Wolterstorff and Thomasson also argue for the conclusion that the character role is filled; they go on to present theories about the nature of the things that fill it.

¹⁸ There are two other theories, at least, that might be described as in essential agreement with mine. One is the theory of fictional objects that Edward Zalta presents as one of the fruits of 'object theory' (see Zalta 1983, 1988). According to Zalta, there are two kinds of predication, 'exemplification' and 'encoding'. Exemplification corresponds roughly to what I call 'having' and 'encoding' to what I call 'holding'. But I do not regard having and holding as two sorts of predication. In my view having is predication—and predication is predication, full stop. I regard 'holding' as a special-purpose relation peculiar to literary discourse, a relation that happens to be expressed in ordinary speech by the words that, in their primary use, express the general logical relation of predication. It must be said, however, that Zalta's 'object theory' is an immensely powerful and ambitious theory of abstract objects. If I were convinced that object theory succeeded as a general, comprehensive theory of abstract objects, I should agree that what I had called 'holding' was just encoding in application to one very special sort of abstract object—fictional characters—and was therefore a species of predication. An evaluation of the general claims of 'object theory', however, would take us far beyond the scope of this chapter.

The second theory is that presented by Nathan Salmon in Salmon (1998). Salmon's theory of fictional objects is certainly in fundamental agreement with mine. His theory of fiction, however, endorses a thesis I reject: that typical sentences contained in a work of fiction (as opposed to fictional discourse) express propositions—in almost all cases false, and necessarily false, propositions—about fictional characters. According to Salmon, the sentence "The game's afoot, Watson!" cried Holmes', written by Conan Doyle, expresses the necessarily false proposition that a certain fictional character has (not holds but has) the property of having cried out the words 'The game's afoot, Watson!' (But Salmon does not maintain that Doyle, in writing this sentence, asserted the proposition it expresses.) Cf. n. 15.

¹⁹ Kinds, according to Wolterstorff, correspond one-one to properties (and, Russellian monsters aside, a property is given by every open sentence or condition on objects, however complex), but are apparently a distinct ontological category. The kind 'Neanderthal man', for example, a kind of human being, is an abstraction, a universal—membership in it is what is universal among Neanderthals and among the members of no more inclusive class—and it is an eternal, necessarily existent Platonic object that would exist not only if there were no Neanderthals but if there were no physical world. It is intimately related to the property 'being a Neanderthal', but is numerically distinct from it and from every other property.

²⁰ This is a marvellously subtle notion. See David Lewis's marvellously subtle Lewis (1978).

spoke French better than Russian, and that there is a man named ‘Pierre Bezúkhov’. The conjunction of all the propositions true in a work we call its *world*. A proposition may *include* a kind: it does so if its truth entails that that kind has members. Thus, the proposition that some Greeks are wise includes the kinds ‘human being’, ‘Greek’, and ‘wise Greek’. A kind may *incorporate* a kind: kind *A* incorporates kind *B* if it is impossible for something of kind *A* not to be of kind *B*. Thus, ‘Greek’ incorporates ‘human being’. If a kind is included in the world of a work, and if no other kind included in that world incorporates it, it is maximal within that work—this being the term we set out to define. It is, I should think, intuitively evident that there is a kind that incorporates ‘man named “Pierre Bezúkhov”’ and is maximal within *War and Peace*.

To be a character in a given work, Wolterstorff holds, is just exactly to be a person-kind (a kind that incorporates ‘person’) maximal within that work. (Thus, since there are kinds, since kinds exist—and since person-kinds maximal within works exist—fictional characters exist.) Consider the fictional character Pierre Bezúkhov—that kind maximal within *War and Peace* that incorporates ‘man named “Pierre Bezúkhov”’. As every object must, Pierre Bezúkhov has a ‘complete’ set of properties: for every property, he has either that property or its complement. He has such properties as being self-identical, being a kind, being an eternal object, being an important character in *War and Peace*, and not being a human being. He does *not* have such properties as being human, being the son of Count Cyril Bezúkhov, or having lived in the nineteenth century. But, of course, he does *incorporate* these properties (we may say that a character incorporates a *property* if it incorporates the kind associated with that property).²¹ Wolterstorff and I, of course, mean the same thing by ‘have’, and ‘incorporation’ (this is not a term he actually uses) plays the role in his theory that ‘holding’ plays in mine. That is to say: if I

²¹ Certain conventions are on display in this paragraph: ‘A character may, without any explanation or baptismal ceremony, be referred to by a name if it incorporates the kind associated with the property of having that name’; ‘A character is normally referred to as “he” if it incorporates the kind “male”’

It should be remarked that similar conventions apply in my own theory: in particular, this one: a name can be applied to a character in critical discourse if the property of having that name is held by or ascribed to that character. (See the second paraphrase of ‘Sherlock Holmes does not exist’ in the text.) And such a name will be a full-fledged proper name. Some have apparently thought that my appeal to the existence of such a convention to explain the fact that, for example, ‘Tom Sawyer’ names a certain fictional character is inconsistent with the thesis that a full-bodied Kripkean theory of proper names applies to the names of fictional characters. It is not. It does not even appear to be, as far as I am able to judge. By way of analogy, imagine the following convention, another ‘automatic naming’ convention: in a certain culture any male baby with green eyes automatically receives, and *must* receive, the name ‘Robin’ (but females and non-green-eyed males can also be given the name ‘Robin’ if their parents are so inclined). Imagine that Sally, a woman of that culture, hears that the Smiths have had a green-eyed male baby, and proceeds, without giving any thought to what she is doing, to refer to the new arrival as ‘Robin Smith’, despite the fact that no one has performed a ceremony in which a name was conferred on the baby. It is obvious that ‘Robin Smith’ in Sally’s mouth is functioning as a proper name for the new child, and that nothing in this case contradicts Kripke’s theory of proper names.

were to decide that characters were kinds, I would also decide that holding was incorporation.

By being specific about the ontology of fictional characters, and by (in effect) replacing my vague, ostensibly explained notion of ‘holding’ with the relatively precise and explicitly defined notion of incorporation, Wolterstorff lays his theory open to many difficulties that my theory avoids (or puts itself in a position to respond to more easily) by the clever expedient of being vague. I will mention five.

First, there is the ‘Joe DiMaggio’ difficulty: persons casually referred to in a story (whether, to speak the language of everyday fictional discourse, they are real people like DiMaggio or fictional people) become characters in the story. Thus, since the world of *The Old Man and the Sea* includes (in virtue of some casual thoughts of old Santiago’s about baseball) a person-kind, maximal in that novel, that includes the kind ‘famous ball player named “Joe DiMaggio”’ (if there is such a kind as ‘Joe DiMaggio’, the kind-analogue of the property *being Joe DiMaggio*, it includes that kind, too). Thus, it follows from Wolterstorff’s theory that Joe DiMaggio is a character in *The Old Man and the Sea*. Or suppose that, as seems plausible, it is true in the world of *War and Peace* that everyone has a paternal grandfather. Then a person-kind that incorporates ‘paternal grandfather of a man named “Count Cyril Bezúkhov”’ is maximal in the novel; that is to say, one of Pierre’s great-grandfathers is a character in the novel, even though he is never explicitly mentioned in the novel.

Secondly, there is the ‘Robinson twins’ difficulty.²² Suppose a story contains a reference to ‘the Robinson twins’, but says nothing to differentiate them. Then there is *one* person-kind, maximal in the story, that incorporates the kind ‘being one of a pair of twins named “Robinson”’. That is to say, *one and only one* of the characters in the story is named ‘Robinson’ and has a twin.

Thirdly, there is the fact that, if Wolterstorff is right, every property incorporated by a character is *essentially* incorporated by that character. So, for example, Lewis Carroll could not have so arranged literary matters that the character we in fact call ‘Alice’ was asked the riddle ‘Why is a raven like a rolling-pin?’ instead of ‘Why is a raven like a writing-desk?’; a character who was asked the former riddle would have been a different character, even if the *Alice* books were changed in no other way. And, assuming that characters exist in unfinished stories, the characters in unfinished stories change (not ‘descriptively’ but ‘numerically’: they do not come to incorporate new properties—which would be impossible; rather, they vanish from the unfinished work and are replaced by distinct characters) almost every time the author adds to or revises his or her manuscript.

²² See Walton (1983, esp. 187–8). I believe it was Robert Howell who first discussed cases of this general kind. See Howell (1979).

Fourthly, there is the fact that a character who incorporates inconsistent properties (which a character certainly might; I have been told that one of the characters in *War and Peace* is in two places at once) incorporates all properties.²³

Finally, Wolterstorff's characters are eternal, necessarily existent objects. They are thus not literally created by the authors who are normally described as their creators. Mr Pickwick, for example, exists at all times and in all possible worlds. In writing *The Pickwick Papers*, Dickens perhaps caused the entity we call 'Mr Pickwick' to have the property *being a fictional character*, but he did not bring it into existence. A corollary is: the same character could, in principle, occur in causally independent works by different authors.

I now turn to Thomasson. Her theory, like Wolterstorff's, shares my two central assumptions, and, like his theory, is much more specific than mine about the nature of fictional characters. Her theory is, as one might say, the mirror image of Wolterstorff's. According to Thomasson, fictional characters are not necessary and eternal; they are, rather, in the most literal sense, created—brought into existence—by their authors. They hold the properties they hold ('hold' is my word, not Thomasson's) for the most part accidentally: 'Tom Sawyer', the very character that is in fact called 'Tom Sawyer', might have had only nine toes; Alice, she and not another, might have grown taller before she grew smaller. Fictional characters, Thomasson tells us, are 'abstract artifacts' and seem to differ from ordinary or concrete, material artefacts mainly in being abstract: their causal, temporal, and modal features are remarkably like those of concrete works of art, such as paintings and statues. When we compare Thomasson's Tom Sawyer with Wolterstorff's, the advantages seem to be all on Thomasson's side. The way she tells us is the literally correct way to talk about fictional characters is the way we do talk about fictional characters, for we talk of Mark Twain's creating Tom Sawyer, and we talk as if Tom did not exist before Mark Twain created him and would not have existed if Mark Twain had not created him. In Thomasson's view, the creativity of an author is literally creativity; in Wolterstorff's view (as in the Meinongian's), the creativity of an author is something more like an ability to find interesting regions of logical space, regions that exist independently of the author and, indeed, independently of the whole of concrete reality. Wolterstorff's authors bring us news from Plato's heaven; Thomasson's authors *make* things.

The only problem I can see with this appealing theory is that it is not at all clear that it is metaphysically possible. Can there really be abstract things that are made? Some might find it implausible to suppose that even God could literally

²³ My own theory, at least as it is presented in 'Creatures of Fiction' (van Inwagen 1977), faces the same difficulty, owing to my perhaps unwise stipulation that 'ascription' is closed under entailment. But I could have avoided the difficulty by refraining from making the stipulation, or restricting it in some way. It would not be at all easy for Wolterstorff to modify his theory in such a way that it does not have this feature.

create an abstract object. Only God can make a tree, granted, but can even God make a poem—that is, cause the object that is the poem to begin to exist? (I think it is clear that Thomasson has no special problem in this area as regards fictional characters; if an author can bring a poem or a novel—as opposed to a manuscript—into existence, there would seem to be no reason to suppose an author could not bring a character into existence.) One very plausible argument for the conclusion that it is possible to bring abstract objects into existence is provided by sets. If I can bring any objects into existence, it would seem that I can bring sets into existence. If I bring *A* and *B* into existence, then, surely, (if there are sets at all) I bring the set $\{A, \{A, B\}\}$ into existence, for the existence of that set supervenes on the existence of *A* and *B*. And that set is certainly an abstract object. (To adapt a point of Nelson Goodman's, nothing in the world of 'nominalistically acceptable things' could ground or explain the non-identity of the set $\{A, \{A, B\}\}$ with the set $\{B, \{A, B\}\}$.) Similarly, one might suppose, if the existence of characters supervenes on the existence of, say, manuscripts, and if an author can bring a manuscript into existence, then the author can thereby—*must* thereby—bring certain characters into existence.

But this analogy is not decisive. Uncontroversial examples of abstract objects other than sets (not themselves wholly uncontroversial) that can be brought into existence might be hard to find, and Thomasson's abstract artefacts are nothing like sets. (Sets have the wrong properties. If a character were, say, a set of linguistic items of some sort, characters would have properties more like those Wolterstorff's theory ascribes to characters than those Thomasson's theory ascribes to characters.) Thomasson, in fact, appeals to another sort of analogy:

... it is a common feature of many cultural and institutional entities that they can be brought into existence merely by being represented as existing. Just as marriages, contracts, and promises may be created through the performance of linguistic acts that represent them as existing, a fictional character is created by being represented in a work of literature. (Thomasson 1998)

But is it evident that when two people marry, or when a contract or a promise is made, an object called a 'marriage' or a 'promise' or a 'contract' thereby comes into existence? It seems to me to be much more plausible to say that in such cases 'all that happens' is that things already in existence acquire new properties or come to stand in new relations: the property *having promised to teach Alice to drive*, for example, or the relation *is married to*. If I say that this is 'all that happens'—that no new things come into existence in such cases—I commit myself to the thesis that quantificational sentences like 'Some marriages are happier than others' or 'A contract made under duress is not binding' can be paraphrased as sentences whose variables range only over people, properties, relations, times, and such other things as we were probably going to have to 'quantify over' in any case. It seems to me very plausible to suppose that the required paraphrases are possible. And if they are not possible, if we find that we must quantify over, say, marriages, what

would be so objectionable about regarding a marriage as an abstract object?—a set, say, that contained one man and one woman and nothing else: a ‘marriage’ would be said to be ‘in force’, or to ‘have been entered into’, or some such suitable phrase, at a moment just in the case that its members then became married to each other. I certainly have no strong tendency to believe that when Alice and James marry, a new object called ‘their marriage’ comes into existence. If no such new object does come into existence, this might be because there *never* is any such object as their marriage, or it might be because, although there is such an object as their marriage, it existed before they were married and ‘was entered into’ by them at the moment they were married. I’m not sure whether there is a ‘right’ or ‘wrong’ in these matters, and if there is, I have a hard time seeing why it would be of any great philosophical importance what was right and what was wrong. What is of some philosophical importance is this: it is not a philosophical datum that ‘many cultural and institutional entities. . . can be brought into existence merely by being represented as existing’. And we are therefore left without any strong reason for believing that it is metaphysically possible for anything to have the properties Thomasson ascribes to ‘abstract artifacts’. I hasten to add that I know of no strong reason for thinking that it is *not* possible for anything to have these properties.

Wolterstorff’s theory is unintuitive in many respects (it cannot be reconciled with many of things we are naturally inclined to believe about fictional entities), but it asks us to believe only in things that we, or the Platonists among us, were going to believe in anyway. Thomasson’s theory respects what we are naturally inclined to believe about fictional entities, but it achieves its intuitive character by, as it were, brute force: by postulating objects that have the features we are naturally inclined to think fictional entities have. The metaphysical possibility of objects having the requisite combination of properties is supported only by analogy to ‘cultural and institutional entities’, and I have given reasons for being sceptical about whether there really are any such things, and reasons for at least some uncertainty about whether, if there are, they have the properties they must have to serve as the other term in Thomasson’s analogy.

If what I have said is correct, then, although there are good reasons for believing that there are such things as fictional characters, existent objects that bear some relation that is not ‘having’ to properties like *committed suicide by throwing herself under a train* and *was born on the banks of the Mississippi in the 1830s*, the question of the metaphysical nature of these objects (whether, for example, they are eternal and necessarily existent) is very far from having been given a decisive answer.

REFERENCES

-
- Cartwright, Richard (1954). ‘Ontology and the Theory of Meaning’. *Philosophy of Science*, 21: 316–25. Repr. in Cartwright, *Philosophical Essays*. Cambridge, Mass.: MIT Press, 1987, 1–12.

- Castañeda, Hector-Neri (1979). 'Fiction and Reality: Their Relations and Connections'. *Poetics*, 8: 31–62.
- Chisholm, Roderick M. (1972). 'Beyond Being and Non-Being', in Herbert Feigl, Wilfrid Sellars, and Keith Lehrer (eds.), *New Readings in Philosophical Analysis*. New York: Appleton-Century-Crofts, 1972, 15–22.
- Church, Alonzo (1958). 'Ontological Commitment'. *Journal of Philosophy*, 55: 1008–14.
- Crittenden, Charles (1991). *Unreality: The Metaphysics of Fictional Objects*. Ithaca: Cornell University Press.
- Gellner, Ernest (1979). 'The Last Pragmatist, or the Behaviourist Platonist', in Gellner, *Spectacles and Predicaments: Essays in Social Theory*. Cambridge: Cambridge University Press, 199–208.
- Haller, Rudolf, and Rudolf Kindiger (eds.) (in collaboration with Roderick M. Chisholm) (1969–73). *Alexius Meinong Gesamtausgabe*. Graz: Akademische Druck- und Verlagsanstalt.
- Howell, Robert (1979). 'Fictional Objects: How they Are and How they Aren't'. *Poetics*, 8: 129–77.
- Laurence, Stephen, and Cynthia Macdonald (eds.) (1998). *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Basil Blackwell.
- Lewis, David (1978). 'Truth in Fiction'. *American Philosophical Quarterly*, 15: 37–46.
- (1983). *Philosophical Papers*, vol. i. New York: Oxford University Press.
- (1990). 'Noneism and Allism'. *Mind*, 99: 23–31.
- and Stephanie Lewis (1970). 'Holes'. *Australasian Journal of Philosophy*, 48: 206–12. Repr. in Lewis (1983: 3–9).
- Meinong, Alexius (1969). 'Über Gegenstandstheorie', in Haller and Kindiger (1969–73).
- Parsons, Terence (1975). 'A Meinongian Analysis of Fictional Objects'. *Grazer Philosophische Studien*, 1: 73–86.
- (1980). *Non-Existent Objects*. New Haven: Yale University Press.
- Plantinga, Alvin (1974). *The Nature of Necessity*. Oxford: Clarendon Press.
- Quine, W. V. O. (1960). *Word and Object*. Cambridge, Mass.: MIT Press.
- (1961). *From a Logical Point of View*, 2nd edn. Cambridge, Mass.: Harvard University Press.
- Routley, Richard (1980). *Exploring Meinong's Jungle and Beyond: An Investigation of Noneism and the Theory of Items*, Departmental Monograph no. 3. Canberra: Philosophy Department, Research School of Social Sciences, Australian National University.
- Salmon, Nathan (1998). 'Nonexistence'. *Nous*, 32: 277–319.
- Thomasson, Amie L. (1998). 'The Artifactual Theory of Fiction'. Paper presented at the CSLI Conference 'Empty Names, Fiction, and the Puzzles of Non-Existence', Stanford University, Mar. 1998.
- (1999). *Fiction and Metaphysics*. Cambridge: Cambridge University Press.
- van Inwagen, Peter (1977). 'Creatures of Fiction'. *American Philosophical Quarterly*, 14: 299–308. Repr. in van Inwagen (2001: 37–56).
- (1981). 'Why I don't Understand Substitutional Quantification'. *Philosophical Studies*, 39: 281–5. Repr. in van Inwagen (2001: 32–6).
- (1983). 'Fiction and Metaphysics'. *Philosophy and Literature*, 7: 67–77.
- (1985). 'Pretense and Paraphrase', in Peter J. McCormick (ed.), *The Reasons of Art*. Ottawa: University of Ottawa Press, 1985, 414–22.
- (1998). 'The Nature of Metaphysics', in Laurence and Macdonald (1998: 11–21).

- (2001). *Ontology, Identity, and Modality: Essays in Metaphysics*. Cambridge: Cambridge University Press.
- Walton, Kendall L. (1983). Review of Wolterstorff (1980). *Journal of Philosophy*, 80: 179–93.
- (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, Mass.: Harvard University Press.
- Wolterstorff, Nicholas (1980). *Worlds and Works of Art*. Oxford: Clarendon Press.
- Zalta, Edward (1983). *Abstract Objects*. Dordrecht: Reidel.
- (1988). *Intensional Logic and the Metaphysics of Intensionality*. Cambridge, Mass.: MIT Press.